

# Evidence for metabolite composition underlying consumer preference in Sub-Saharan African *Musa* spp.

Margit Drapal<sup>a</sup>, Delphine Amah<sup>b</sup>, Brigitte Uwimana<sup>c</sup>, Allan Brown<sup>d</sup>, Rony Swennen<sup>c,e</sup>, Paul D. Fraser<sup>a,\*</sup>

<sup>a</sup> School of Biological Sciences, Royal Holloway, University of London, Egham Hill, Egham, Surrey TW20 0EX, UK

<sup>b</sup> International Institute of Tropical Agriculture, Ibadan, Nigeria

<sup>c</sup> International Institute of Tropical Agriculture, Sendusu, Uganda

<sup>d</sup> International Institute of Tropical Agriculture, Arusha, Tanzania

<sup>e</sup> Department of Biosystems, KU Leuven University, W. De Croylaan 42, 3001 Leuven, Belgium

## ARTICLE INFO

### Keywords:

EAHB  
Mchare  
Plantain  
Metabolite profiling  
Phenylpropanoids  
Breeding resource

## ABSTRACT

Breeding programs for disease resistant bananas in Sub-Saharan Africa generated resistant bananas, which did not meet fruit quality characteristics preferred by consumers. The present study aimed to establish chemotypes underlying preferred cooking bananas of Matooke, Mchare and plantain and less preferred Mbidde bananas, used for beer brewing. The metabolite data of Mbidde showed higher levels of metabolites associated with sour, sweet, and astringent taste; as well as different textural properties associated with cell wall composition and lignin content. Significant differences in the majority of specialised and primary metabolites were observed in the pulp of cooking banana groups. Analysis of peel tissue indicated similar metabolic differences in the protective layer surrounding the pulp and suggested a distinct genetic regulation of phenylpropanoid and flavonoid pathways between the genome groups. In summary, the present data can be used to establish metabolic traits associated with consumer preference, which can augment modern breeding programs.

## 1. Introduction

Bananas contribute greatly to food security in many regions and are a source of income for millions of people in Sub-Saharan Africa, which is the secondary centre for banana diversification (Tenkouano et al., 2019). The main genotypes grown in this region are *Musa acuminata* derived diploid Mchare (AA) and triploids Mbidde and Matooke (AAA-EA) and plantain (AAB). Matooke and its closely related beer banana, known as Mbidde, are clone sets of Mutika-Lujugira, often referred to as East African Highland bananas (EAHB) (Perrier et al., 2019). Among the most important bananas grown in this region, are those used for cooking. Mbidde is an exception and primarily used for beer brewing. All of these bananas are harvested at green mature stage, representing the most starch-rich stage of banana ripening (Akankwasa et al., 2021; Ndayihanzamaso et al., 2020). The estimated banana consumption in East Africa exceeds 200 kg per capita annually and is mainly produced by small-scale farms (Perrier et al., 2019).

Bananas are susceptible to a wide range of diseases e.g. banana weevil, burrowing nematodes, black sigatoka, bacterial wilt and Fusarium wilt, which decrease the yield and/or quality of banana fruits (Ndayihanzamaso et al., 2020). Fusarium wilt, widely known as Panama disease, is probably the most well-known banana disease, as it led to the total collapse of “Gros Michel” in the 1960s (Ploetz, 1994). Present in most banana-growing regions, the disease is caused by an infection of *Fusarium oxysporum* f. sp. *cubense* and affects a wide range of accessions regardless of their genomic composition (Ndayihanzamaso et al., 2020). Another prominent fungus is *Pseudocercospora fijiensis* (sexual morph: *Mycosphaerella fijiensis*) the causal agent of the foliar disease black sigatoka (Tenkouano et al., 2019). Hence, the National Agricultural Research Organization (NARO) of Uganda and the International Institute of Tropical Agriculture (IITA) started developing breeding strategies in the 1990s for disease resistant EAHB and plantain hybrids in Uganda and Nigeria, respectively. Unfortunately, the plantain hybrids with improved resistance traits showed unfavourable texture properties

**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.

\* Corresponding author.

E-mail address: [P.Fraser@rhul.ac.uk](mailto:P.Fraser@rhul.ac.uk) (P.D. Fraser).

<https://doi.org/10.1016/j.foodchem.2023.137481>

Received 16 May 2023; Received in revised form 21 July 2023; Accepted 11 September 2023

Available online 15 September 2023

0308-8146/© 2023 Published by Elsevier Ltd.

(e.g. softer fruit) compared to the parental landraces (Ferris et al., 1999). The IITA-NARO breeding program focused on Matooke breeding and uses Matooke germplasm, several of the new hybrids are considered to be Mbidde bananas. Mbidde bananas are highly astringent, a sensation traditionally linked to high tannin content (Bashmil et al., 2021; Wuyts et al., 2006). The majority (~90 %) of the developed plantain and Matooke hybrids are not selected by breeders for on-farm testing and rejected by the farmers and consumers due to taste, aroma and visual characteristics, akin to the situation found with Mbidde accessions (Akankwasa et al., 2021; Tumuhimbise et al., 2019). Hence, only a very small fraction of developed hybrids are accepted by consumers (Madalla et al., 2022).

Tannins and lignin, responsible for astringency and texture properties, are biopolymers of phenylpropanoid derived compounds (further referred to as phenylpropanoids) and part of the plant defence mechanism against biotic and abiotic stresses (Wuyts et al., 2006). Additionally, phenylpropanoids act as antioxidants with many health promoting properties in the human diet and are naturally converted to aroma volatiles in the banana tissue (Sidhu & Zafar, 2018; Tsamo et al., 2015). During banana ripening, several gradual degradation processes have been observed, including the conversion of starch to sugars and changes of organic acids and flavonoids content, and lead to the sweet soft pulp of ripe bananas (Nascimento et al., 2019; Vermeir et al., 2009). Previous studies measured the antioxidant capacity of banana and established that antioxidant activity of metabolites is best established through liquid chromatography-mass spectrometry (LC-MS) characterisation of phenylpropanoids present in the sample (Bashmil et al., 2021). Hence, the present study utilised LC-MS analysis previously established for banana samples including method optimisation and a database comprising the concentration range of metabolites annotated in banana leaf, peel and pulp (Drapal et al., 2020).

The present study elucidated components related to banana fruit quality through metabolomics approaches. Two analytical platforms were used to observe a broad range of hydrophilic metabolites in pulp and peel tissue of bananas and plantains. The resulting dataset was used to assess (i) the inter- and intra-genome group heterogeneity, (ii) the metabolic differences of Mbidde accessions compared to edible accessions of Matooke and (iii) compared to Mchare and plantain.

## 2. Material and methods

### 2.1. Plant material

*Musa* spp. used in this study were cultivated at the research farms of IITA (Table S1). Mchare genotypes (AA) were cultivated at the IITA breeding station in Arusha (Tanzania), plantain genotypes (AAB) at the IITA banana station in Ibadan (Nigeria), Matooke genotypes (AAA-EA) at the IITA breeding station in Sendusu (Uganda) and Mbidde (AAA-EA) samples were outsourced from the National Agricultural Research Organization (NARO) in Uganda. Plants were grown under standard field conditions at a spacing of 3 × 2 m. Pisang lilin (ITC) and AACv Rose were cultivated in Arusha, Ibadan and Sendusu to account for environmental effects on the metabolism. Banana fruits were harvested at the mature green stage when fruits were completely filled. Peel and pulp were separated, immediately frozen in liquid nitrogen and lyophilised. The lyophilised samples were sent to Royal Holloway University of London in three instalments (Table S1).

### 2.2. Metabolite analysis

The samples were analysed as three sample sets. Accessions of set A and B, representative for the genotype groups plantain, Matooke, Mchare and Mbidde, were analysed again in set B and C (Table S1). Each sample was ground to a fine powder and randomised into sample batches of 23 samples and one extraction blank. The sample batches were extracted with methanol/water/chloroform and analysed, as

previously described (Price et al., 2016). Aliquots of the polar extraction phase were used for gas chromatography mass spectrometry (GC-MS; 150 µL) and liquid chromatography mass spectrometry (LC-MS; 100 µL). Internal standards D<sub>4</sub>-succinic acid (10 µg/sample) and genistein (2.5 µg/sample) were added for the respective analysis platform.

Analysis was performed with a 7890A gas chromatography (GC) system coupled with a mass spectrometer (MS) 5795C MSD (Agilent Technologies, Inc.) and a 6560 Ion Mobility Q-TOF coupled to a 1290 Infinity II (Agilent Technologies, Inc.) as previously described (Drapal et al., 2020).

LC-MS data was converted into a tabulated format with Agilent MassHunter Profinder 10.0 (Agilent Inc.). Settings included a retention time (RT) restriction 0–8.5 min and a peak filter > 1500 counts. RT and mass tolerances were 0.2 min and 20 ppm ± 2 mDa and molecular features needs to be present in at least 30 % of all sample files. Integration algorithm “ChemStation” with skim valley ratio 20 and peak-to-valley ratio 1000 were used.

AMDIS 2.71 (NIST) was used for peak picking, deconvolution and identification of GC-MS data. Metabolites were identified through authentic standards and NIST<sup>®</sup> 17 MS library (National Institute of Standards and Technology, USA).

### 2.3. Lignin analysis

An aliquot (180 mg) of the finely ground powder was used to prepare a protein-free cell wall extract for three selected accessions of Matooke, Mchare and plantain and for five accessions of Mbidde. The dried cell wall extracts were used for lignin quantification with the optimised acetyl bromide assay, previously published by Moreira-Vilar et al. (2014). A standard curve was created with a lignin standard (Merck, CAS: 8068-05-1).

### 2.4. Data processing and statistical analysis

All samples were normalised by the internal standard and sample weight. The accessions of set A and B were then normalised to set C with the accessions that were repeatedly analysed in set C. The data were then normalised by Pisang lilin and AACv Rose grown in all three locations. The final tabulated data set for GC-MS and LC-MS contained metabolites or molecular features present in > 80 % of the samples, which were averaged for each accession.

Statistical analysis for genome group comparisons was performed with XLSTAT (Addinsoftware). ANOVA was applied for one-way and two-way comparisons, including Tukey HSD and Benjamini-Hochberg post-hoc for the latter. Statistically significant values were used to perform metabolic pathway enrichment analysis in MetaboAnalyst 5.0 (Pang et al., 2021) to identify molecular features. MetaboAnalyst was also used to create volcano plots and establish molecular features significantly different with a 1.5-fold change in Mbidde compared to Matooke, Mchare and plantain. Bar charts were created with GraphPad Prism 9.4.1 (GraphPad Software, LLC).

## 3. Results

### 3.1. Metabolite composition of banana pulp tissue and the variation between and among four banana types

Metabolite profiling was performed on banana pulp of 10 Mchare, five Mbidde, 17 Matooke and 21 plantain accessions from IITA (Table S1). The accessions of the present study were grown in three different locations (Nigeria, Tanzania and Uganda). The untargeted analysis of the polar extracts of pulp tissue detected ~ 1300 molecular features including 53 unique features distributed between the banana groups. A feature was established as unique in the present study, if the feature was detected in all biological and technical replicates of one banana group and could not be detected in at least 50 % or > 75 % of

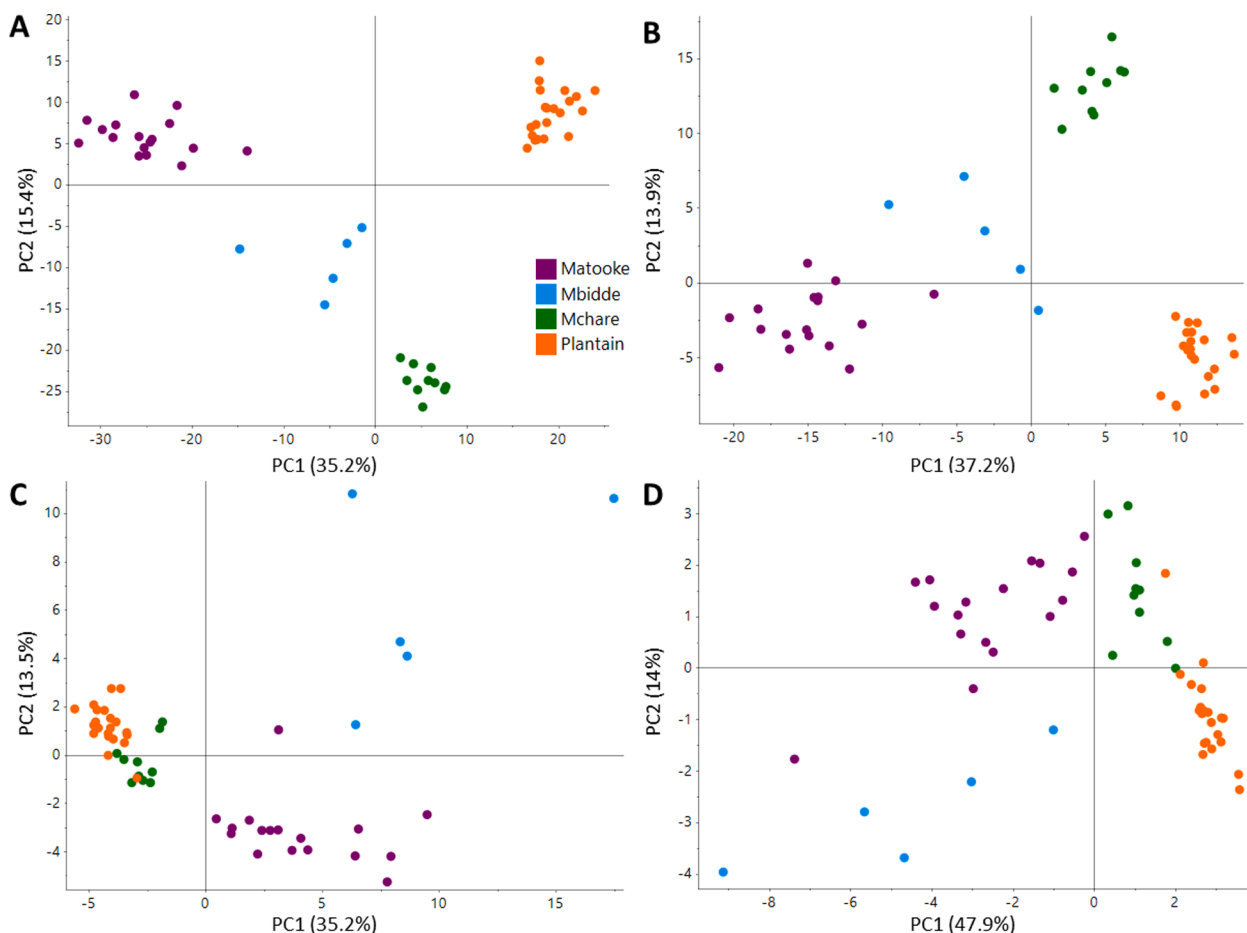
samples in another group (Table S2). Half of the molecular features were unique to Matooke and the least number of unique features (five) were present in Mchare. Ten unique molecular features could be putatively identified as glucosides of anthocyanidins, catechin and other unknown compounds.

To account for any environmental influence on the metabolome or variation in sample processing methods, two genotypes (Pisang lilin and AACv Rose) were grown in all three locations and used to normalise the dataset as part of the data pre-processing. The resulting data set, used for data mining and statistical analysis, only included compounds detected in the control plants (Pisang lilin and AACv Rose) and in at least 30 % of the accessions of the panel. The untargeted approach by liquid chromatography mass spectrometry (LC-MS) detected over 1000 molecular features and analysis by gas chromatography mass spectrometry (GC-MS) identified 75 metabolites including sugars, amino acids, intermediates of the TCA cycle and other hydrophilic metabolites (Table S3). The data was displayed as PCA score plots (Fig. 1A and C) and showed a clear separation of the four banana types Mchare, Mbidde, Matooke, and plantain for the LC-MS data (Fig. 1A), whereas the GC-MS data indicated a closer similarity of Mchare and plantain accessions (Fig. 1C). In both plots, the separation was explained by approximately half of the data set represented by the first two principal components. Similar results were detected for the peel tissue of Mbidde, Matooke and plantain accessions (Fig. S1).

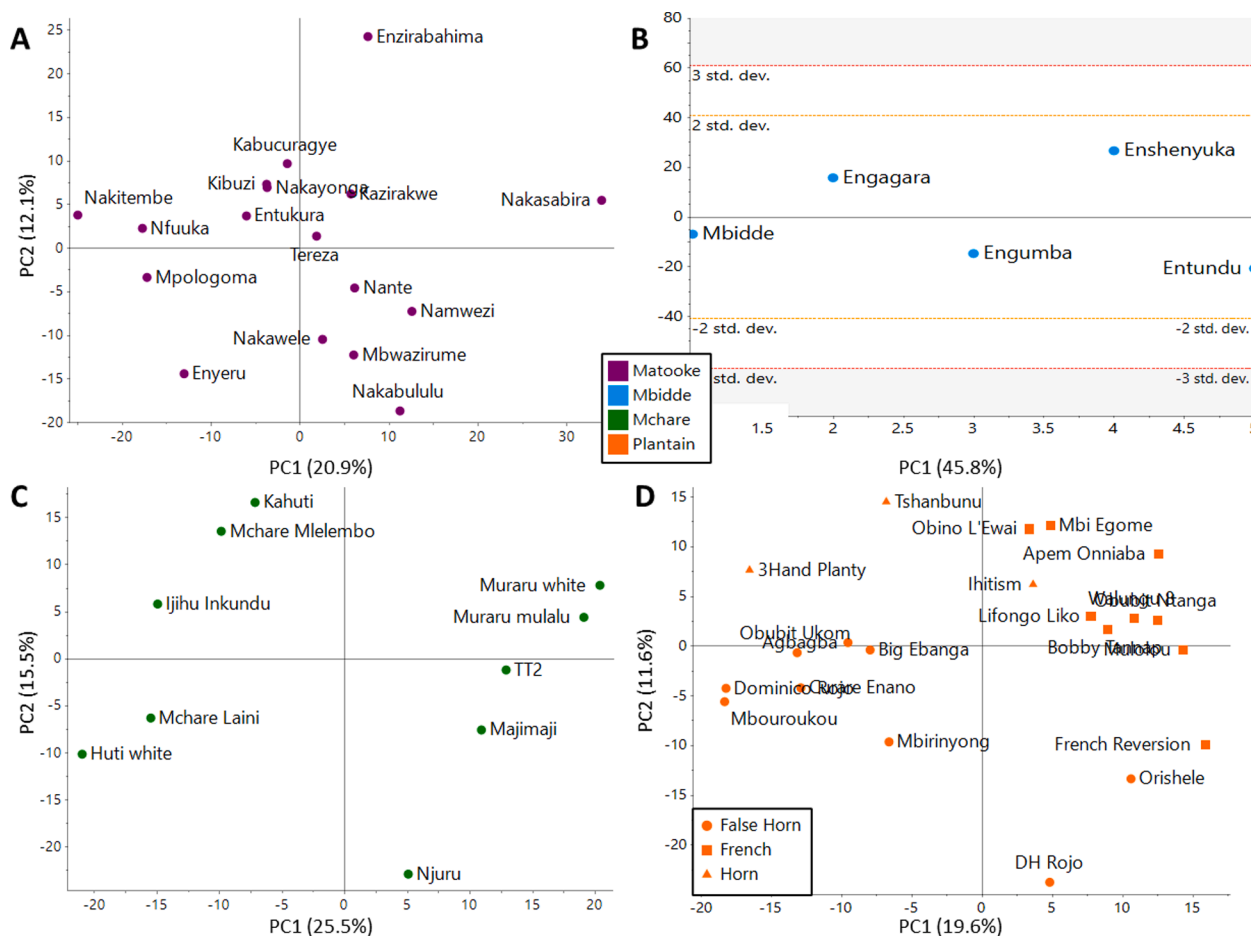
Approximately a third of the detected molecular features and identified metabolites were present in all accessions analysed. PCA score plots, based solely on these selected data, were created to visualise the similarities/differences of the genome groups and banana types. The PCA for selected LC-MS data (Fig. 1B) showed a similar pattern to the

score plot including all molecular features (Fig. 1A). The PCA score plot of selected GC-MS data (Fig. 1D) showed a closer grouping of Mchare, Mbidde, Matooke and plantain compared to the analysis with all identified metabolites (Fig. 1B) and a higher representation (62 %) of the metabolites in the first two principal components. This indicates that the metabolites present in all accessions were present in similar concentrations, followed similar patterns and showed less variation.

The metabolic variability within Mchare, Mbidde, Matooke and plantain was assessed through PCA (Fig. 2). Each of the score plots indicated that the first two principal components comprised ~ 35 % of the molecular features detected by LC-MS. Mbidde accessions showed the lowest metabolic variability. This was visualised in score plot (Fig. 2B) through separation of the accessions along only one principal component, representing 45.8 % of the molecular features. The selection for plantain accessions included genotypes from three subsets, False Horn, French and Horn. No separation between the three subsets could be observed in the PCA score plot. PLS-DA (Fig. 2D) showed that a separation between the three subsets can be forced and indicated that False Horn and French are distinguishable from each other. The score plots of the GC-MS data (Fig. S2) showed 32 % and over 52 % of the metabolites were represented by the first two principal components for plantain and all other groups, respectively. Plantain accessions could not be separated by supervised methods PLS-DA or OPLS-DA based on GC-MS data representing primary metabolism. Contrary to the LC-MS data, Mbidde accessions were separated on at least two principal components (Fig. S2B), explaining 56 % of the variation. Engagara and Enshenyuka were located in different quadrants compared to the other three Mbidde accessions. The same analysis of peel tissue of Mbidde, Matooke and plantain (Fig. S3) showed the same results with a higher



**Fig. 1.** Metabolite profiling of pulp tissue of *Musa* accessions with LC-MS (A, B) and GC-MS (C, D). Data is displayed as PCA score plots and includes (A, C) all metabolites/molecular features of the processed data table or (B, D) only metabolites/molecular features detected in all accessions analysed.



**Fig. 2.** Principal component analysis of banana pulp of Matooke (A), Mbidde (B), Mchare (C) and PLS-DA of plantain (D). Data visualised as score plots, comprises molecular features detected in the polar extracts by LC-MS. Subsets False Horn, French and Horn for plantain accessions are indicated by different shapes (see legend).

percentage of the data represented by the first two principal components in the GC-MS data compared to the LC-MS data.

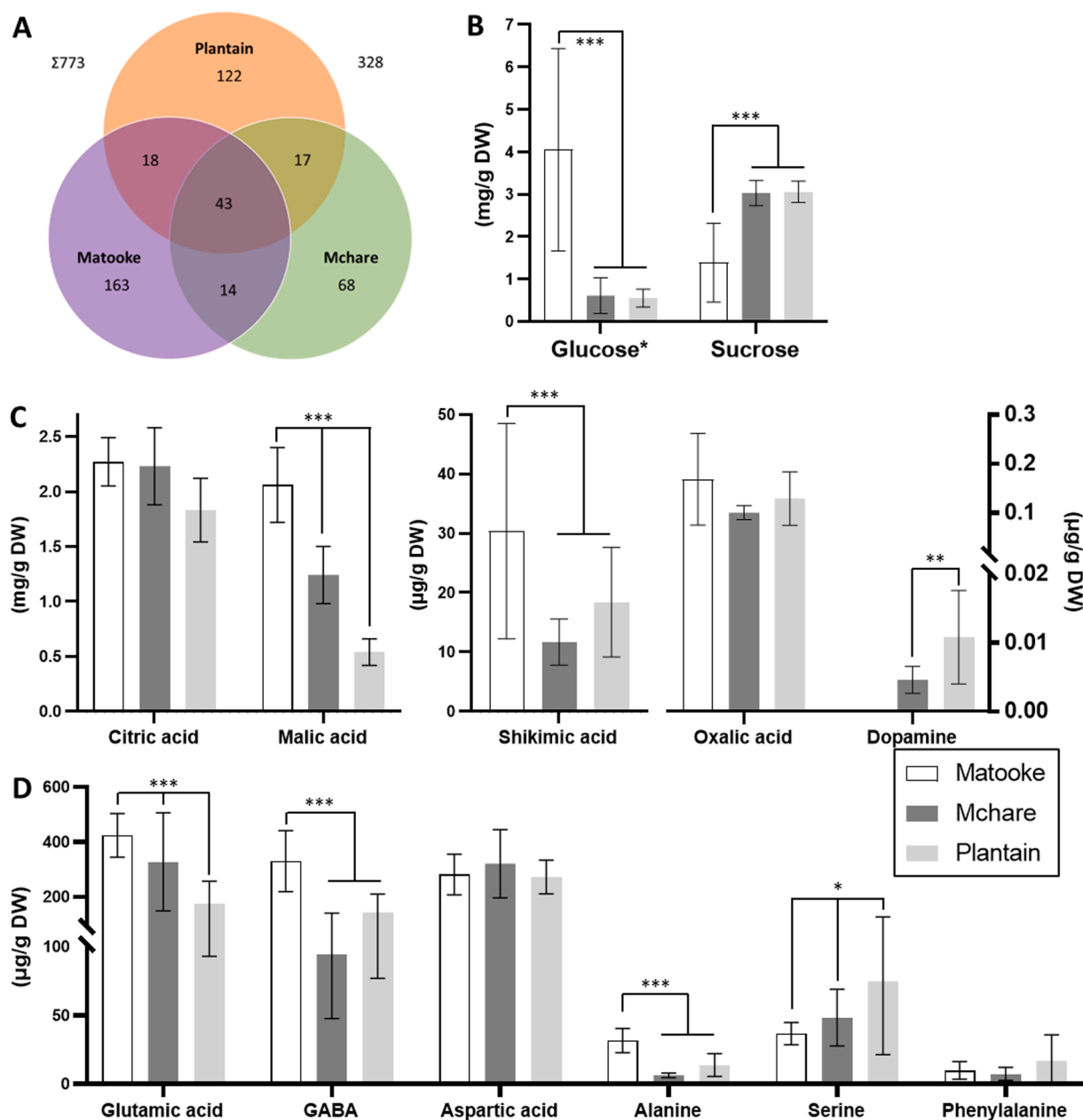
### 3.2. Similarities and differences between pulp composition of Matooke, Mchare and plantain

Statistical analysis of the metabolite profiling data of pulp tissue of Matooke, Mchare and plantain observed 730 of 773 molecular features as significantly different (Fig. 3A). 328 of these molecular features were significantly different between all three genome groups and included e. g. aspartic acid, afzelechins, quercetin glucosides, ferulic acid and feruloyl malic acid. <7 % of molecular features were only significantly different in one-to-one comparisons of the three genome groups. Furthermore, the data showed that 46 % of the molecular features were only significantly different in one compared to the other two genome groups. Matooke displayed the highest amount of differences (163), followed by plantain (122) and Mchare (68). Identification of molecular features highlighted different compositions of phenylpropanoids. Quantities of the metabolites were established relative to the internal standard and the percentage of each metabolite was calculated to establish the most prominent phenylpropanoid pathways (Fig. S4, Table 1). Matooke accessions had similar levels of quercetin, afzelechin and epiafzelechin, followed by kaempferol; and a similar amount of their glycosylation products, including quercetin glucosides. Plantain accessions also had a similar content of free flavonoids with kaempferol representing 30 %; whereas quercetin was the most prominent flavonoid (68 %) in Mchare, followed by kaempferol with 14 %. In Mchare and plantain accessions, two quercetin diglucosides constituted for > 90 %

of the observed glycosylated flavonoids. A quantitative comparison of these metabolites between the three genome groups highlighted that the content of phenylpropanoids varied between the three genome groups. However, no trends with regards to metabolic regulation of specific pathways or compounds could be established from the quantitative differences between the genome groups.

For the GC-MS data, metabolites identified in pulp tissue of Matooke, Mchare and plantain accessions were displayed as bar charts (Fig. 3B-D). Some of the metabolites present in all accessions, e.g. phenylalanine, aspartic acid, citric acid and oxalic acid, were not significantly different between the genome groups. Other metabolites, e. g. sucrose, glutamic acid and serine, showed a significant difference between genome groups and a large standard deviation (30–70 %) within the genome groups. A few metabolites, e.g. glucose, malic acid, alanine and GABA, were significantly different and showed a reduced overlap of error bars. Glucose, which represented the combined levels of glucose, fructose, galactose and mannose, was ~ 4-fold higher in Matooke compared to Mchare and plantain. Alanine followed the same trend with ~ 4-fold higher levels in Matooke. Malic acid was detected in the highest levels in Matooke, followed by Mchare and plantain. Matooke had ~ 1.1-fold higher levels than Mchare and ~ 3.8-fold higher than plantain. For GABA and glutamic acid, the standard deviation was calculated as 20–54 % for the three genome groups. Matooke had ~ 2.4-fold higher levels of glutamic acid compared to plantain; and ~ 3.5- and 2.3-fold higher levels of GABA compared to Mchare and plantain, respectively. Dopamine was not detected in Matooke accessions and was significantly higher (1.5-fold) in plantain than Mchare accessions.

The ANOVA results of peel tissue of Matooke and plantain accessions



**Fig. 3.** Differences of molecular features (A) and identified metabolites (B-D) in banana pulp of Matooke, Mchare and plantain. (A) ANOVA was performed on molecular features detected by LC-MS in Matooke, Mchare and plantain. Total number ( $\Sigma$ ) and features significantly different between all three genome groups are stated beside the Venn diagram. Features are grouped as no difference between all three genome groups (middle of the diagram), statistically significant in only two genome groups (single overlap) or in only one genome group (listed beside the genome group label). (B-D) Concentrations of selected sugars, TCA cycle intermediates and amino acids are displayed as bar charts. Glucose\* represents the combined levels of glucose, fructose, galactose and mannose detected by GC-MS. Statistically significant changes are displayed as  $P < 0.05$  (\*),  $<0.01$  (\*\*),  $<0.001$  (\*\*\*).

showed that over two thirds of the significantly different features/metabolites in the pulp tissue, were also significantly different in peel tissue. These metabolites followed two trends: i) similar ratio between Matooke and plantain as observed in pulp or ii) opposite trend e.g. citric acid was higher in pulp and lower in peel of Matooke. Metabolites following the first trend included free and glycosylated flavonoids (e.g. quercetin, rutin) and free and esterified hydroxycinnamic acids (e.g. ferulic acid, coumaroyl hexoside). Metabolites following the second trend included the majority of amino acids and intermediates of the TCA cycle. Intermediates of the nitrogen shunt were only significantly different in pulp tissue with a higher content in Matooke accessions.

### 3.3. Elucidation of distinctive metabolites associated with *Mutika/Lujugira* subgroup *Mbidde* in pulp

The metabolic composition of Mbidde accessions were compared within the *Mutika/Lujugira* subgroup to Matooke and to two other genome groups Mchare and plantain (Fig. 4). Analysis included OPLS-DA and volcano plot to establish the significantly different molecular features and metabolites. For OPLS-DA, both comparisons showed a distinct separation of Mbidde accessions and a higher variation within the groups than between the groups (Fig. 4A and B). The volcano plot highlighted ~ 40 % of the molecular features detected by LC-MS were significantly different in Mbidde and had a fold change higher than 1.5. The comparison to Mchare and plantain indicated a similar amount of

Table 1

Metabolites significantly different in pulp tissue of Mchare, Matooke and plantain compared to Mbidde accessions. Metabolite levels are listed as average  $\pm$  standard deviation ( $\mu\text{g/g}$  dry wt.). The metabolites were significantly different between Mbidde and Matooke (bold) or Mbidde and Mchare/plantain (underlined). N.d. indicates that the metabolite was below the limit of detection. \* listed as the sum of mono- or di-saccharides detected by GC–MS.

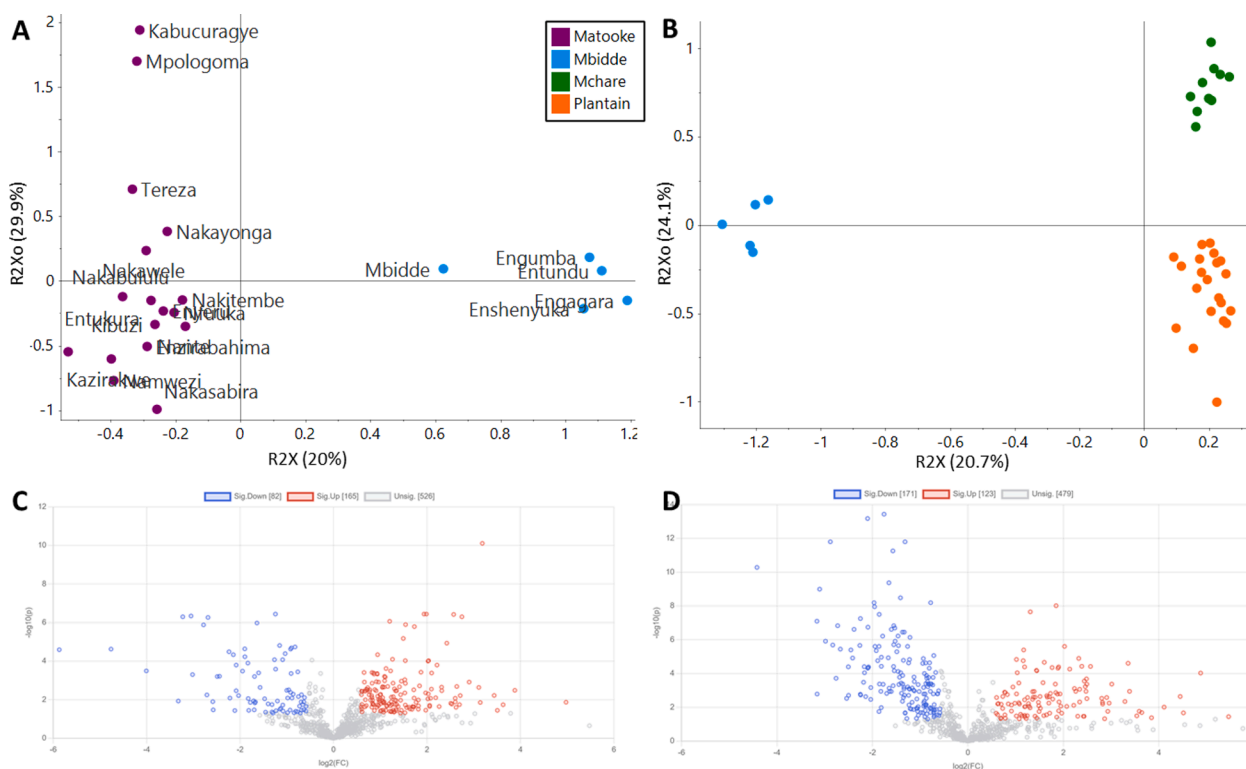
Metabolite	Mbidde	Matooke	Mchare	Plantain
Monosaccharide *	<b>9425.7 <math>\pm</math> 3543.8</b>	4046.9 $\pm$ 2387.9	611.4 $\pm$ 415.5	552.4 $\pm$ 211.4
Disaccharide *	<u>1936.6 <math>\pm</math> 1256.6</u>	1392.9 $\pm$ 928.2	3032.6 $\pm$ 296.3	3063.1 $\pm$ 249.8
<i>TCA cycle</i>				
Pyruvic acid	<b>12.7 <math>\pm</math> 7.38</b>	28 $\pm$ 14.2	21 $\pm$ 9.91	25.2 $\pm$ 18.1
Citric acid	<b>1023.2 <math>\pm</math> 190.7</b>	680.3 $\pm$ 260.8	971.2 $\pm$ 81.6	900.1 $\pm$ 236.1
<i>cis</i> -Aconitic acid	<b>481.6 <math>\pm</math> 151.5</b>	652.3 $\pm$ 129	552.8 $\pm$ 114.7	346.3 $\pm$ 83.3
Isocitric acid	<b>23621.2 <math>\pm</math> 4439.2</b>	1443.9 $\pm$ 5953.4	15119.2 $\pm$ 5960.2	14093.5 $\pm$ 7130.6
Fumaric acid	<b>260.4 <math>\pm</math> 248.4</b>	445.2 $\pm$ 112.6	242.4 $\pm$ 49.3	152.1 $\pm$ 43.1
Malic acid	<u>14624 <math>\pm</math> 6163.9</u>	16500 $\pm$ 2689.3	9939.3 $\pm$ 2091.4	4327.4 $\pm$ 935.9
<i>Amino acids</i>				
Leucine	<b>169.4 <math>\pm</math> 98.8</b>	39.1 $\pm$ 52.6	n.d.	7.52 $\pm$ 13.2
Valine	<u>235.3 <math>\pm</math> 89.4</u>	131.6 $\pm$ 55.1	44.7 $\pm$ 18	117.9 $\pm$ 86.9
Alanine	<u>311.4 <math>\pm</math> 212.8</u>	253.3 $\pm$ 70.6	50.9 $\pm$ 13.9	104.8 $\pm$ 66.1
Glycine	<u>222.2 <math>\pm</math> 104.2</u>	178.3 $\pm$ 89.3	28.2 $\pm$ 13.8	38 $\pm$ 31.3
Aspartic acid	<u>2.51 <math>\pm</math> 1.15</u>	3.73 $\pm$ 0.77	8.09 $\pm$ 1.84	9.49 $\pm$ 1.17
Lysine	<b>191.3 <math>\pm</math> 185.1</b>	84.6 $\pm$ 30.5	1.41 $\pm$ 4.47	15.2 $\pm$ 17.8
Asparagine	<b>346.8 <math>\pm</math> 417.9</b>	89.2 $\pm$ 34.6	29.2 $\pm$ 89.9	17.7 $\pm$ 16.7
Isoleucine	<b>85.3 <math>\pm</math> 32.1</b>	39.7 $\pm$ 36.7	2.15 $\pm$ 4.54	12.4 $\pm$ 13.6
Glutamic acid	<b>1996.7 <math>\pm</math> 808.7</b>	3308.9 $\pm$ 639.9	2325.6 $\pm$ 1278.5	1242.2 $\pm$ 622.8
GABA	<b>1399.3 <math>\pm</math> 384.5</b>	2640.7 $\pm$ 883.7	753.9 $\pm$ 372.7	1095.5 $\pm$ 535.6
Ornithine	<b>81.6 <math>\pm</math> 156.2</b>	n.d.	n.d.	31.6 $\pm$ 30
Proline	<u>169.2 <math>\pm</math> 63.4</u>	147.4 $\pm$ 51.6	n.d.	n.d.
Tryptophan	<u>11.2 <math>\pm</math> 4.07</u>	3.0 $\pm$ 1.83	1.29 $\pm$ 0.99	4.8 $\pm$ 4.02
Phenylalanine	<u>20.2 <math>\pm</math> 8.46</u>	5.67 $\pm$ 3.1	7.43 $\pm$ 5.13	8.18 $\pm$ 7.18
<i>Shikimic acid</i>	<b>14.6 <math>\pm</math> 3.15</b>	30.4 $\pm$ 18.1	11.7 $\pm$ 3.88	18.4 $\pm$ 9.2
<i>Phenolic acids</i>				
<i>trans</i> -Caffeic acid	<b>14.7 <math>\pm</math> 9.95</b>	0.69 $\pm$ 2.84	1.27 $\pm$ 2.7	0.49 $\pm$ 1.55
Ferulic acid	<b>5.41 <math>\pm</math> 0.48</b>	8.25 $\pm$ 2.33	4.69 $\pm$ 0.29	5.65 $\pm$ 1.08
Feruloyl malic acid	<b>5.72 <math>\pm</math> 3.42</b>	11.1 $\pm$ 2.85	3.65 $\pm$ 0.81	1.36 $\pm$ 0.5
Ferulic acid hexoside	<u>3.1 <math>\pm</math> 3.41</u>	1.25 $\pm$ 1.54	40.7 $\pm$ 39.5	113.8 $\pm$ 113.9
Sinapoyl malic acid	<b>1.28 <math>\pm</math> 0.43</b>	1.81 $\pm$ 0.38	0.83 $\pm$ 0.17	0.41 $\pm$ 0.06
Sinapic acid hexoside	<u>41.0 <math>\pm</math> 10.2</u>	49.6 $\pm$ 8.44	67.7 $\pm$ 10.6	56.5 $\pm$ 15.4
Galloyl shikimic acid	<b>0.09 <math>\pm</math> 0.1</b>	0.24 $\pm$ 0.05	0.07 $\pm$ 0.05	0.22 $\pm$ 0.06
<i>Polyphenols and putative flavonoids</i>				
Procyanidin B2	<b>0.32 <math>\pm</math> 0.28</b>	0.09 $\pm$ 0.05	4.21 $\pm$ 1.47	0.89 $\pm$ 1.57
Procyanidin B4	<b>0.28 <math>\pm</math> 0.23</b>	0.09 $\pm$ 0.05	4.76 $\pm$ 1.46	1.02 $\pm$ 0.81

Table 1 (continued)

Metabolite	Mbidde	Matooke	Mchare	Plantain
Flavonoid glucoside I	<u>0.2 <math>\pm</math> 0.09</u>	0.38 $\pm$ 0.16	1.71 $\pm$ 1.46	1.47 $\pm$ 0.75
Flavonoid glucoside II	<b>0.15 <math>\pm</math> 0.06</b>	0.39 $\pm$ 0.12	4.65 $\pm$ 1.07	1.06 $\pm$ 0.31
Flavanone 7-O-glucoside	<b>0.58 <math>\pm</math> 0.29</b>	1.68 $\pm$ 0.23	1.03 $\pm$ 0.21	0.21 $\pm$ 0.05
Flavone	<b>1.17 <math>\pm</math> 1.03</b>	0.25 $\pm$ 0.18	0.52 $\pm$ 0.24	0.78 $\pm$ 0.31
<i>Flavonols</i>				
Dihydroflavonol	<u>0.14 <math>\pm</math> 0.04</u>	0.2 $\pm$ 0.03	0.02 $\pm$ 0.01	0.09 $\pm$ 0.02
Kaempferol	<b>1.2 <math>\pm</math> 0.47</b>	1.94 $\pm$ 0.47	1.32 $\pm$ 0.21	1.24 $\pm$ 0.22
Kaempferol-rutinoside	<b>1.6 <math>\pm</math> 0.55</b>	0.81 $\pm$ 0.61	0.81 $\pm$ 0.19	2.48 $\pm$ 0.95
Dihydrokaempferol	<u>1.29 <math>\pm</math> 0.53</u>	0.19 $\pm$ 0.07	0.9 $\pm$ 0.25	0.59 $\pm$ 0.22
Dihydrokaempferol hexoside	<b>0.06 <math>\pm</math> 0.09</b>	0.52 $\pm$ 0.09	0.25 $\pm$ 0.12	1.26 $\pm$ 0.55
Quercetin	<b>1.95 <math>\pm</math> 1.26</b>	9.71 $\pm$ 3.75	6.54 $\pm$ 3.59	0.43 $\pm$ 0.13
Rutin	<b>2.07 <math>\pm</math> 0.96</b>	0.83 $\pm$ 0.7	1.13 $\pm$ 0.22	4.46 $\pm$ 0.99
Quercetin diglucoside I	<u>62.7 <math>\pm</math> 38.4</u>	1.07 $\pm$ 0.5	46.5 $\pm$ 17.4	26.5 $\pm$ 6.73
Quercetin diglucoside II	<b>40.8 <math>\pm</math> 24.2</b>	1.5 $\pm$ 0.55	45.1 $\pm$ 16.8	24.2 $\pm$ 6.92
Quercetin hexose malic acid	<u>0.57 <math>\pm</math> 0.22</u>	0.06 $\pm$ 0.03	0.01 $\pm$ 0	0.23 $\pm$ 0.07
<i>Flavan-3-ols</i>				
Afzelechin	<b>0.47 <math>\pm</math> 0.29</b>	14.7 $\pm$ 9.24	0.11 $\pm$ 0.05	0.28 $\pm$ 0.11
Epiafzelechin	<b>1.46 <math>\pm</math> 1.06</b>	12.6 $\pm$ 5.45	0.12 $\pm$ 0.09	0.45 $\pm$ 0.21
<i>Other</i>				
Dopamine	<b>45.1 <math>\pm</math> 39.7</b>	n.d.	7.4 $\pm$ 15.7	61.8 $\pm$ 54.5
Nicotinic acid	<u>38.7 <math>\pm</math> 4.4</u>	50.8 $\pm$ 6.28	51.6 $\pm$ 10	80.5 $\pm$ 21.7
Gluconic acid	<u>230 <math>\pm</math> 42.1</u>	349.3 $\pm$ 95.3	11.8 $\pm$ 24.9	124.2 $\pm$ 62.5

features were higher or lower compared to Mbidde. This was different to the comparison to Matooke, which showed that twice as many features were significantly higher in Matooke than in Mbidde accessions.

Identification of the molecular features was performed based on authentic standards and fragmentation patterns and the identified metabolites were quantified relative to the internal standard (Table 1). Mbidde accessions had 2.4- to 15-fold higher levels of monosaccharides compared to all other accessions and 1.6-fold lower levels of disaccharides compared to Matooke. The intermediates of the TCA cycle were only different to Matooke with 1.1- to 2-fold lower levels of pyruvic, *cis*-aconitic and fumaric acid and 2- to 4-fold higher levels of citric and isocitric acid in Mbidde. The exception was malic acid which was 1.1-fold higher in Mbidde compared to Mchare and plantain. The majority of detected amino acids were significantly higher (1.8- to 24-fold) in Mbidde compared to Matooke, Mchare/plantain or all three. However, amino acids, involved in the nitrogen shunt, were lower in Mbidde. Aspartic acid was  $\sim$  3-fold lower in Mbidde compared to Mchare and plantain accessions. Glutamic acid and GABA were 2-fold lower in Mbidde compared to Matooke. Phenylalanine and tryptophan, the precursors for the phenylpropanoid pathway, were 2- to 6-fold higher in Mbidde compared to the other three groups. *trans*-Caffeic acid was over 14-fold higher in Mbidde compared to all other groups. Shikimic acid, aglycone flavonols (quercetin and kaempferol) and flavanonols (afzelechin and epiafzelechin) were 1.5- to 30-fold lower in Mbidde compared to Matooke, whereas glycosylated flavonols (rutin, kaempferol-rutinoside and quercetin diglucosides) were present in 2- to 60-fold higher levels. The quercetin diglucosides were also 2- to 50-fold higher in Mbidde compared to Mchare and plantain. For



**Fig. 4.** Comparison of Mbidde with Matooke, Mchare and plantain accessions. Data was based on molecular features detected in banana pulp. OPLS-DA was performed for (A) Mbidde to Matooke and (B) Mbidde to Mchare and plantain. Volcano plots were created to visualise molecular features significantly higher (blue) or lower (red) in Mbidde compared to (C) Matooke and (D) Mchare and plantain. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

dihydrokaempferol and its hexoside, the opposite trend was observed with respectively 6-fold higher and 8-fold lower levels present in Mbidde compared to the other three groups. Ferulic acid, its malic acid ester and galloyl shikimic acid were also present in 1.5- to 2.6-fold lower levels in Mbidde and procyanidin B2 and B4 with 3-fold higher levels. These differences were, with the exception of feruloyl malic acid, only observed between Mbidde and Matooke. Dopamine was only detected in Mbidde, plantain and Mchare and the latter had ~ 7-fold lower levels compared to the first two.

### 3.4. Distinctive metabolites associated with Mutika/Lujugira subgroup Mbidde in peel

The comparison of peel tissue of Mbidde to Matooke and plantain showed that half of the detected molecular features were significantly different (Fig. S5). These molecular features matched to 86 % and 76 % of significantly different features detected in pulp tissue of Matooke and plantain, respectively. Furthermore, they represented 22 % and 30 % of molecular features significantly different in peel. Mbidde showed ~ 10 % more statistically significant differences compared to plantain than Matooke. Metabolite quantities in peel tissue were compared to pulp tissue and indicated that the majority of trends described above could also be detected in peel (Table S3). The exceptions were most amino acids e.g. alanine and glycine, which were present in 1.5- and 2.4-fold lower levels in Mbidde compared to plantain and glutamic acid in 2-fold higher levels compared to Matooke. The majority of intermediates of the TCA cycle were below the limit of detection and only isocitric acid was significantly different with 2-fold higher and 1.5-fold lower levels in Mbidde compared to Matooke and plantain, respectively. Procyanidin Bs were 2-fold higher in plantain compared to Mbidde. Dopamine levels were overall higher in peel tissue and Mbidde had 5- to 10-fold higher levels compared to plantain and Matooke, respectively. Sinapic acid and its malate ester, quinic acid and an unknown flavonol and its glycoside

were only significantly different in peel. The first was 5-fold higher in Mbidde compared to plantain, quinic acid and the flavonol were 3-fold higher and ~ 100-fold lower in Mbidde compared to Matooke.

### 3.5. Cell wall and lignin content in pulp and peel of four banana groups

Based on the significant differences of several flavonoids, the content of cell wall and lignin was measured in pulp tissue of selected Mbidde, Matooke, Mchare and plantain accessions (Fig. 5A). For Mbidde, Matooke and plantain, peel tissue was also collected and analysed. The standard deviation of cell wall content in pulp and peel varied between 10 and 26 % and 10–43 % for lignin content. Hence, no significant differences could be established in the cell wall content of pulp tissues and only peel tissue of plantain accessions had a distinctly lower cell wall content. The lignin content in the cell wall extracted from pulp and peel tissue, similar to the cell wall content, was not significantly different between the four groups. However, plantain accessions had ~ 1.5-fold lower average value compared to the other three groups.

## 4. Discussion

Previous metabolomics studies of *Musa* spp mainly focused on Cavendish cultivars and/or comparison of triploid accessions (AAA and AAB) at the ripe fruit stage (Nascimento et al., 2019; Pandey et al., 2016; Tsamo et al., 2015). A few studies included a more diverse selection of *Musa* spp. and indicated bananas as a rich source of antioxidants in the human diet. The previous data showed the presence of phenylpropanoids in all *Musa* accessions with a genetic influence on their composition (Borges et al., 2019; Drapal et al., 2020). The green mature stage of *Musa* accessions is the most starch-rich maturity stage and represents the harvest stage for bananas consumed after cooking (Campos et al., 2022). Hence, the present study aimed to elucidate the metabolite composition of green mature banana fruit of Mutika/

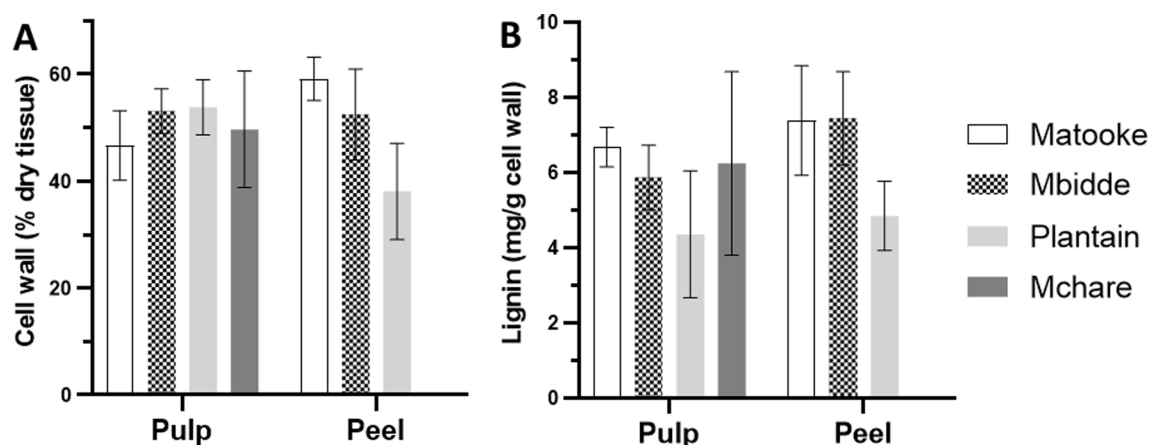


Fig. 5. Quantification of cell wall (A) and lignin (B) content in peel and pulp of Matooke, Mbidde, plantain and Mchare accessions. Bars represent average and standard deviation for three accessions of Matooke, plantain and Mchare and five accessions of Mbidde. Two technical replicates were measured for each accession.

Lujugira (Matooke and Mbidde), Mchare and plantain accessions and establish metabolic identifiers related to banana fruit quality.

#### 4.1. Influence of genome on metabolic composition of banana pulp

The metabolomics analysis applied included hydrophilic primary and specialised metabolites, which currently provides a good representation of the metabolome. A very clear separation of the Matooke, Mbidde, Mchare and plantain metabolomes was evident. A similar separation was previously observed for mature banana pulp of a *Musa* spp panel with diploid, auto- and allotriploid genotypes (Drapal et al., 2020). This separation was partially based on phenylpropanoid levels, which were lowest in AAA and highest in AAB bananas. The same composition was observed in the present study for Matooke and Mbidde (AAA) and plantain (AAB) accessions. Cenci et al. (2019) established that the B genome is responsible for the increased phenylpropanoid levels through a higher expression of flavonoid genes. However, Mchare accessions (AA) contained similar levels of phenylpropanoids compared to plantain, which could be the result of the high allelic contribution of *M. acuminata* ssp. *banksii* on these two genome groups (Perrier, 2009). Mutika/Lujugira had an equal contribution of ssp. *banksii* and *zebrina*, which might be the underlying factor for the equal presence of different aglycone and conjugated flavonoids compared to the dominant presence of one or two flavonoids in plantain and Mchare. Karamura et al. (2016) established a distinct separation of Mutika/Lujugira based on SSR markers, which emphasises the different genetic background of Matooke and Mbidde accessions. The identification of unique metabolites indicated that the banana groups differ in glycosylation of phenylpropanoids rather than phenylpropanoid pathways. The high and low levels of additional glycosylation in Matooke and Mchare, respectively, could not be directly correlated to the allelic contribution of wild *M. acuminata* and will require additional transcriptional, translational or enzymatic elucidation (Perrier, 2009). This is likely related to the allelic contribution of wild *M. acuminata* similar to the quantitative differences of phenylpropanoids discussed above. The more focused analysis, including only metabolites detected in all samples, indicated that the genetic regulation of specialised pathways also influenced the regulation of core metabolism e.g. glycolysis, TCA cycle and amino acids, as previously reported (Drapal et al., 2020). This data furthermore emphasised the similarity between Mchare and plantain compared to Mutika/Lujugira.

Volatile analysis of accessions used in the present study could further elucidate the different flavour profiles of Matooke, Mchare and plantain. The majority of features, significantly different between the three groups, were identified as phenylpropanoids. Similar phenylalanine levels and different composition of phenylpropanoids, in particular the

high levels of procyanidins in Mchare, indicates that the genome groups vary in their regulation of phenylpropanoid pathways. Depending on the genetic regulation of volatile biosynthesis, this could lead to very specific aroma profiles and can be used for sensory analysis of consumer preference in different regions and countries. The metabolite data showed a low intra-group heterogeneity of ~ 32 % for plantain and Matooke and ~ 40 % for Mchare. This corroborates previous studies, reporting genetically extremely uniform accessions within Matooke, Mchare and plantain in Africa despite morphological diversity and should facilitate the transcriptomic analysis of phenylpropanoid biosynthesis and degradation (Campos et al., 2022; Perrier et al., 2019; Tsamo et al., 2015).

#### 4.2. Metabolite composition underlying phenotypic traits of banana fruit in Mutika/Lujugira

Fruit quality is an important factor for the acceptance and introduction of resistant Matooke hybrids and includes visual characteristics, peel properties, cooking time as well as absence of sweetness, good Matooke taste and smooth mouthfeel (Akankwasa et al., 2021; Tumuhimbise et al., 2019). Mbidde accessions, which are genetically uniform with Matooke at the genomic DNA level (Kitavi et al., 2020), have an astringent bitter taste, due to higher tannin content, and are solely used for beer brewing (Bashmil et al., 2021). However, there is no up-to-date literature report elucidating the specific metabolic composition of Mbidde underlying the undesired sensory properties. The present study established that ~ 30 % of molecular features were significantly different between Mbidde and Matooke, suggesting a different epigenetic or post-translational regulation between the two clone sets (Kitavi et al., 2020). The differences included higher levels of organic acids (1.1- to 3-fold), monosaccharides (2- to 15-fold) and procyanidins B (3-fold), probably resulting in a more sour, sweet and astringent taste of Mbidde pulp tissue. Furthermore, amino acids referring umami taste as well as precursors for lignin and condensed tannin biosynthesis were present in lower levels (3-fold). Tannin and lignin content contribute to the taste and textural properties of food. Tannins are classed as a bitter, anti-nutritional factor and can inhibit the uptake of nutritional compounds. The reduced content of tannin precursors, increased levels of procyanidins (condensed tannins) and previously reported higher levels of total tannin content in Mbidde accessions, indicates an up-regulation of the condensed tannin pathway in this clone set (Bashmil et al., 2021). Lignin is an important component of the cell wall as mechanical support and physical barrier. The metabolite data showed a reduction of lignin precursors in Mbidde compared to Matooke, which indicates a higher content of lignin, similar to tannin. The quantification of lignin in Mbidde showed ~ 14 % lower levels compared to Matooke and a ~ 22 %

higher percentage of cell wall per g dry tissue. Both values were not statistically significant between the two clone sets and would suggest the presence of an additional cell wall structure (e.g. latex) and therefore, different texture properties in Mbidde (Kyamuhangire et al., 2006). Studies of the cooking process in chickpea and other vegetables showed no change of the tasteless lignin content (Herranz et al., 1983). This suggests that the lignin and cell wall composition of raw green mature Mbidde bananas can be used as a “marker” for undesired traits in new Matooke hybrids.

Glycosylation is an important modification of phenylpropanoids and flavonoids for bio-distribution, metabolism, storage of precursors for later re-activation and stabilisation of the metabolites during the storage/accumulation period in the plant and enhances the respective biological function for the human diet (Kytidou et al., 2020). As mentioned above, Mbidde contained higher amounts of aglycone flavonoids and Matooke significantly higher amounts of glycosylated flavonoids. This could imply an increased (i) flavonoid biosynthesis in Mbidde, (ii) flavonoid storage in Matooke or (iii) flavonoid degradation to aromatic volatiles in Mbidde. All of these hypotheses require further analysis, preferably a combination of transcriptomics, volatile analysis and fluxomics.

#### 4.3. Metabolite differences in banana pulp of plantain and Mchare compared to Mbidde

Mbidde is the only banana group that is solely used for beer brewing, contrary to bananas from other groups, which are used for beer brewing as well as eaten raw or cooked. Hence, Mbidde was used in the present study as a comparator for other genome groups, to assess aspects of fruit quality related to metabolic composition e.g. flavour, colour and nutrition. The untargeted data showed an even higher number of molecular features (~40 %) significantly different between plantain and Mchare compared to Mbidde. This substantiates the distinct genetic background of Mutika/Lujugira based on SSR markers, as discussed above (Karamura et al., 2016). Similar to Matooke, the metabolite composition of plantain and Mchare indicates that Mbidde bananas should be perceived as more sour and astringent (Kyamuhangire et al., 2006). This was based on plantain and Mchare containing less malic acid (1.5- to 3.5-fold) and gluconic acid (2- to 20-fold) and less *trans*-caffeic acid (>11-fold), aglycone and glycosylated forms of flavonoids (~1.5- to 2.5-fold) compared to Mbidde (Ulla et al., 2016). Frank et al. (2007) observed that *trans*-caffeic acid as well as its thermal degradation products contribute to the bitter taste. Previous literature highlighted that phenylpropanoids and flavonoids with sugar moieties, despite their non-bitter taste, can enhance the perception of bitterness of non-astringent metabolites (Chen et al., 2022). Additionally, significantly higher levels of aspartic acid (>3.2 fold) in plantain and Mchare should influence the perception of umami (Morris et al., 2010). The levels for mono- and di-saccharides showed a ~ 16-fold decrease and 1.6-fold increase in plantain and Mchare compared to Mbidde. Depending on the sweetening power of the individual mono- and di-saccharides present, these differences will influence the perception of sweetness (Carocho et al., 2017). However, for exact quantification of individual sugars and their impact on fruit quality, a targeted metabolomics method will need to be applied in combination with sensory assessment.

The present data showed a distinct chemotype of plantains, with the highest levels of phenylpropanoid precursors for lignin, the highest content of cell wall per dry weight and the lowest content of lignin. Metabolite identification highlighted that the hexosides of ferulic and sinapic acid were the dominant phenylpropanoids. Previous literature indicated that the glycosylation is a storage response and the phenylpropanoids cannot be accessed by enzymes for monolignol biosynthesis, leading to the contradictory data of high precursor and low lignin content (Kytidou et al., 2020; Perez-Fons et al., 2019; Tsamo et al., 2015). Furthermore, the levels of procyanidins B were ~ 14-fold higher in plantains compared to the other two genome groups. This suggests that a

large portion of the cell wall in plantains is composed of condensed tannins, which is reflected in the high levels of procyanidins. Plantains were the only genome group containing a B genome in the present study, which suggests that the B genome contains regulatory genes for condensed tannin biosynthesis (Carreel et al., 2002; Perrier, 2009). Tannins have a range of biological activities related to fruit quality including antimicrobial, antioxidant and anti-inflammatory, which represent important traits for farmer and consumer preference and can be elucidated through a combination of system biology techniques and phenotyping.

#### 4.4. Influence of peel on banana fruit quality

In *Musa* species, both the pulp and peel develop from the ovary during parthenocarp. The former originates from the outer lining of the loculus and the latter from the vascular skin of the ovary. Previous literature established migration of water, minerals and metabolites such as tannins from peel to pulp (Chang & Hwang, 1990). Previous studies also reported an increase of anti-nutritional components (e.g. tannins) in the fruit tissue under abiotic stress and pesticide treatment (Pandey et al., 2016). Tannins and other phenolics are naturally enriched in the pulp during ripening and metabolised through polymerisation to reduce the astringent mouthfeel caused by these metabolites. During the ripening process, ethylene is mainly produced in the pulp and diffused to the peel (John & Marchal, 1995). These observations highlight the prevalent interaction between pulp and peel tissue. The peel presents a protective layer for the pulp and contains a wide range of phenylpropanoids involved in constitutive and induced defence mechanisms, e.g. cell wall strengthening and signalling (Pandey et al., 2016). Phenylpropanoids are traditionally hydrophilic compounds referring a bitter and/or astringent taste, which suggests the peel as a source of metabolites, potentially transported to the pulp, with an undesired effect on the taste aspect of fruit quality. Hence, the present study included metabolite analysis of peel tissue, which due to capacity of processing facilities was only available for Matooke, Mbidde and plantain. The majority of metabolic differences detected in pulp tissue of Mbidde compared to Matooke and plantain, could also be observed in the peel tissue and corresponded to less than a third of the metabolic differences detected in the peel. This is consistent with the metabolic plasticity expected of the pulp and peel tissue based on their function as juicy, soft, edible flesh and protective skin, respectively. Mbidde was more similar to the other Mutika/Lujugira clone set Matooke than the genetically different plantains, as discussed for the pulp. The identification of phenylpropanoids highlighted different regulation of procyanidins in plantain and aromatic amino acids and lignin precursors in Mbidde. The genetic regulation of these pathways could be a result of the genome group as well as the adaptation processes to different environments (Pott et al., 2020). For example, plantains are cultivated in the humid lowlands and rainforests of Central and West Africa, Matooke in the region of the Great Lakes and Mchare in the highlands of East Africa and other humid altitude zones of Tanzania (Perrier et al., 2019). Based on the abiotic and biotic stresses present in the respective environments, the genome groups will have adapted their metabolic regulation for the production of phytoanticipins in leaf and fruit tissue (le Gall et al., 2015). The correlation of metabolic similarities and differences between pulp and peel was previously reported (Drapal et al., 2020) and could provide the foundation for the development of early screening tools for resistance and fruit quality traits. Experimental studies for this purpose will require ovary tissue, peel and pulp tissue of mature green and ripe stages of banana development and improved identification of molecular features. The latter is an on-going obstacle of metabolomics due to the fast amount of metabolites present in plants and their potential for modifications (Muhich et al., 2022).

## 5. Conclusion

The present study highlighted significant differences in metabolite composition between the banana groups grown and consumed in Sub-Saharan Africa. The low intra-group heterogeneity indicated defined pulp chemotypes among the banana groups, referring unique sensory properties. Matooke, Mchare and plantain are cultivated in different climate zones in Africa with consumer preference being biased towards the local banana group. However, consumers agree on the generally undesired sensory properties of Mbidde accessions, which are used for the brewing of beer. The present data highlighted a broad range of key metabolites responsible for the undesired sensory properties and poor consumer uptake. The data showed increases of metabolites associated with sour (e.g. citric acid), sweet (e.g. glucose), bitter (*trans*-caffeic acid) and astringent (e.g. quercetin diglucosides) taste and decrease of metabolites associated with umami (e.g. aspartic acid) in Mbidde accessions. The range of metabolites detected in Matooke, Mchare and plantain can be used as a reference for selecting hybrids during breeding, with a metabolite composition that falls within the observed range. Furthermore, the metabolite data suggested that metabolic properties of Mbidde accession can be used as a marker for unfavourable consumer traits associated with new disease resistant EAHB hybrids.

## CRedit authorship contribution statement

M.D., A.B., R.S. and P.D.F. devised the concept. D.A., B.U., A.B. and R.S. selected plant materials. M.D. generated the metabolomics datasets, assembled the figures, compiled supplementary tables and drafted the manuscript. A.B., R.S. and P.D.F. aided the interpretation of results, edited the manuscript and secured funding. All authors approve the content and interpretation of the final draft.

## Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Prof. Paul D. Fraser reports financial support was provided by CGIAR Research Program on Roots, Tubers and Bananas (RTB).

## Data availability

As stated in the manuscript, raw data is available under <https://data.mendeley.com/drafts/rntnvdby8s>

## Acknowledgments

The authors would like to thank NARO for providing Mbidde pulp tissue and 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. Funding from the BBSRC-GCRF NUTRI-FOOD (BB/T008946/1) project to PDF is also acknowledged.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.foodchem.2023.137481>.

## References

- Akankwasa, K., Marimo, P., Tumuhimbise, R., Asasira, M., Khakasa, E., Mpirirwe, I., Kleih, U., Forsythe, L., Flidel, G., Dufour, D., & Nowakunda, K. (2021). The East African highland cooking bananas 'Matooke' preferences of farmers and traders: Implications for variety development. *International Journal of Food Science & Technology*, 56(3), 1124–1134. <https://doi.org/10.1111/ijfs.14813>
- Bashmil, Y. M., Ali, A., BK, A., Dunshea, F. R., & Suleria, H. A. R. (2021). Screening and Characterization of Phenolic Compounds from Australian Grown Bananas and Their Antioxidant Capacity. *Antioxidants*, 10(10), 1521. <https://doi.org/10.3390/antiox10101521>.
- Borges, C. V., Amorim, E. P., Leonel, M., Gomez, H. A. G., dos Santos, T. P. R., Ledo, C. A. da S., Belin, M. A. F., de Almeida, S. L., Minatel, I. O., & Lima, G. P. P. (2019). Post-harvest physicochemical profile and bioactive compounds of 19 bananas and plantains genotypes. *Bragantia*, 78, 284–296. [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0006-87052019000200284&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0006-87052019000200284&nrm=iso).
- Campos, N. A., Colombié, S., Moing, A., Cassan, C., Amah, D., Swennen, R., Gibon, Y., & Carpentier, S. C. (2022). From fruit growth to ripening in plantain: A careful balance between carbohydrate synthesis and breakdown. *Journal of Experimental Botany*, 73(14), 4832–4849. <https://doi.org/10.1093/jxb/erac187>
- Carocho, M., Morales, P., & Ferreira, I. C. F. R. (2017). Sweeteners as food additives in the XXI century: A review of what is known, and what is to come. *Food and Chemical Toxicology*, 107, 302–317. <https://doi.org/10.1016/j.fct.2017.06.046>
- Carreel, F., de Leon, D., Lagoda, P., Lanaud, C., Jenny, C., P Horry, J., & du Montcel, H. (2002). Ascertaining maternal and paternal lineage within Musa by chloroplast and mitochondrial DNA RFLP analyses. *Genome*, 45, 679–692. <https://doi.org/10.1139/g02-033>.
- Cenci, A., Hueber, Y., Zorrilla-Fontanesi, Y., van Wesemael, J., Kissel, E., Gislar, M., Sardos, J., Swennen, R., Roux, N., Carpentier, S. C., & Rouard, M. (2019). Effect of paleopolyploidy and allopolyploidy on gene expression in banana. *BMC Genomics*, 20(1), 244. <https://doi.org/10.1186/s12864-019-5618-0>
- Chang, W. H., & Hwang, Y. J. (1990). EFFECT OF ETHYLENE TREATMENT ON THE RIPENING, POLYPHENOLOXIDASE ACTIVITY AND WATER-SOLUBLE TANNIN CONTENT OF TAIWAN NORTHERN BANANA AT DIFFERENT MATURITY STAGES AND THE STABILITY OF BANANA POLYPHENOLOXIDASE. *Acta Horticulturae*, 275, 603–610. <https://doi.org/10.17660/ActaHortic.1990.275.75>.
- Chen, Y.-H., Zhang, Y.-H., Chen, G.-S., Yin, J.-F., Chen, J.-X., Wang, F., & Xu, Y.-Q. (2022). Effects of phenolic acids and quercetin-3-O-rutinoside on the bitterness and astringency of green tea infusion. *Npj Science of Food*, 6(1), 8. <https://doi.org/10.1038/s41538-022-00124-8>
- Drapal, M., Amah, D., Schöny, H., Brown, A., Swennen, R., & Fraser, P. D. (2020). Assessment of metabolic variability and diversity present in leaf, peel and pulp tissue of diploid and triploid Musa spp. *Phytochemistry*, 176, Article 112388. <https://doi.org/10.1016/j.phytochem.2020.112388>.
- Ferris, R. S. B., Ortiz, R., & Vuylsteke, D. (1999). Fruit quality evaluation of plantains, plantain hybrids, and cooking bananas. *Postharvest Biology and Technology*, 15(1), 73–81. [https://doi.org/10.1016/S0925-5214\(98\)00067-2](https://doi.org/10.1016/S0925-5214(98)00067-2)
- Frank, O., Blumberg, S., Kunert, C., Zehentbauer, G., & Hofmann, T. (2007). Structure Determination and Sensory Analysis of Bitter-Tasting 4-Vinylcatechol Oligomers and Their Identification in Roasted Coffee by Means of LC-MS/MS. *Journal of Agricultural and Food Chemistry*, 55(5), 1945–1954. <https://doi.org/10.1021/jf0632280>
- le Gall, H., Philippe, F., Domon, J.-M., Gillet, F., Pelloux, J., & Rayon, C. (2015). Cell wall metabolism in response to abiotic stress. *Plants*, 4(1), 112–166. <https://doi.org/10.3390/plants4010112>
- Herranz, J., Vidal-Valverde, C., & Rojas-Hidalgo, E. (1983). Cellulose, Hemicellulose and Lignin Content of Raw and Cooked Processed Vegetables. *Journal of Food Science*, 48(1), 274–275. <https://doi.org/10.1111/j.1365-2621.1983.tb14847.x>
- John, P., & Marchal, J. (1995). Ripening and biochemistry of the fruit. In *Bananas and Plantains* (pp. 434–467). Netherlands: Springer. [https://doi.org/10.1007/978-94-011-0737-2\\_15](https://doi.org/10.1007/978-94-011-0737-2_15).
- Karamura, D., Kitavi, M., Nyine, M., Ochola, D., Ocimati, W., Muhangi, S., Talengera, D., & Karamura, E. B. (2016). Genotyping the local banana landrace groups of East Africa. *Acta Horticulturae*, 1114, 67–74. <https://doi.org/10.17660/ActaHortic.2016.1114.9>.
- Kitavi, M., Cashell, R., Ferguson, M., Lorenzen, J., Nyine, M., McKeown, P. C., & Spillane, C. (2020). Heritable epigenetic diversity for conservation and utilization of epigenetic germplasm resources of clonal East African Highland banana (EAHB) accessions. *Theoretical and Applied Genetics*, 133(9), 2605–2625. <https://doi.org/10.1007/s00122-020-03620-1>
- Kyamuhangire, W., Kreckling, T., Reed, E., & Pehrson, R. (2006). The microstructure and tannin content of banana fruit and their likely influence on juice extraction. *Journal of the Science of Food and Agriculture*, 86(12), 1908–1915. <https://doi.org/10.1002/jsfa.2553>
- Kytidou, K., Artola, M., Overkleeft, H. S., & Aerts, J. M. F. G. (2020). Plant Glycosides and Glycosidases: A Treasure-Trove for Therapeutics. *Frontiers in Plant Science*, 11. <https://doi.org/10.3389/fpls.2020.00357>
- Madalla, N. A., Massawe, C., Shimmwela, M., Mbongo, D., Kindimba, G., Kubiriba, J., ... Ortiz, R. (2022). TARIBAN1, TARIBAN2, TARIBAN3, and TARIBAN4 'Matooke' Cooking Banana Cultivars for the Great Lakes Region of Africa. *HortScience*, 57(12), 1588–1592. <https://doi.org/10.21273/HORTSCI16854-22>.
- Moreira-Vilar, F. C., Siqueira-Soares, R. de C., Finger-Teixeira, A., Oliveira, D. M. de, Ferro, A. P., da Rocha, G. J., Ferrarese, M. de L. L., dos Santos, W. D., & Ferrarese-Filho, O. (2014). The Acetyl Bromide Method Is Faster, Simpler and Presents Best Recovery of Lignin in Different Herbaceous Tissues than Klason and Thioglycolic Acid Methods. *PLoS ONE*, 9(10), e110000. <https://doi.org/10.1371/journal.pone.0110000>.
- Morris, W. L., Shepherd, T., Verrall, S. R., McNicol, J. W., & Taylor, M. A. (2010). Relationships between volatile and non-volatile metabolites and attributes of processed potato flavour. *Phytochemistry*, 71(14), 1765–1773. <https://doi.org/10.1016/j.phytochem.2010.07.003>
- Muhich, A. J., Agosto-Ramos, A., & Kliebenstein, D. J. (2022). The ease and complexity of identifying and using specialized metabolites for crop engineering. *Emerging Topics in Life Sciences*, 6(2), 153–162. <https://doi.org/10.1042/ETLS20210248>
- Nascimento, T. P., Castro-Alves, V. C., Castelan, F. P., Calhau, M. F. N. S., Saraiva, L. A., Agopian, R. G., & Cordenunsi-Lysenko, B. R. (2019). Metabolomic profiling reveals that natural biodiversity surrounding a banana crop may positively influence the

- nutritional/sensorial profile of ripe fruits. *Food Research International*, 124, 165–174. <https://doi.org/10.1016/j.foodres.2018.07.050>
- Ndayihanzamaso, P., Mostert, D., Matthews, M. C., Mahuku, G., Jomanga, K., Mpanda, H. J., Mduma, H., Brown, A., Uwimana, B., Swennen, R., Tumuhimbise, R., & Viljoen, A. (2020). Evaluation of Mchare and Matooke Bananas for Resistance to *Fusarium oxysporum* f. sp. cubense Race 1. *Plants*, 9(9), 1082. <https://doi.org/10.3390/plants9091082>
- Pandey, A., Alok, A., Lakhwani, D., Singh, J., Asif, M. H., & Trivedi, P. K. (2016). Genome-wide expression analysis and metabolite profiling elucidate transcriptional regulation of flavonoid biosynthesis and modulation under abiotic stresses in banana. *Scientific Reports*, 6, 31361. <https://doi.org/10.1038/srep31361> <https://www.nature.com/articles/srep31361#supplementary-information>
- Pang, Z., Chong, J., Zhou, G., de Lima Morais, D. A., Chang, L., Barrette, M., Gauthier, C., Jacques, P.-É., Li, S., & Xia, J. (2021). MetaboAnalyst 5.0: Narrowing the gap between raw spectra and functional insights. *Nucleic Acids Research*, 49(W1), W388–W396. <https://doi.org/10.1093/nar/gkab382>
- Perez-Fons, L., Bohorquez-Chaux, A., Irigoyen, M. L., Garceau, D. C., Morreel, K., Boerjan, W., Walling, L. L., Becerra Lopez-Lavalle, L. A., & Fraser, P. D. (2019). A metabolomics characterisation of natural variation in the resistance of cassava to whitefly. *BMC Plant Biology*, 19(1), 518. <https://doi.org/10.1186/s12870-019-2107-1>
- Perrier, X. (2009). Combining Biological Approaches to Shed Light on the Evolution of Edible Bananas. 2009, 7, 18. <https://doi.org/10.17348/era.7.0.199-216>
- Perrier, X., Jenny, C., Bakry, F., Karamura, D., Kitavi, M., Dubois, C., Hervouet, C., Philippon, G., & de Langhe, E. (2019). East African diploid and triploid bananas: A genetic complex transported from South-East Asia. *Annals of Botany*, 123(1), 19–36. <https://doi.org/10.1093/aob/mcy156>
- Ploetz, R. (1994). Panama disease: Return of the first banana menace. *International Journal of Pest Management*, 40, 326–336.
- Pott, D. M., Vallarino, J. G., Cruz-Rus, E., Willmitzer, L., Sánchez-Sevilla, J. F., Amaya, I., & Osorio, S. (2020). Genetic analysis of phenylpropanoids and antioxidant capacity in strawberry fruit reveals mQTL hotspots and candidate genes. *Scientific Reports*, 10(1), 20197. <https://doi.org/10.1038/s41598-020-76946-x>
- Price, E. J., Wilkin, P., Sarasan, V., & Fraser, P. D. (2016). Metabolite profiling of Dioscorea (yam) species reveals underutilised biodiversity and renewable sources for high-value compounds. *Scientific Reports*, 6, 29136. <https://doi.org/10.1038/srep29136>
- Sidhu, J. S., & Zafar, T. A. (2018). Bioactive compounds in banana fruits and their health benefits. *Food Quality and Safety*, 2(4), 183–188. <https://doi.org/10.1093/fqsafe/fyy019>
- Tenkouano, A., Lamien, N., Agogbua, J., Amah, D., Swennen, R., Traoré, S., ... Ortiz, R. (2019). Promising High-Yielding Tetraploid Plantain-Bred Hybrids in West Africa. *International Journal of Agronomy*, 2019, 1–8. <https://doi.org/10.1155/2019/3873198>
- Tsamo, C. V. P., Herent, M. F., Tomekpe, K., Happi Emaga, T., Quetin-Leclercq, J., Rogez, H., Larondelle, Y., & Andre, C. (2015). Phenolic profiling in the pulp and peel of nine plantain cultivars (*Musa* sp.). *Food Chem*, 167, 197–204. <https://doi.org/10.1016/j.foodchem.2014.06.095>
- Tumuhimbise, R., Buregyeya, H., Kubiriba, J., Tushemereirwe, W. K., Barekye, A., Tendo, R. S., Namagembe, B., Muhangi, S., Kazigye, F., Talengera, D., Tindamanyire, J., Akankwasa, K., Nabulya, G., Namaganda, J., William, W., Kushaba, A., Namuddu, M., Oyesigye, N., Namanya, P., ... Karamura, E. (2019). 'NABIO808' (Syn. 'NAROBAN5'): A tasty cooking banana cultivar with resistance to pests and diseases. *Crop Breeding and Applied Biotechnology*, 19(4), 491–495. <https://doi.org/10.1590/1984-70332019v19n4c71>
- Ulla, H., Antti, K., Oskar, L., & Mari, S. (2016). Genetic basis of flavor sensitivity and food preferences. In *Flavor* (pp. 203–227). Elsevier. <https://doi.org/10.1016/B978-0-08-100295-7.00010-4>
- Vermeir, S., Hertog, M. L. A. T. M., Vankerschaver, K., Swennen, R., Nicolai, B. M., & Lammertyn, J. (2009). Instrumental based flavour characterisation of banana fruit. *LWT - Food Science and Technology*, 42(10), 1647–1653. <https://doi.org/10.1016/j.lwt.2009.05.024>
- Wuyts, N., De Waele, D., & Swennen, R. (2006). Extraction and partial characterization of polyphenol oxidase from banana (*Musa acuminata* Grande naine) roots. *Plant Physiology and Biochemistry*, 44(5), 308–314. <https://doi.org/10.1016/j.plaphy.2006.06.005>