

Article

# Gas Chromatography-Mass Spectrometry and Single Nucleotide Polymorphism-Genotype-By-Sequencing Analyses Reveal the Bean Chemical Profiles and Relatedness of *Coffea canephora* Genotypes in Nigeria

Chinyere F. Anagbogu <sup>1,2,3,\*</sup> , Christopher O. Ilori <sup>2</sup>, Ranjana Bhattacharjee <sup>4</sup> ,  
Olufemi O. Olaniyi <sup>3</sup> and Diane M. Beckles <sup>1</sup>

<sup>1</sup> Department of Plant Sciences, University of California, Davis, CA 95616, USA; dmbeckles@ucdavis.edu

<sup>2</sup> Department of Crop Protection and Environmental Biology, University of Ibadan, Ibadan 200284, Nigeria; chrisilori@yahoo.com

<sup>3</sup> Crop Improvement Division, Cocoa Research Institute of Nigeria, Ibadan 200255, Nigeria; yinkamoji@yahoo.com

<sup>4</sup> Bioscience Laboratory, International Institute of Tropical Agriculture, Ibadan 200285, Nigeria; r.bhattacharjee@cgiar.org

\* Correspondence: flora2na@yahoo.com; Tel.: +234-803-764-6412

Received: 4 September 2019; Accepted: 1 October 2019; Published: 18 October 2019



**Abstract:** The flavor and health benefits of coffee (*Coffea* spp.) are derived from the metabolites that accumulate in the mature bean. However, the chemical profiles of many *C. canephora* genotypes remain unknown, even as the production of these coffee types increases globally. Therefore, we used Gas Chromatography-Mass Spectrophotometry to determine the chemical composition of *C. canephora* genotypes in Nigeria—those conserved in germplasm repositories and those cultivated by farmers. GC-MS revealed 340 metabolites in the ripe beans, with 66 metabolites differing ( $p$ -value < 0.05) across the represented group. Univariate and multivariate approaches showed that the ‘Niaouli’ genotypes could be clearly distinguished from ‘Kouillou’ and ‘Java’ genotypes, while there was almost no distinction between ‘Kouillou’ and ‘Java,’. Varietal genotyping based on bean metabolite profiling was synchronous with that based on genome-wide Single Nucleotide Polymorphism analysis. Across genotypes, the sucrose-to-caffeine ratio was low, a characteristic indicative of low cup quality. The sucrose-to-caffeine ratio was also highly correlated, indicative of common mechanisms regulating the accumulation of these compounds. Nevertheless, this strong correlative link was broken within the ‘Niaouli’ group, as caffeine and sucrose content were highly variable among these genotypes. These ‘Niaouli’ genotypes could therefore serve as useful germplasm for starting a Nigerian *C. canephora* quality improvement breeding program.

**Keywords:** *Coffea canephora*; metabolomics; genomics; cup quality; health benefits

## 1. Introduction

Coffee (*Coffea* spp.) is one of the most consumed beverages globally, and there is a clear correlation between the sensory quality of a coffee variety and its market value [1]. Although *Coffea canephora* does not have the cup quality of the more popular *Coffea arabica*, it continues to be widely grown, especially in regions where farming is low intensive because of its relative tolerance to a range of biotic and abiotic stresses [2]. As a result of this, and its competitive pricing, the *C. canephora* market share has increased in recent years, totaling 40% of the coffee traded globally [3]. Importantly, *C. canephora* forms a critical source of income for millions of smallholder coffee farmers in developing countries [3]. In

Nigeria, it accounts for 90% of coffee production [4]. Since quality is a major selection criterion for coffee improvement [5], it is important to determine the chemical profiles of *C. canephora* varieties which are important to many resource-poor farmers.

Little is known about the genetic background and the quality profile of the genotypes used for coffee cultivation in Nigeria. The early establishment of a *C. canephora* polyclonal seed garden in Nigeria was based on five sexually compatible genotypes of ‘Java Robusta’, ‘Kouillou’ and ‘Niaouli’ varieties at the Cocoa Research Institute of Nigeria (CRIN), for replication and distribution to farmers [6]. The use of these varieties led to an increase in coffee yields from 800 kg to 1.4 ton per hectare [6]. We recently used Single Nucleotide Polymorphism-Genotype-By-Sequencing analyses to determine the relatedness of some of the conserved and cultivated *C. canephora* genotypes in Nigeria [7]. The data showed that the genetic diversity of the conserved CRIN germplasm was narrow and comprised of three genetic groups (‘Niaouli’, ‘Kouillou’ and ‘Java Robusta’)—of which, one of the groups (‘Niaouli’) dominated the cultivated accessions [7]. The chemical composition of the beans from these genotypes has yet to be investigated. Therefore, their sensory profiles are unknown.

Understanding the metabolomics profile of coffee is essential for coffee cup quality improvement. Metabolite profiling has been used to discern coffees of different origins, e.g., Asia, Africa, and South America [8], and in linking high grade Colombian *C. arabica* beans to its higher sucrose content [9]. Metabolite profiling is also the most efficient way to determine the relative level of key chemicals in coffee that affect quality [10]. Several compounds with important contributions to the complex chemistry of coffee roasting have already been identified [11,12]. For example, two abundant metabolites, caffeine and chlorogenic acid at high amounts, contribute to bitterness and lead to low cup quality [13,14]. Among the fatty acids, linoleic and palmitic acids correlate with high-quality coffee, while oleic and stearic acids are associated with inferior flavors [15]. In *C. arabica*, many of the alcohols, aldehydes, hydrocarbons, and ketones present in the beans are associated with poor aroma, volatility and acidic profile, causing an earthy flavor [16].

While the metabolite profiles of coffee beans from various regions are well represented in literature [8], the chemical composition of the beans of *C. canephora* genotypes in Nigeria is unknown, and their quality from a chemical perspective is lacking. A combination of genome-wide Single Nucleotide Polymorphisms and metabolomic analyses could help to link variation in genotype to variation in phenotype (mainly quality parameters). Some Single Nucleotide Polymorphisms (SNPs) may even cause significant changes in the accumulation of some metabolites, which may be manifested at the phenotypic level as alterations in disease resistance and cup quality [5,17]. Different analytical techniques such as High-Performance Liquid Chromatography (HPLC) [18,19], Gas Chromatograph-Mass Spectrometry (GC-MS) and Proton Nuclear Magnetic Resonance ( $^1\text{H}$ -NMR) [20] have been successfully used for the classification of coffee genotypes. Metabolomics of coffee beans will help to decipher the features contributing to good quality and their profiles, while the genomics study will help in linking good quality metabolomic profiles to DNA polymorphisms.

The aim of this study is to determine the chemical fingerprint of the coffee germplasm maintained at CRIN and those grown by Nigerian farmers, and to relate this to their SNP-variation. Specifically, we aim to: (a) assess the metabolomic profiles of *C. canephora* genotypes cultivated in Nigeria, (b) determine the chemical diversity of the genotypes in relation to SNP diversity and (c) identify genotypes with potential good cup quality to be incorporated in the breeding program for cup quality improvement.

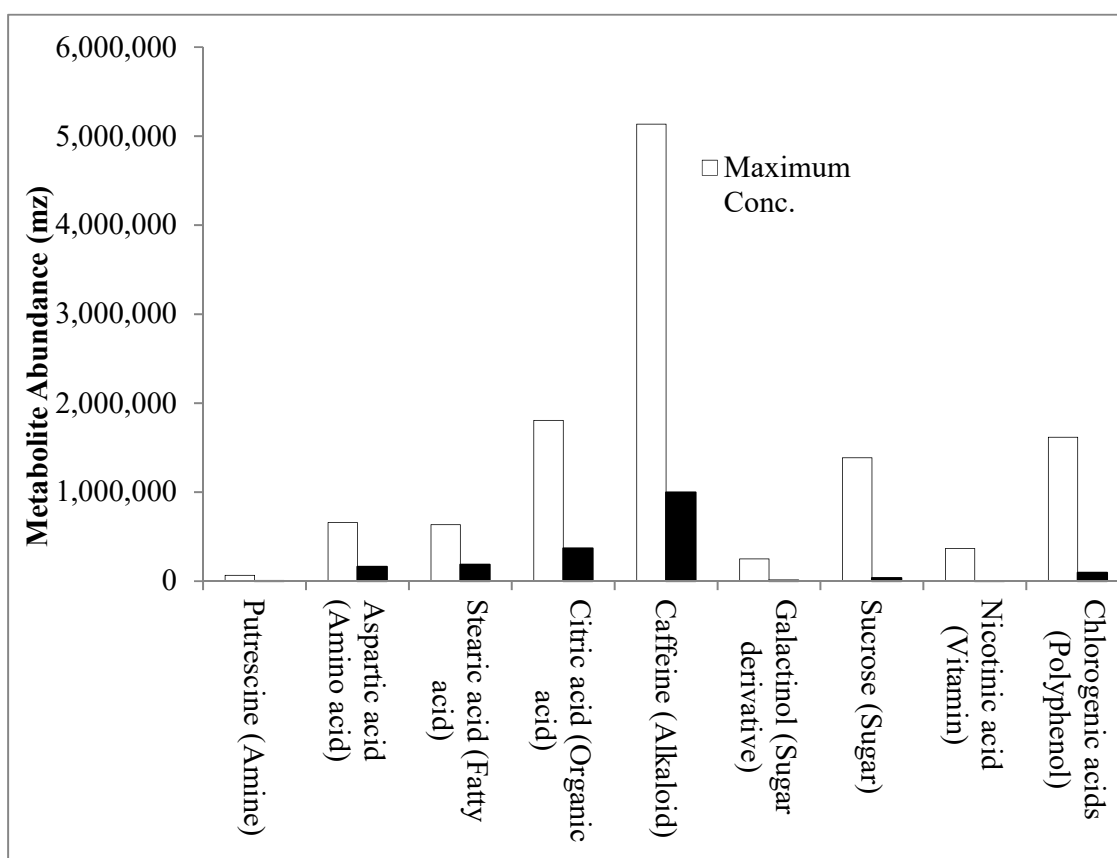
## 2. Results and Discussion

### 2.1. Metabolite Profiles of the Three *C. canephora* Genotypes: ‘Niaouli’, ‘Kouillou’ and ‘Java Robusta’

A total of 340 untargeted metabolites were detected—of which, 163 (Table S1) were identified across all three varieties of *C. canephora*, ‘Niaouli’, ‘Kouillou’ and ‘Java Robusta’, used in the study. A similar number of metabolites (182) was detected in Asian palm civet coffee using GC-MS [21]. The metabolites identified in our study were grouped into eight biochemical classes: amines, amino acids,

sugars, sugar derivatives, organic acids, fatty acids, phenolic acids/alkaloids and inorganic compound (Table S1).

The major compounds identified in the Nigerian varieties were compared to those in other studies. Caffeine, chlorogenic acids, quinic acid, citric acid and sucrose were the most abundant metabolites identified in these varieties (Figure 1), which is similar to the predominant compounds identified in *C. arabica* derived from Asian palm civet [21]. Stearic acid, palmitic acid, linoleic acid and pelargonic acid were the major fatty acids in the three Nigerian genotypes (Table S1). This fatty acid profile was similar to that described by Dong et al., who studied the green coffee beans of seven cultivars of *C. canephora* grown in Hainan Province, China [22]. In their study, and similar to ours, linoleic acid and palmitic acid were among the most abundant fatty acids, but oleic acid, and arachidic acid were also high in content [22]. The most abundant amino acids in our study did not overlap with those identified by Dong et al. [22]. In this study, aspartic acid, glutamic acid, proline, and tryptophan were highest in abundance (Table S1), while Dong et al. identified leucine, lysine and arginine to be the most dominant amino acids in their study [22].



**Figure 1.** Variation in the relative abundance of coffee metabolites across genotypes. Metabolites that occur with maximal abundance within each of the eight main metabolite classes are indicated by the white bars. Also shown (indicated as a black bar) is the minimal abundance of that metabolite across genotypes.

Next, we looked at the compounds known to influence coffee quality. The relative abundance of important cup quality precursors, including amino acids, fatty acids, and sugars was low, while the compounds associated with poor hedonistic value, such as caffeine, organic acids, and polyphenols were relatively high (Figure 1). This suggested that the sensory attributes of these Nigerian varieties may not match that of high-quality coffees [14,19,23,24] and that there is a need for a concerted coffee breeding program in Nigeria focusing on the improvement of sensory attributes.

Another important aim was to look for variability in metabolites among the Nigerian genotypes. One-way ANOVA identified 66 (~20%) metabolites that differed among the coffee varieties within the 'Niaouli' group (Nia\_1, Nia\_2 and Nia\_3), the 'Kouillou' group (C111 and C36) and the 'Java Robusta' (T1049) group ( $p < 0.01$  and  $FDR < 0.05$ ; Table 1 and Figure S1). Many metabolites could be differentiated between the 'Niaouli,' and, the 'Kouillou' genotypes, and between the 'Niaouli,' and 'Java Robusta' (herein called 'Java') genotypes. However, in contrast, there were few differences in metabolite abundance between 'Kouillou' and 'Java'. As a result, the latter two genotypes were grouped together and described as 'Kouillou/Java.'

**Table 1.** Metabolomic markers that could be used to differentiate 'Java/Kouillou' from 'Niaouli' genotypes as detected by one-way ANOVA and post-hoc analyses ( $f > 4.0$ ;  $p < 0.01$ ;  $-\text{Log}_{10}(p) > 2.0$ ;  $FDR < 0.05$ ). Fisher's LSD identified 'Java/Kouillou' to be higher in content in the metabolites listed here compared to 'Niaouli.' The  $f$ -value is derived from the  $F$ -statistic test for significance, the  $p$ -value is used to test variability between two groups,  $-\text{LOG}_{10}(p)$  determines the significant levels, such as \*, \*\*, and \*\*\*. The False discovery rate (FDR) assists in testing significant result from  $p$ -value.

Metabolomic Markers	$f$ . Value	$p$ . Value	$-\text{Log}_{10}(p)$	FDR
6404	37.478	$1.41 \times 10^{-10}$	9.8511	$4.30 \times 10^{-8}$
125788	23.897	$1.35 \times 10^{-8}$	7.8688	$2.06 \times 10^{-6}$
Citramalic acid	21.799	$3.29 \times 10^{-8}$	7.4826	$2.96 \times 10^{-6}$
Tryptophan	21.428	$3.88 \times 10^{-8}$	7.4112	$2.96 \times 10^{-6}$
34007	14.774	$1.17 \times 10^{-6}$	5.9327	$4.29 \times 10^{-5}$
Palatinitol	14.587	$1.30 \times 10^{-6}$	5.8846	$4.29 \times 10^{-5}$
134465	14.39	$1.47 \times 10^{-6}$	5.8333	$4.29 \times 10^{-5}$
2193	14.304	$1.55 \times 10^{-6}$	5.8109	$4.29 \times 10^{-5}$
Erythritol	13.379	$2.74 \times 10^{-6}$	5.5625	$6.96 \times 10^{-5}$
134464	13.05	$3.38 \times 10^{-6}$	5.4711	$7.36 \times 10^{-5}$
Threitol	12.92	$3.68 \times 10^{-6}$	5.4344	$7.48 \times 10^{-5}$
1-methylgalactose NIST	12.663	$4.35 \times 10^{-6}$	5.3615	$8.29 \times 10^{-5}$
127358	12.044	$6.58 \times 10^{-6}$	5.1815	0.0001
Gluconic acid	11.973	$6.91 \times 10^{-6}$	5.1606	0.0001
2-hydroxyglutaric acid	11.239	$1.15 \times 10^{-5}$	4.9383	0.000146
3182	11.072	$1.30 \times 10^{-5}$	4.8861	0.000152
Maleic acid	10.247	$2.39 \times 10^{-5}$	4.6219	0.000241
Sophorose	10.215	$2.45 \times 10^{-5}$	4.6114	0.000241
16594	9.5222	$4.18 \times 10^{-5}$	4.3787	0.000399
4850	9.3508	$4.79 \times 10^{-5}$	4.3195	0.000443
125830	8.252	0.000119	3.9236	0.00107
Butane-2,3-diol NIST	8.0335	0.000144	3.8414	0.001253
1,2-anhydro-myo-inositol NIST	8.0039	0.000148	3.8301	0.001253
Pseudo uridine	7.6369	0.000205	3.6888	0.001688
6-deoxyglucitol	7.2396	0.000294	3.5316	0.00236
Mannitol	7.0457	0.000352	3.4532	0.002755
102728	6.4686	0.000612	3.2131	0.004446
133018	5.6757	0.001363	2.8655	0.009037
Hexadecylglycerol NIST	4.8794	0.003208	2.4938	0.019967
Arachidic acid	4.5156	0.004832	2.3159	0.026794
Sorbitol	4.4837	0.005011	2.3001	0.02724
trans-4-hydroxyproline	4.0293	0.008518	2.0697	0.042538
beta-gentiobiose	3.8623	0.010403	1.9829	0.048073

Table 1. Cont.

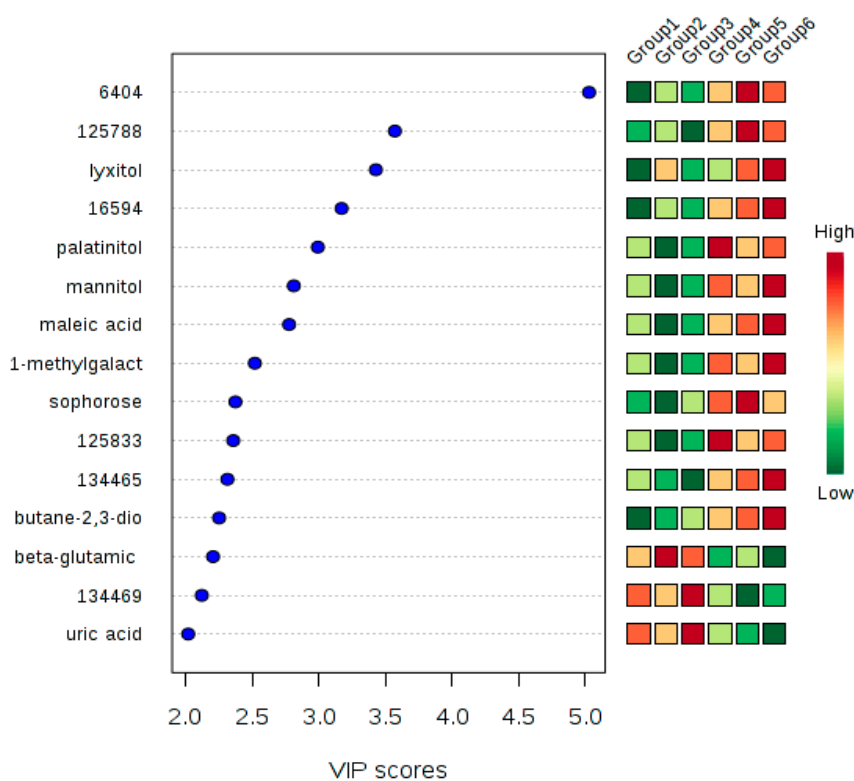
Metabolomic Markers	f. Value	p. Value	−Log10(p)	FDR
Fisher's Least Square Difference (LSD) identified 'Niaouli' to be higher in content in the metabolites listed here compared to 'Java/Kouilou'.				
Threonine	16.164	$5.27 \times 10^{-7}$	6.2779	$3.22 \times 10^{-5}$
Uric acid	15.162	$9.30 \times 10^{-7}$	6.0315	$4.29 \times 10^{-5}$
Normicotine	13.222	$3.03 \times 10^{-6}$	5.5189	$7.10 \times 10^{-5}$
Adipic	12.564	$4.64 \times 10^{-6}$	5.3331	$8.33 \times 10^{-5}$
17094	12.205	$5.90 \times 10^{-6}$	5.229	$9.47 \times 10^{-5}$
Pentitol	11.778	$7.90 \times 10^{-6}$	5.1025	0.000109
5-hydroxynorvaline NIST	11.414	$1.02 \times 10^{-5}$	4.992	0.000135
Tyrosol	11.076	$1.30 \times 10^{-5}$	4.8875	0.000152
Proline	10.719	$1.68 \times 10^{-5}$	4.7747	0.00019
Methanolphosphate	10.547	$1.91 \times 10^{-5}$	4.7195	0.000201
Trisaccharide	5.7878	0.001214	2.9159	0.008225
110009	4.2989	0.006203	2.2074	0.032621
Isocitric acid	4.114	0.007704	2.1133	0.039825
Fisher's LSD identified 'Java' to be higher in content in the metabolites listed here compared to 'Kouilou'				
Lyxitol	12.382	$5.24 \times 10^{-6}$	5.2807	$8.88 \times 10^{-5}$
Glycerol	5.4388	0.001748	2.7574	0.01111

## 2.2. Metabolomic Markers for Differentiating Genotypes

Metabolomic markers can be useful in differentiating genotypes grown in different regions and in selecting genotypes with metabolite profiles for use in breeding programs [10,25]. Such metabolites were identified in this study by using the analysis of variance (ANOVA), and more potential markers were uncovered using variable importance in projection (VIP) at a Partial Least Square (PLS) regression value above 2 (Figure 2). Most of the discriminatory metabolomic markers were in abundance in 'Kouilou/Java' compared to 'Niaouli' (Table 1; Figure 2). These distinct metabolites were mainly sugar derivatives, while the others were organic acids, amino acids and amines. The most distinctive metabolomic marker differentiating 'Niaouli' from 'Kouilou/Java' was an unknown compound '6404,' identified using both the Significant Analysis of Microarrays (SAM) and one-way ANOVA analyses. It was higher in 'Kouilou/Java' than 'Niaouli'. However, its role in flavor characteristics is unknown and needs to be further studied.

**Sugar derivatives.** These are compounds that are usually derived from monosaccharides, but which have undergone further chemical modification. They include sugar alcohols, sugar acids, amino sugars and deoxy sugars. Of these, the sugar alcohol palatinitol was higher in 'Kouilou/Java' compared to 'Niaouli' (Table 1; Figure 2) and could therefore serve as a potential biomarker to differentiate between these genotypes. The farmers' cultivated accessions, which were all of the 'Niaouli' group, had low palatinitol content, and this metabolite can be used to distinguish among *C. canephora* varieties in Nigeria ( $p = 1.30 \times 10^{-6}$ ).

**Amino acids.** Generally, amino acids make positive contributions to coffee hedonistic value. They react favorably with sugars during the Maillard reaction, producing pleasant aroma precursors [26]. Tryptophan, proline, and threonine were the amino acids also identified as potential biomarkers. Threonine and proline have high-quality attributes [27], while tryptophan is linked to low bean quality; it is a specific marker linked with bean immaturity, and the presence of high levels negatively impact the flavor of roasted coffee [28] Tryptophan levels were higher, while proline and threonine were lower ( $p = 3.88 \times 10^{-8}$ ), in 'Kouilou/Java' compared to 'Niaouli'. This amino acid profile suggested that 'Kouilou/Java' has an inferior amino acid profile compared to 'Niaouli' (Table 1). Notably, the farmers' accessions, which were all of the 'Niaouli' genotypes, were characterized by high threonine and proline with  $p$ -values of  $1.68 \times 10^{-5}$  and  $5.27 \times 10^{-7}$ , respectively.



**Figure 2.** Potential metabolomic markers differentiating ‘Niaouli’ (genotypes groups 1–3) from ‘Kouillou’ (groups 4–5) and ‘Java’ (group 6). The variable importance in projection (VIP) scores on the x-axis provide an estimate of the contribution of a given predictor (metabolites shown on the y-axis) to the Partial Least Square (PLS) regression above 2. The higher the VIP score, the better the metabolite is as a predictor of the discrimination among genotypes.

**Polyphenols.** Coffee is one of the most important sources of polyphenols. The polyphenols with the potential to act as biomarkers of Nigerian *C. canephora* coffee beans were tyrosol, gluconic acid and nornicotine. Tyrosol and nornicotine were high in ‘Niaouli’ compared to ‘Kouillou/Java’ (Table 1). Tyrosol is an essential polyphenol in olive oil which protects the actin filament network from oxidized Low-Density Lipoproteins. The tyrosol derived from coffee could also serve a protective function [29].

**Fatty acids.** In coffee beans, fatty acids contribute to the desirable aroma precursors of roasted coffee beans [30] but high fatty acids metabolites are linked to low grade (quality) coffee [29]. We found that hexadecyl glycerol and arachidic acid were statistically higher in ‘Kouillou/Java’ compared to ‘Niaouli’. Interestingly, glycerol was one of the few metabolites that could differentiate between ‘Java’ and ‘Kouillou’ and it was higher in the latter, compared to the former (Table 1).

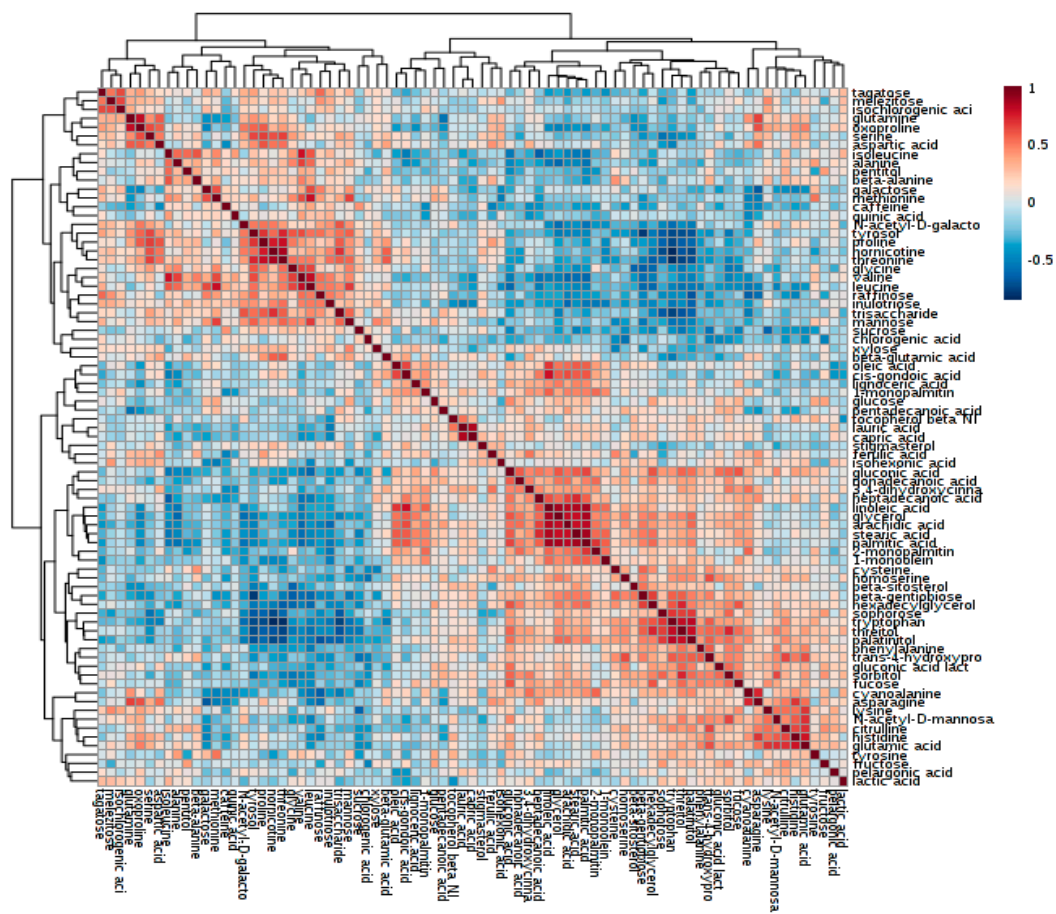
### 2.3. Metabolite-to-Metabolite Correlations and Their Potential Influence on Cup Quality and Other Beneficial Traits

The relative abundance of multiple metabolites and their interaction greatly affects cup quality and other beneficial properties of coffee [31]. To study these interactions, heatmap analyses of metabolite-to-metabolite correlations were drawn, identifying metabolites that are positively and negatively correlated in abundance across genotypes (Figure 3a,b).

**Caffeine.** This metabolite correlated moderately with sucrose ( $r^2 = 0.42$ ; Table 2). This moderate association of sucrose with caffeine should have the effect of reducing the bitterness caused by the high caffeine content of many *C. canephora* types [31]. Although caffeine contributes to bitterness in coffee, it has long been valued for its beneficial cognitive and other health effects [32], even whether there are negative consequences for overconsumption [33].

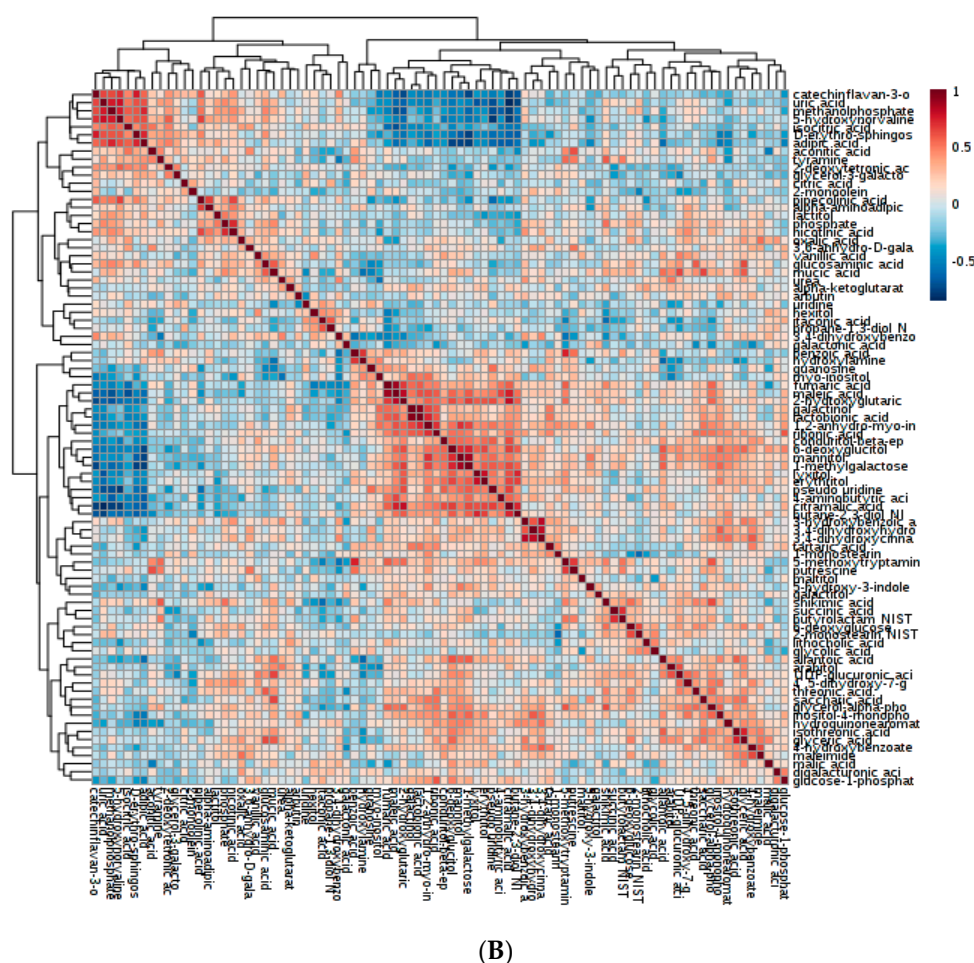
**Table 2.** Metabolites showing similar correlation pattern with sucrose and caffeine ( $p < 0.05$ ).

Metabolite Identified as Markers	Correlation Coefficient ( $r^2$ )	
	Sucrose	Caffeine
Caffeine	0.4079	-
16548	0.4044	0.5182
6404	-0.3705	-
Palatinitol	-0.4961	-
68	0.4991	0.4502
Tryptophan	-0.4599	-



(A)

Figure 3. Cont.



**Figure 3.** Heatmaps showing metabolite-to-metabolite correlation patterns. Two maps were developed for clarity (A). Heatmap showing the correlation among sugars, amino acids, polyphenols and fatty acids. (B). Heatmap showing the correlative pattern among organic acids, sugar derivatives, amines and vitamins.

**Sugars.** Sucrose and fructose were the two major sugars found in the coffee beans of *C. canephora* cultivated in Nigeria (Supplementary Table S1). High-quality coffee usually accumulates high levels of sucrose [9]. The positive Pearson's correlation coefficient for sucrose and caffeine observed in this study was in contradiction to the negative correlation found by Caporaso et al. [34]. In our work, sucrose and caffeine were found to be positively correlated with two unknown compounds '16548' and '68' (Table 2). Interestingly, a negative correlation was found between sucrose and palatinitol (Table 2); this could be explained by the fact that sucrose is a precursor of palatinitol biosynthesis [35]. Metabolite '6404' is potentially a metabolomic marker, differentiating cultivated accession 'Niaouli' from 'Java/Kouillou' (Table 1). This metabolite could contribute to low cup quality as it was negatively correlated with sucrose ( $r = -0.37$ ).

**Phenolic acids.** Chlorogenic acid and quinic acid were the most abundant polyphenols (Supplementary Table S1). These compounds cause a high degree of bitterness and could contribute to the low cup quality [36] of Nigerian *C. canephora*, since they are higher in abundance relative to sucrose, amino acids and the fatty acids, compounds that are associated with high cup quality [27,37,38]. However, while chlorogenic acids negatively affect taste, they do have beneficial health properties as antioxidants.

**Amino acids.** Several amino acids identified as good cup quality precursors were found to be positively associated in coffee genotypes used in the study. These included aspartic acid, glutamine,



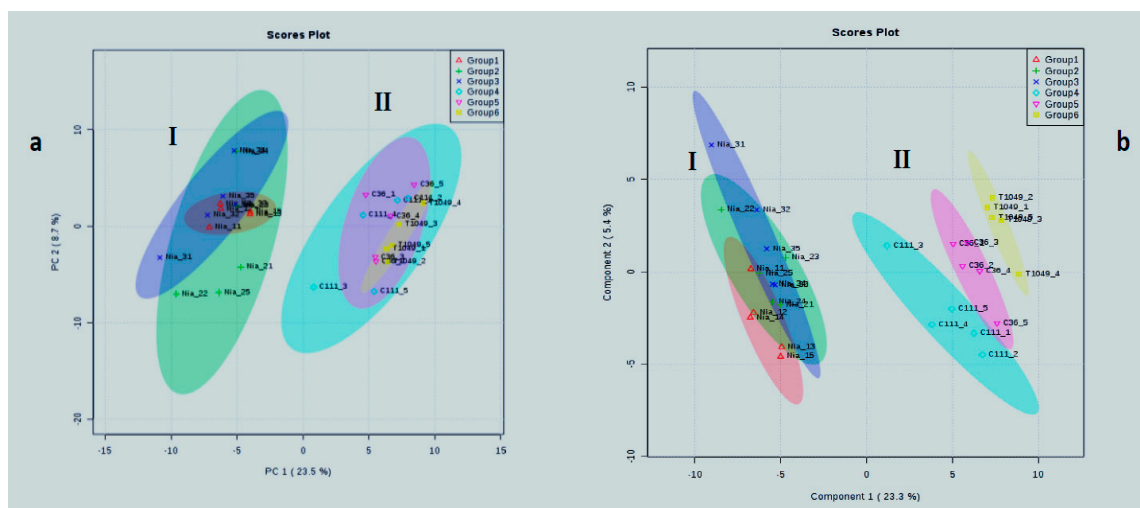
oxoproline, serine, N-acetyl-D-galactosamine, beta-glutamic acid, proline and threonine (Figure 3a). Of these, it is known that proline produces pleasant, flowery and fragrant aromas; aspartic acid and serine generates pleasant and fruity aromas; and threonine produces a pleasant caramel-like odor [27]. Aspartic acid and glutamic acid could be explored as potential markers differentiating Nigerian *C. canephora* genotypes. These amino acids accumulated to relatively high levels (Supplementary Table S1) compared to that reported by Arnold et al., [39], where asparagine and glutamic acid were the two major amino acids found in the coffee beans in their study.

**Fatty Acids.** Arachidic acid, stearic acid, palmitic acid, linoleic acid and glycerol are fatty acids linked to coffee with better sensory qualities [30] and, in this study, we found these metabolites to be positively correlated with each other (Supplementary Table S2).

**Sugar alcohols.** Apart from cup quality improvement, the identification of genotypes with high drought tolerance traits is of great importance in the face of climate change. Here, galactinol and beta-gentiobiose were positively and highly correlated, ( $r^2 = 0.98$ ) (Figure 3b). Galactinol, together with raffinose, are essential in protecting plant cells from the oxidative damage caused by various types of stress conditions [40]. Galactinol and beta-gentiobiose had the highest positive correlation coefficient ( $r^2 = 0.97$ ), followed by lactobionic acid and beta-gentiobiose ( $r^2 = 0.96$ ). Galactinol synthase is a key enzyme in the synthesis of the raffinose family of oligosaccharides, which function as osmoprotectants in plant cells [40]. The Pearson correlation was statistically significant ( $p < 0.05$ ) among different chemicals.

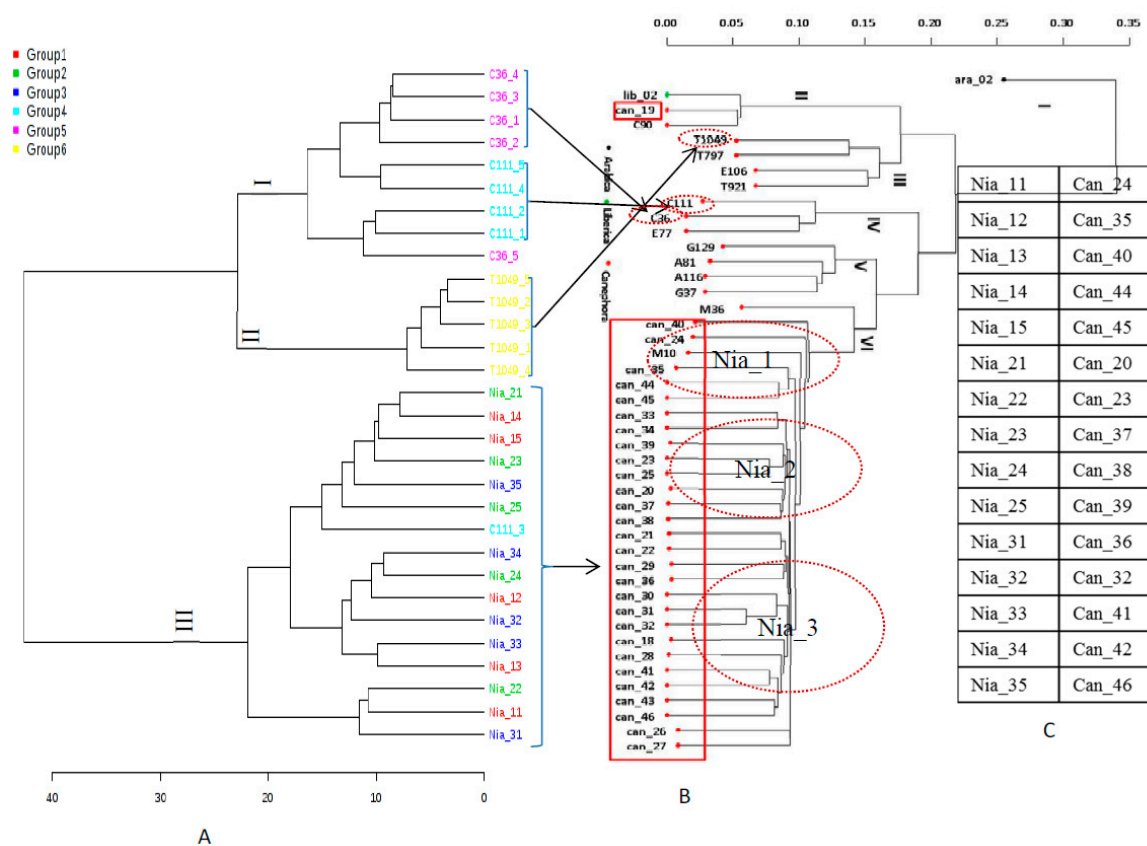
#### 2.4. Metabolite and Genomic Diversity within and among Varieties

Principal Component Analysis (PCA) and Partial Least Squares (PLS) both detected group differences. PCA is an unsupervised approach. PLS, in contrast, is a supervised approach that seeks to maximize separation between groups by reducing within group variation [41,42]. This explains why clear separation between ‘Java’ and ‘Kouillou’ was possible with Partial Least Squares-Discriminant Analysis (PLS-DA), but not with PCA (Figure 4). Both PCA and PLS detected 32.2% and 28.7% variations, respectively (Figure 4a,b). They also indicated that there are two metabolomic diverse groups, which we designated as clusters I and II. Cluster I comprised of all the ‘Niaouli’ genotypes, while cluster II was made up of the ‘Java’ and ‘Kouillou’ genotypes (Figure 4a,b).



**Figure 4.** Multivariate classification of Nigerian *Coffea canephora* coffee genotypes. (a) Principal Component Analysis (PCA) and (b) Partial Least Squared (PLS) Analysis. PLS is a supervised method that minimizes within group variability and maximizes intergroup variability to achieve the greatest separation (discrimination) among groups.

Because of post-transcriptional modification, a change in metabolite level may result in phenotypic changes more than alterations in DNA sequence [10,43]. Still, it was of interest to compare how the genotypes would be classified based on SNPs vs. based on their metabolite profiles. Genomic analysis through the characterization of 433,048 SNPs grouped the genotypes into three genetic units (clusters III, IV and VI) (Figure 5a). Hierarchical analysis of the GC-MS data also led to a similar clustering of the genotypes i.e., (I, II and III) (Figure 5a). Among the coffee genotypes studied, genetic diversity based on SNPs can be linked to chemical or metabolomics diversity, as the data proved to be complementary. Single Nucleotide Polymorphism analysis of both the phenolic and vitamin E pathways revealed metabolite-specific genetic diversity among the rice varieties examined [44].

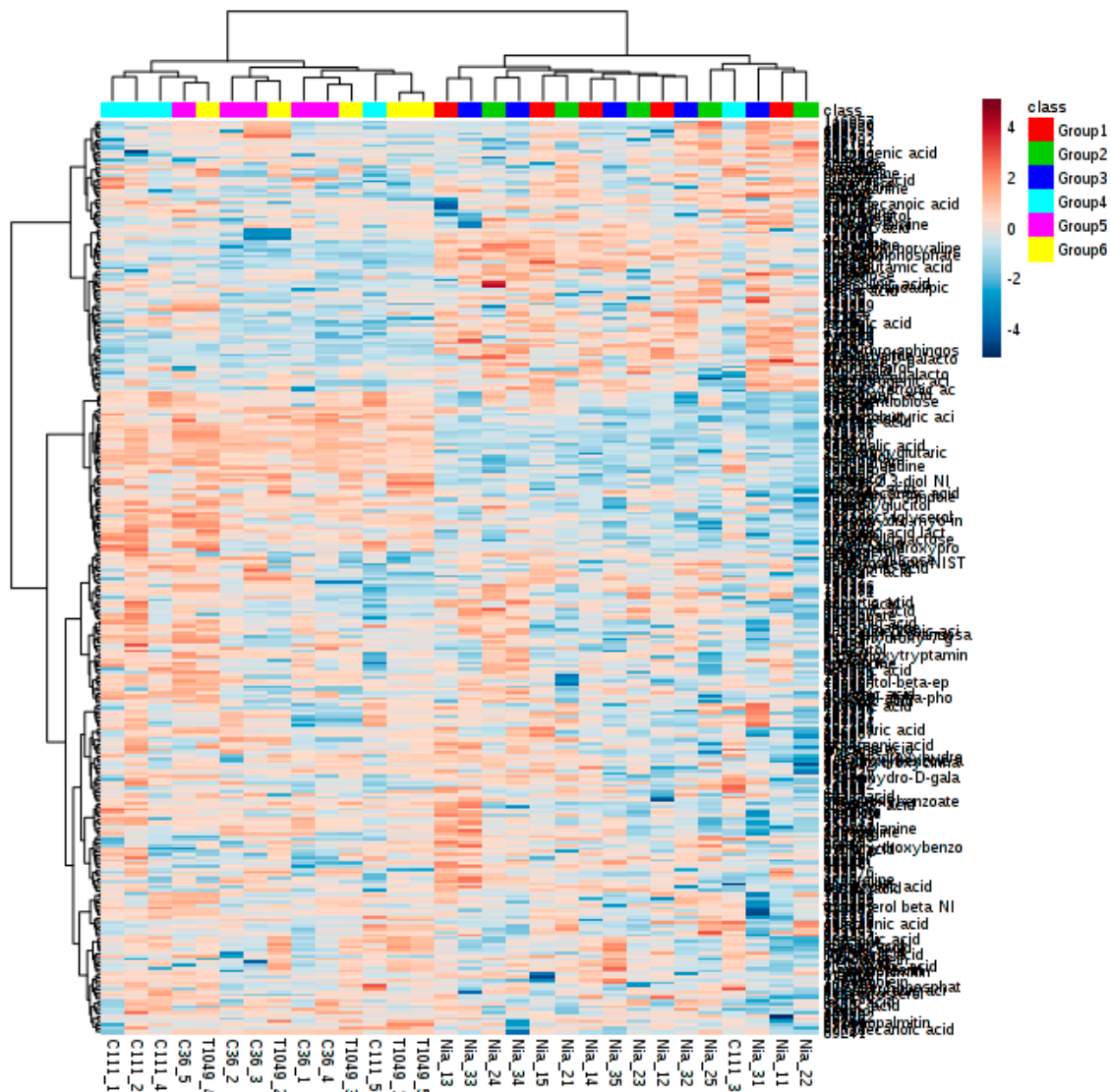


**Figure 5.** Comparing the metabolite profiles and Single Nucleotide Polymorphism (SNP) diversity of *C. canephora*. (A) Hierarchical clustering of the metabolomics data derived from 30 coffee samples. (B) Hierarchical clustering of Single Nucleotide Polymorphism-Genotype-By-Sequencing data derived from 47 coffee samples, including those used in the metabolomics analysis. Dotted red circles denote genotypes used for metabolomics analysis. The red rectangular box contains the ‘Niaouli’ genotypes. (C) This table matches the ‘Niaouli’ genotypes in (A) to those in (B), where Nia\_11 is the same as Can\_24.

### 2.5. Genotypes with Favorable Bean Quality Traits

Our analysis showed that some genotypes contained high and low levels of some important chemicals linked to coffee quality and health benefits (Table 3). High levels of sucrose and low levels of aminobutyric acid (GABA), quinic acid, choline, acetic acid and fatty acids were reported to be the metabolomic markers of high-grade green coffee [9]. The Nia\_14 and Nia\_15 genotypes of ‘Niaouli’ have a low caffeine content and relatively higher sucrose content (Table 3). Thus, bitterness may be reduced within these genotypes, and the caffeine level may make it suitable to some consumers who have health concerns. Determining whether there are SNP differences in the gene coding for these “favorable” metabolites may point to a genomic cause for differences in metabolite levels among these

genotypes (Table 3) that might be exploited in coffee breeding and coffee functional genomic analysis. Sucrose and caffeine showed a broad range of concentrations (Table 3 and Figure 6). The variability of sucrose was 39-fold, while caffeine was 5-fold, offering targets for breeding programs. A high level of natural variability for caffeine and sucrose in green coffee beans was also observed in other studies [24,33]. Metabolites differentiating high- (*C. arabica*) and low-quality coffee (*C. canephora*) are summarized in Table 4.



**Figure 6.** The hierarchical clustering analysis (HCA) showing metabolite diversity and intensity levels in relation to genotypes. Dark brown (4) represents high intensity, while dark blue (−4) represents very low intensity.

**Table 3.** Genotypes grouped based on the relative levels of the key metabolites identified from hierarchical analysis (Figure 6).

Metabolites	Metabolite Levels				
	Very High	High	Medium	Low	Very Low
Caffeine		Nia_24, Nia_25		Nia_15, Nia_14	
CGA		Nia_25		Nia_14	C111_2
Sucrose		Nia_25, Nia_22, Nia_15, Nia_14	Nia_22	Nia_24	C36_5
Quinic acid		Nia_25, Nia_31, Nia_34	C36_1	Nia_14	
Butane-2,3 diol		C36_2		All Nia except Nia_25	Nia_22
Saccharic acid		Nia_15		Nia_22, Nia_11, Nia_31	
Ferulic acid		Nia_33			
Tryptophan		C36_5		Nia_24	
Putrescine	C36_5			Nia_22, Nia_25, Nia_32, Nia_33	C111_5
Proline		Nia_33			C111_3

Very high (4 and above); high (2 to 4); medium (0 to 2); low (0 to -2); and very low (-2 to -4).

**Table 4.** Relative levels of metabolites that contribute to coffee quality based on published data.

Coffee Type and Perceived Quality	Relative Metabolite Levels			Source
	High	Low		
Palm civet (Superior)	Citric acid, malic acid, and glycolic acid	Quinic acid, caffeine, and caffeic acid		[21]
<i>Coffea arabica</i> (Good)	Sucrose, triglyceride, and threonine, proline	Caffeine, chlorogenic acid, aminobutyric acid (GABA), quinic acid, choline, acetic acid, and fatty acid		[9,19,27]
<i>C. canephora</i> (Poor)	Caffeine and chlorogenic acid	Sucrose		[45,46]

Levels of sucrose are exceptionally higher in the ‘Niaouli’ genotypes, with the highest concentration found in Nia\_25 and Nia\_22, while those with the lowest concentration were C111\_1 and C36\_5. It could be deduced that the genetic group ‘Nia\_2’ has high sucrose synthase enzyme and could be linked to SNPs changes. A low sucrose concentration in *C. canephora* green coffee beans has been attributed to the limited capacity of sucrose synthase to synthesize sucrose at the final stage of coffee grain development [45–47]. It will be interesting to further evaluate the sucrose synthase sequence of these genotypes having high and low sucrose content to determine whether there is a SNP change at this locus.

### 3. Materials and Methods

#### 3.1. Single Nucleotide Polymorphism Genotype-By-Sequencing Analysis

DNA from 48 samples comprising 18 accessions from Cocoa Research Institute of Nigeria (CRIN) and 30 accessions from farmers’ fields were analyzed. The CRIN accessions consisted of five *Coffea* species including *Coffea arabica*, *Coffea abeokutae*, *Coffea liberica*, *Coffea stenophylla* and *Coffea canephora*.

The *C. canephora* varieties used were ‘Kouillou’, ‘Gold Coast’, ‘Java Robusta’, ‘Niaouli’, ‘Uganda’ and ‘Java Robusta Ex Gamba’. More information about the methods used for this analysis is available [7].

The farmers’ accessions known to be the ‘Niaouli’ variety were grouped into 3 genotypes (Nia\_1, Nia\_2 and Nia\_3) based on the hierarchical clustering of 433,650 SNPs using SNPRelate software (<http://www.Rproject.org>) [48]. These genotypes and those within the ‘Kouillou’ and ‘Java Robusta’ groups were used for metabolomics analysis. These varieties (‘Niaouli’, ‘Kouillou’ and ‘Java Robusta’) represented three genetic structures detected by GBS-SNP analysis on *C. canephora* repository in Nigeria [7]. These genotypes were selected because they were used to establish the pioneer model polyclonal seed garden plot, which will be replicated in farmers’ fields.

### 3.2. GC-MS Analysis

The genotypes used for metabolomics were cultivated *C. canephora* accessions in Nigeria [7], consisting of ‘Niaouli’, ‘Kouillou’ and ‘Java Robusta’ varieties. They were further classified into one of six groups based on the result from SNP-GBS analysis [7]. ‘Niaouli’ is comprised of three genotypes: Nia\_1, Nia\_2 and Nia\_3, (classified as Groups 1, 2 and 3, respectively), ‘Kouillou’ is comprised of two genotypes: C111 and C36, (classified as Groups 4, and 5, respectively) and, ‘Java Robusta’ is comprised of one genotype: T1049, and was classified as Group 6. There were five replicates for each group (genotypes), giving a total of 30 samples (Table 5).

**Table 5.** Sample genotypes and symbols.

Variety	‘Niaouli’			‘Kouilou’		‘Java Robusta’
	Group 1 (Nia_1)	Group 2 (Nia_2)	Group 3 (Nia_3)	Group 4 (C111)	Group 5 (C36)	Group 6 (T1049)
Sample Symbols	Nia_11	Nia_21	Nia_31	C111_1	C36_1	T1049_1
	Nia_12	Nia_22	Nia_32	C111_2	C36_2	T1049_2
	Nia_13	Nia_23	Nia_33	C111_3	C36_3	T1049_3
	Nia_14	Nia_24	Nia_34	C111_4	C36_4	T1049_4
	Nia_15	Nia_25	Nia_35	C111_5	C36_5	T1049_5

Reddish matured (ripened), coffee bean (Figure 7b) of these genotypes were collected in ice bags and immediately transferred to  $-80\text{ }^{\circ}\text{C}$ . The endosperms of the coffee bean (Figure 7c) were excised using a sterile blade and re-transferred to  $-80\text{ }^{\circ}\text{C}$ . These endosperms were lyophilized, ground into powder with Udy mill (Udy Corporation) and sealed prior to metabolomics analysis.



**Figure 7.** *C. canephora* coffee beans. Two maturity stages of the beans, (a) unripe, and (b) ripe are shown. Note, only the ripe beans were used for the analysis. (c) The endosperm, which was the portion of the coffee bean dissected and used to extract metabolites for chemical analysis.

The metabolomics analysis was performed according to Fiehn et al. [49]. The analyte was extracted from the sample using a solvent containing isopropanol/acetonitrile/water at the volume ratio of 3:3:2. The supernatant was concentrated in a Centrivap cold trap vacuum concentrator (<http://www.labconco.com>) at room temperature for 4 h. The extracts were immediately derivatized for GC–Time-of-Flight

(TOF) mass spectrometry analysis by adding 90  $\mu\text{L}$  of N-methyl-N-trimethylsilyltrifluoroacetamide and 1% (v/v) trimethylchlorosilane (1 mL bottles; Pierce) to the extract and shaken at 37 °C for 30 min. The reaction mixture was transferred to a 2 mL clear glass auto-sampler vial with micro-insert (Agilent; <http://www.agilent.com>) and closed using an 11 mm T/S/T crimp cap.

Samples (0.5  $\mu\text{L}$ ) were injected between 2 and 24 h after derivatization in an Agilent 6890 gas chromatograph controlled by using LecoChromaTOF software version 2.32 (<http://www.leco.com>). The analytical GC column was Restek corporation Rtx-5Sil MS (30 m length  $\times$  0.25 mm internal diameter with 0.25  $\mu\text{m}$  film made of dimethyl/diphenylpolysiloxane at the volume ratio 95:5) and the mobile phase used was Helium.

The data was acquired on a Mass spectrometry Instrument, a Leco Pegasus IV time-of-flight mass spectrometer controlled using LecoChromaTOF software version 2.32, at a mass resolving power  $R = 600$  from  $m/z$  85–500 at 20 spectra per second, and 1550 V detector voltage, without turning on the mass defect option. Recording ended after 1200 s.

### 3.3. Statistical Analysis of Metabolomics Data

All metabolites, including those identified and those not identified using the National Institute of Standards and Technology mass spectral library, were used for data analysis. Both univariate and multivariate statistical approaches were performed with metaboAnalystR [50]. One-way Analysis of Variance (ANOVA) test was performed to ascertain the significant variables, and they were expressed as  $f$ - and  $p$ -values. The level of statistical significance ( $-\text{Log}_{10}(p)$ ) was determined, followed by post-hoc analyses to correct the  $p$ -value and thus generate the False Discovery Rate (FDR). Fisher's least significant difference method (LSD) was used to identify groups that differ in their metabolite profile. Pearson's correlation coefficient between the metabolites was calculated using metaboAnalystR [50]. Principal Component Analysis (PCA) and Partial Least Squares-Discriminant Analysis (PLS-DA) were performed using the prcomp package within the R statistical package [51]. Prominent potential discriminatory metabolites were identified with variable importance in projection (VIP) scores at a PLS regression value above 2. Calculations for these analyses were based on singular value decomposition [50]. Hierarchical Clustering Analysis was used to classify genotypes by employing Pearson's correlation as a similarity measure. The clustering algorithm used was Ward's linkage, which minimizes the sum of squares of any two clusters [50]. The metabolomic analyses results were compared to hierarchical clustering of SNP data obtained by SNP relate [50].

## 4. Conclusions

For the first time, metabolites in the beans of Nigerian *C. canephora* were evaluated. The farmers' cultivated accessions have unique metabolite profiles and genomic structures compared to the conserved accessions. There was high variability within the varieties for caffeine and sucrose—the two key compounds in coffee. However, the high caffeine, quinic acid and chlorogenic acid may contribute to low quality and bitterness. Genomic technologies such as transgenesis, molecular marker-assisted breeding, genomics, proteomics, and metabolomics should complement traditional breeding efforts for hastening the genetic improvement of coffee. The alteration of major metabolites that contribute to coffee taste, such as caffeine and sucrose, has been one of the strategies used in developing high cup quality coffee. Based on the flavor profile we have assayed, it was found that the ratio of caffeine to sucrose is higher than desirable for good cup quality. It remains to be seen whether there are polymorphisms for known enzymes and proteins that regulate the levels of these compounds among our genotypes. Such data can be obtained by careful examination of our GBS-SNP data, and the utility of this information will be further extended by using them as genetic markers for breeding. By altering this design, the desirable flavor profile will be achieved. It will be of great interest to characterize those genotypes conferring relatively low caffeine, and high sucrose, together with other important quality precursors to determine the cause of such variation and identify some important quality trait markers.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2223-7747/8/10/425/s1>, Table S1: The eight classes of metabolites identified in the Nigerian *C. canephora* coffee genotypes, Table S2: Fatty acids with high Pearson's Coefficient correlative scores ( $r^2 > 0.80$ ;  $p < 0.05$ ), Figure S1: Metabolites detected from one-way Analysis of variance that significantly varied across genotypes.

**Author Contributions:** C.F.A., D.M.B., C.O.I., and R.B. designed the project; D.M.B., C.O.I., and R.B. supervised the project; C.F.A. and D.M.B. analyzed the data; C.F.A., D.M.B. and O.O.O. wrote the manuscript; all authors read and approved the final manuscript.

**Funding:** The Norman E. Borlaug Leadership Enhancement in Agriculture Program (Borlaug LEAP) through a grant to the University of California-Davis by the United States Agency for International Development supported this research. The opinions expressed herein are those of the authors and do not necessarily reflect the views of USAID. Funding was also provided by the Federal Ministry of Agriculture in Nigeria through a grant to the Cocoa Research Institute of Nigeria.

**Acknowledgments:** We appreciate the International Institute of Tropical Agriculture and the University of California Davis, USA, for providing free access to their facilities. We also thank May Thitisaksakul, Karin Alborno, Shaoyun Dong, and Jorge Dubcovsky at UC Davis, USA, for their advice and technical support.

**Conflicts of Interest:** Authors declare no conflict of interest.

## Abbreviations

CRIN, Cocoa Research Institute of Nigeria; SNPs, Single Nucleotide Polymorphisms; GC-MS, Gas Chromatography-Mass Spectrometry; HPLC, High-Performance Liquid Chromatography;  $^1\text{H-NMR}$ , Proton Nuclear Magnetic Resonance; Time of Flight, TOF; NIST, National Institute of Standards and Technology; ANOVA, One-Way Analysis of Variance; FDR, False Discovery Rate; LSD, Least Significant Difference; PCA, Principal Component Analysis; PLS-DA, Partial Least Squares-Discriminant Analysis; VIP, Variable Importance in Projection; SAM, Significant Analysis of Microarrays; R, mass resolving power;  $m/z$ , mass/charge; and V, Voltage.

## References

1. Dicum, G.; Luttinger, N. *The Coffee Book: Anatomy of an Industry from the Crop to the Last Drop*; The New Press: New York, NY, USA, 2006.
2. Oxfam. *The Coffee Market: A Background Study*; Oxfam: London, UK, 2001.
3. Davis, A.P.; Chadburn, H.; Moat, J.; O'Sullivan, R.; Hargreaves, S.; Lughadha, E.N. High extinction risk for wild coffee species and implications for coffee sector sustainability. *Sci. Adv.* **2019**, *5*. [[CrossRef](#)] [[PubMed](#)]
4. Ayegboyin, K.O.; Famaye, A.O.; Akinrinde, E.A.; Adejobi, K.B.; Akanbi, O.S.O. Growth performance and nutrient uptake of *Coffea canephora* Prieur Ex. Froehner grown in contrasting soils. *Int. J. Res.* **2015**, *1*, 25–31.
5. Leroy, T.; Ribeyre, F.; Bertrand, B.; Charmetant, P.; Dufour, M.; Montagnon, C.; Marraccini, P.; Pot, D. Genetics of coffee quality. *Br. J. Plant. Phys.* **2006**, *18*, 229–242. [[CrossRef](#)]
6. Omolaja, S.S. Biochemical Characterization, Improvement and Propagation of *Coffea canephora*. Cocoa Research Institute of Nigeria Annual Report; 2000; pp. 32–35. Available online: <http://www.crin.gov.ng> (accessed on 2 October 2019).
7. Anagbogu, C.F.; Bhattacharjee, R.; Ilori, C.; Tongyoo, P.; Dada, K.E.; Muiyiwa, A.A.; Beckles, D.M. Genetic diversity and re-classification of coffee (*Coffea canephora* Pierre ex A. Froehner) from South Western Nigeria through genotyping-by-sequencing-single nucleotide polymorphism analysis. *Genet. Resour. Crop. Evol.* **2019**, *66*, 685–695. [[CrossRef](#)]
8. Choi, M.Y.; Choi, W.; Park, J.H.; Lim, J.; Kwon, S.W. Determination of coffee origins by integrated metabolomic approach of combining multiple analytical data. *Food Chem.* **2010**, *121*, 1260–1268. [[CrossRef](#)]
9. Kwon, D.J.; Jeong, H.J.; Moon, H.; Kim, H.N.; Cho, J.H.; Lee, J.E.; Hong, K.S.; Hong, Y.S. Assessment of green coffee bean metabolites dependent on coffee quality using a  $^1\text{H-NMR}$ -based metabolomics approach. *Food Res. Int.* **2015**, *67*, 175–182. [[CrossRef](#)]
10. Beckles, D.M.; Roessner, U. Plant metabolomics: Applications and opportunities for agricultural biotechnology. *Plant Biotech. Agric.* **2012**, 1st Edition. 67–81.
11. Flament, I. *Coffee Flavor Chemistry*; John Wiley & Sons Ltd.: New York, NY, USA, 2002; p. 424.
12. Ribeiro, J.S.; Augusto, F.; Salva, T.J.G.; Thomaziello, R.A.; Ferreira, M.M.C. Prediction of sensory properties of Brazilian Arabica roasted coffees by headspace solid phase microextraction-gas chromatography and partial least squares. *Anal. Chim. Acta* **2009**, *634*, 172–179. [[CrossRef](#)]

13. Dessalegn, Y.; Labuschagne, M.T.; Osthoff, G.; Herselman, L. Genetic diversity and correlation of bean caffeine content with cup quality and green bean physical characteristics in coffee (*Coffea arabica* L.). *J. Sci. Food Agr.* **2008**, *88*, 1726–1730. [[CrossRef](#)]
14. Belay, A.; Gholap, A.V. Characterization and determination of chlorogenic acids (CGA) in coffee beans by UV-Vis spectroscopy. *Afr. J. Pure Appl. Chem.* **2009**, *3*, 234–240.
15. Joet, T.; Bertrand, B.; Dussert, S. Environmental Effects on Coffee Seed Biochemical Composition and Quality Attributes: A Genomic Perspective. In Proceedings of the 25th International Conference on Coffee Science, Armenia, Colombia, 8–13 September 2014; pp. 42–49.
16. Bertrand, B.; Boulanger, R.; Dussert, S.; Ribeyre, F.; Berthiot, L.; Descroix, F.; Joet, T. Climatic factors directly impact the volatile organic compound fingerprint in green Arabica coffee bean as well as coffee beverage quality. *Food Chem.* **2012**, *135*, 2575–2583. [[CrossRef](#)] [[PubMed](#)]
17. Oksman-Caldentey, K.M.; Saito, K. Integrating genomics and metabolomics for engineering plant metabolic pathways. *Curr. Opin. Biotech.* **2005**, *16*, 174–179. [[CrossRef](#)] [[PubMed](#)]
18. Duarte, G.S.; Pereira, A.A.; Farah, A. Chlorogenic acids and other relevant compounds in Brazilian coffees processed by semi-dry and wet post-harvesting methods. *Food Chem.* **2010**, *118*, 851–855. [[CrossRef](#)]
19. Farah, A.; Monteiro, M.C.; Calado, V.; Franca, A.S.; Trugo, L.C. Correlation between cup quality and chemical attributes of Brazilian coffee. *Food Chem.* **2006**, *98*, 373–380. [[CrossRef](#)]
20. Arana, V.A.; Medina, J.; Alarcon, R.; Moreno, E.; Heintz, L.; Schafer, H.; Wist, J. Coffee's country of origin determined by NMR: The Colombian case. *Food Chem.* **2015**, *175*, 500–506. [[CrossRef](#)]
21. Jumhawan, U.; Putri, S.P.; Yusianto; Bamba, T.; Fukusaki, E. Application of gas chromatography/flame ionization detector-based metabolite fingerprinting for authentication of Asian palm civet coffee (Kopi Luwak). *J. Biosci. Bioeng.* **2015**, *120*, 555–561. [[CrossRef](#)]
22. Dong, W.J.; Tan, L.H.; Zhao, J.P.; Hu, R.S.; Lu, M.Q. Characterization of Fatty Acid, Amino Acid and Volatile Compound Compositions and Bioactive Components of Seven Coffee (*Coffea robusta*) Cultivars Grown in Hainan Province, China. *Molecules* **2015**, *20*, 16687–16708. [[CrossRef](#)]
23. CIRAD. *Genes Involved in Coffee Quality*; Montpellier University: Montpellier, France, 2006.
24. Ky, C.L.; Louarn, J.; Dussert, S.; Guyot, B.; Hamon, S.; Noiro, M. Caffeine, trigonelline, chlorogenic acids and sucrose diversity in wild *Coffea arabica* L. and *C-canephora* P. accessions. *Food Chem.* **2001**, *75*, 223–230. [[CrossRef](#)]
25. Steinfath, M.; Strehmel, N.; Peters, R.; Schauer, N.; Groth, D.; Hummel, J.; Steup, M.; Selbig, J.; Kopka, J.; Geigenberger, P.; et al. Discovering plant metabolic biomarkers for phenotype prediction using an untargeted approach. *Plant. Biotech. J.* **2010**, *8*, 900–911. [[CrossRef](#)]
26. Liu, C.; Yang, N.; Yang, Q.; Ayed, C.; Linforth, R.; Fisk, I.D. Enhancing Robusta coffee aroma by modifying flavour precursors in the green coffee bean. *Food Chem.* **2019**, *281*, 8–17. [[CrossRef](#)] [[PubMed](#)]
27. Wong, K.H.; Aziz, S.A.; Mohamed, S. Sensory aroma from Maillard reaction of individual and combinations of amino acids with glucose in acidic conditions. *Int. J. Food Sci. Tech.* **2008**, *43*, 1512–1519. [[CrossRef](#)]
28. Setoyama, D.; Iwasa, K.; Seta, H.; Shimizu, H.; Fujimura, Y.; Miura, D.; Wariishi, H.; Nagai, C.; Nakahara, K. High-Throughput Metabolic Profiling of Diverse Green *Coffea arabica* Beans Identified Tryptophan as a Universal Discrimination Factor for Immature Beans. *PLoS ONE* **2013**, *8*. [[CrossRef](#)] [[PubMed](#)]
29. Giovannini, C.; Straface, E.; Modesti, D.; Coni, E.; Cantafora, A.; De Vincenzi, M.; Malorni, W.; Masella, R. Tyrosol, the major olive oil biophenol, protects against oxidized-LDL-induced injury in Caco-2 cells. *J. Nutr.* **1999**, *129*, 1269–1277. [[CrossRef](#)] [[PubMed](#)]
30. Figueiredo, L.P.; Borém, F.M.; Ribeiro, F.C.; Giomo, G.S.; Taveira, J.H.S.; Malta, M.R. Fatty acid profiles and parameters of quality of specialty coffees produced in different Brazilian regions. *Afr. J. Agric. Res.* **2015**, *10*, 3484–3493.
31. Calvino, A.M.; Garciamedina, M.R.; Comettomuniz, J.E. Interactions in Caffeine-Sucrose and Coffee-Sucrose Mixtures - Evidence of Taste and Flavor Suppression. *Chem. Senses* **1990**, *15*, 505–519. [[CrossRef](#)]
32. Nehlig, A.; Daval, J.L.; Debry, G. Caffeine and the Central-Nervous-System - Mechanisms of Action, Biochemical, Metabolic and Psychostimulant Effects. *Brain Res. Rev.* **1992**, *17*, 139–169. [[CrossRef](#)]
33. Kumar, V.; Kaur, J.; Panghal, A.; Kaur, S.; Handa, V. Caffeine: A boon or bane. *Nutr. Food Sci.* **2018**, *48*, 61–75. [[CrossRef](#)]



34. Caporaso, N.; Whitworth, M.B.; Grebby, S.; Fisk, I.D. Non-destructive analysis of sucrose, caffeine and trigonelline on single green coffee beans by hyperspectral imaging. *Food Res. Int.* **2018**, *106*, 193–203. [[CrossRef](#)]
35. Hellmers, F.; Takors, R.; Thum, O. Robust enzyme immobilizates for industrial isomalt production. *Mol. Catal.* **2018**, *445*, 293–298. [[CrossRef](#)]
36. Toci, A.T.; Farah, A. Volatile compounds as potential defective coffee beans' markers. *Food Chem.* **2008**, *108*, 1133–1141. [[CrossRef](#)] [[PubMed](#)]
37. *Coffee: Recent Developments*; Clarke, R.J.; Vitzthum, O.G. (Eds.) Wiley-Blackwell Science Ltd.: Oxford, UK, 2001; pp. 50–67.
38. Grosch, W.; Czerny, M.; Mayer, F.; Moors, A. Sensory Studies on the Key Odorants of Roasted Coffee. In *Caffeinated Beverages*; American Chemical Society: Washington, DC, USA, 2000; Volume 754, pp. 202–209.
39. Arnold, U.; Ludiwig, E.; Kuhn, R.; Moschwitzer, U. Analysis of Free Amino-Acids in Green Coffee Beans. 1. Determination of Amino-Acids after Precolumn Derivatization Using 9-Fluorenylmethylchloroformate. *Z Lebensm Unters For.* **1994**, *199*, 22–25. [[CrossRef](#)] [[PubMed](#)]
40. Nishizawa, A.; Yabuta, Y.; Shigeoka, S. Galactinol and raffinose constitute a novel function to protect plants from oxidative damage. *Plant. Physiol.* **2008**, *147*, 1251–1263. [[CrossRef](#)] [[PubMed](#)]
41. Lee, L.C.; Liong, C.Y.; Jemain, A.A. Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: A review of contemporary practice strategies and knowledge gaps. *Analyst* **2018**, *143*, 3526–3539. [[CrossRef](#)] [[PubMed](#)]
42. Stamova, B.S.; Roessner, U.; Suren, S.; Laudencia-Chinguanco, D.; Bacic, A.; Beckles, D.M. Metabolic profiling of transgenic wheat over-expressing the high-molecular-weight Dx5 glutenin subunit. *Metabolomics* **2009**, *5*, 239–252. [[CrossRef](#)]
43. Fiehn, O. Metabolomics - the link between genotypes and phenotypes. *Plant. Mol. Biol.* **2002**, *48*, 155–171. [[CrossRef](#)]
44. Heuberger, A.L.; Lewis, M.R.; Chen, M.H.; Brick, M.A.; Leach, J.E.; Ryan, E.P. Metabolomic and Functional Genomic Analyses Reveal Varietal Differences in Bioactive Compounds of Cooked Rice. *PLoS ONE* **2010**, *5*, e12915. [[CrossRef](#)]
45. Privat, I.; Foucher, S.; Prins, A.; Epalle, T.; Eychenne, M.; Kandalaf, L.; Caillet, V.; Lin, C.W.; Tanksley, S.; Foyer, C.; et al. Differential regulation of grain sucrose accumulation and metabolism in *Coffea arabica* (Arabica) and *Coffea canephora* (Robusta) revealed through gene expression and enzyme activity analysis. *New Phytol.* **2008**, *178*, 781–797. [[CrossRef](#)]
46. Perrois, C.; Strickler, S.R.; Mathieu, G.; Lepelley, M.; Bedon, L.; Michaux, S.; Husson, J.; Mueller, L.; Privat, I. Differential regulation of caffeine metabolism in *Coffea arabica* (Arabica) and *Coffea canephora* (Robusta). *Planta* **2015**, *241*, 179–191. [[CrossRef](#)]
47. Geromel, C.; Ferreira, L.P.; Guerreiro, S.M.C.; Cavalari, A.A.; Pot, D.; Pereira, L.F.P.; Leroy, T.; Vieira, L.G.E.; Mazzafera, P.; Marraccini, P. Biochemical and genomic analysis of sucrose metabolism during coffee (*Coffea arabica*) fruit development. *J. Exp. Bot.* **2006**, *57*, 3243–3258. [[CrossRef](#)]
48. Zheng, X.; Levine, D.; Shen, J.; Gogarten, S.M.; Laurie, C.; Weir, B.S. A high-performance computing toolset for relatedness and principal component analysis on SNP data. *Bioinformatics* **2012**, *28*, 3326–3328. [[CrossRef](#)] [[PubMed](#)]
49. Fiehn, O.; Wohlgemuth, G.; Scholz, M.; Kind, T.; Lee, D.Y.; Lu, Y.; Moon, S.; Nikolau, B. Quality control for plant metabolomics: Reporting MSI-compliant studies. *Plant. J.* **2008**, *53*, 691–704. [[CrossRef](#)] [[PubMed](#)]
50. Chong, J.; Yamamoto, M.; Xia, J.G. MetaboAnalystR 2.0: From Raw Spectra to Biological Insights. *Metabolites* **2019**, *9*, 57. [[CrossRef](#)] [[PubMed](#)]
51. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing: Vienna, Austria, 2014. Available online: <https://www.R-project.org/> (accessed on 6 October 2019).

