Unraveling varietal differences and nutraceutical potentials of an Indian grape cultivar Manjari Medika and its parents: a high resolution mass spectrometry-based untargeted metabolomics study

Sujata Chatterjee¹#, Nasiruddin Shaikh¹#, Niladri Sekhar Chatterjee²#, Amine Kassouf³,
Archana Dhole¹, Kaushik Banerjee¹*

¹National Referral Laboratory, ICAR-National Research Centre for Grapes, P.O. Manjri Farm, Pune 41207, India
²National Reference Laboratory, ICAR- Central Institute of Fisheries Technology, P.O. Matsyapuri, Cochin 682029, India
³Department of Chemistry and Biochemistry, Faculty of Sciences II, Lebanese University, Jdeideth El Matn, Fanar, Lebanon.

#These authors contributed equally

Corresponding author: Dr. Kaushik Banerjee
Postal mainlining address: ICAR-National Research Centre for Grapes, P.O. Manjari Farm, Solapur Road
Pune – 412307, Maharashtra, India
Phone: +91-20-26956000 (EPABX)
Fax: +91-20-26956099
Email: kbgrape@yahoo.com

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=4567945
Unraveling varietal differences and nutraceutical potentials of an Indian grape Manjari Medika and its parents: a high resolution mass spectrometry-based untargeted metabolomics study

Abstract

The paper provides a comprehensive metabolite profile of Manjari Medika (MM), an Indian teinturier grape juice variety, vis-à-vis its parents, Pusa Navrang (PN) and Flame Seedless (FS) with an untargeted metabolomics approach. By studying their secondary metabolite profiles, the comparative nutraceutical potentials of all of these three varieties were established. Aqueous methanolic extracts of cryoground whole berry samples were analyzed using liquid chromatography-high resolution mass spectrometry. The study identified 483 metabolites, out of which, 410 compounds were shared by all these varieties. The variety MM revealed 73 distinct compounds. The chemoinformatics analysis identified 21 metabolites as distinguishing biomarkers. In MM, the elevated relative abundance of the potentially bioactive metabolites, namely Epicatechin, Trigonellin, Peonidin-3-O-D-glucoside, Oenin, DL-arginine, L-glutamine and Resveratrol suggests its suitability as a potential functional food. The data generated can be used to develop novel grape varieties through strategic breeding programs.

Keywords: Vitis vinifera, Manjari Medika, teinturier grapes, plant metabolomics, anthocyanins, chemoinformatics, liquid chromatography high resolution mass spectrometry

1. Introduction

Grape (Vitis vinifera), an important fruit crop, is cultivated for consumption on the table, as well as for the making of juice, raisin, wine, and processed utility products. Several health-
promoting bioactive compounds, including flavonoids, anthocyanins, stilbenoids, glucosides, and free amino acids, have been identified in this fruit crop. The antioxidant, anti-cancer, and anti-viral properties of these bioactive components and their use in managing cardiovascular and neurological diseases in this matrix have also received scientific attention (Sabra et al. 2021; Troilo et al. 2021). Based on earlier studies of Lattanzio et al. (2006), which reported that colored grape varieties have higher polyphenolic content, and Di Ferdinando et al. (2014), which revealed that polyphenols aid in the plant’s adaptive and defensive abilities, it can be said that polyphenol-rich grapes are more disease-resistant. Worldwide, viticulturists focus on breeding new cultivars with superior qualities, including greater resistance to biotic and abiotic stress, extended shelf life, and functional food properties.

Following the trend, plant breeders at the Indian Council of Agricultural Research’s National Research Center for Grapes (ICAR-NRCG) are continually working toward developing improved grape varieties with high nutritional demands. Located in Pune in Maharashtra (MS), the major grape-growing state of India, the institute developed a new juice variety, Manjari Medika (MM), by crossing Pusa Navrang (PN) and Flame Seedless (FS) (https://nrcgrapes.icar.gov.in/technical%20bulletins/e-techbul-1-manjari-medika.pdf). They particularly chose PN and FS because both of them have acquired prominence in Indian markets for their distinctive phenotypical attributes. Importantly, the main advantage of PN is that it possesses high anthracnose resistance (Murria et al. 2018). Nonetheless, it is seeded and fragile and thus has a short shelf life. Although FS possesses large and firm berries, it has its own limitations: the pulp of this variety is colorless, indicating a lower polyphenol concentration (Tarabih et al. 2020). This might be why FS is sensitive to attacks from pests and physiological illnesses (Leonard and Kim, 2019). Earlier, Sharma et al. (2018) investigated MM’s overall quality and organoleptic properties and indicated its potential as a teinturier
cultivar with the highest total phenolics among the eight grape juice varieties of MS they studied.

Previously, we conducted a metabolomics study on MM grape seed and seed oil, the findings of which proved helpful in determining the varietal authenticity of MM grape seed and seed oil derived from two other world-renowned grape cultivars. Also, a favorable ratio of essential fatty acids (omega-6:omega-3) in MM seeds and seed oil indicated their potential health benefits (Khan et al. 2020a). Another in-house study demonstrated that MM seed oil and seed cake contained higher levels of tocopherols and tocotrienols, illustrating their potential utilization in dietary supplements (Khan et al. 2020b). Elsewhere, our collaborative study revealed that MM seed extract could be used as a therapeutic agent to alleviate methotrexate side effects and other hepatic disorders (Manna et al., 2023). We anticipated that the berry of this particular kind would deliver a greater number of health benefits based on the findings; nevertheless, nothing was known scientifically. A second concern is that the juice of MM and PN appears similar, making them prone to adulteration. Clearly, a comprehensive metabolite profiling of MM that considers varietal differences in relation to its parental lineages is worthwhile.

Metabolomics research is critical for understanding environment-to-gene interactions, phenotyping, and biomarker identification (Razzaq et al. 2019). There are two distinct ways to approach metabolomics: untargeted and targeted; the former can identify each sample metabolite without reference standards to list novel compounds, while the latter quantifies metabolites using reliable reference materials. Because of the advent of high-throughput analytical experiments, scientists have frequently had to deal with multivariate data in a matrix. Untargeted metabolomics has enabled the conversion of large datasets into manageable signals. Moreover, two supervised multivariate statistical analysis methods, namely partial least squares discriminant analysis (PLS-DA) and orthogonal partial least squares discriminant
analysis (OPLS-DA), are used in many food matrices, including grape (Millan et al. 2016, Stranska et al. 2021), and wine (Uttl et al. 2019). Whereas, principal component analysis (PCA) is an unsupervised multivariate technique that is exploratory in nature and provides insight into the sample clustering of classes without predefining classification conditions. Despite significant progress in the development of community-supported metabolite databases (e.g., GNPS, METLIN, and Chemspider), optimizing advanced mass spectrometry instrument-based conditions improves database readings (in-house experience).

The integration of metabolomics and other omics platforms has gained significant importance in breeding programs to accelerate crop improvement, both qualitatively and quantitatively (Kang et al. 2019; Rosato et al. 2018). As evident, varietal differentiation based on grape metabolite profiling has been gaining importance. For instance, a study differentiated Tannat, CS, and Merlot grape varieties of Uruguay based on their anthocyanin profiles by using high performance liquid chromatography (HPLC) coupled to UV-Vis detection (González-Neves et al. 2007). Elsewhere, Pajovic et al. (2014) from Montenegro used HPLC-diode array profiling of polyphenols to classify grape varieties. The excellent mass accuracy, full scan capabilities, high resolving power, and sensitivity of the high resolution mass spectrometers (HRMS) have putatively identified molecular features without the prior need for reference materials (Rochat, 2016). In the last couple of years, scientists have also reported liquid chromatography mass spectrometry (LC-MS) and HRMS-based methodologies for studying grape metabolite profiling. Previously, Flamini et al. (2015) evaluated polyphenols in grapes using a targeted suspect screening method using LC-HRMS and identified certain flavonols and resveratrol derivatives as unique features. Again, Millan et al. (2016) performed a comparative analysis with a targeted metabolomics approach and found significant differences in the metabolite profiles of the seed, skin, and pulp of six Rioja grape cultivars.
As mentioned earlier, there was no robust, reliable, and precise classification model predicting the varietal composition of the MM berry and its parents, namely FS and PN. Furthermore, specific biomarkers important in the differentiation of these studied varieties have yet to be identified. Given the gap in knowledge, the aim of the study was to investigate the metabolite profile of MM using LC-HRMS and compare it with that of its parental lines. We were inspired from our recent collaborative untargeted metabolomics and chemometric species-level authentication study using LC-HRMS to discriminate *Tinospora cordifolia* (a medicinal species) from other identical ones, namely *T. sinensis* and *T. crispa* (Sarkar *et al.* 2023). To the best of our knowledge, this is the first study of its kind to compare and distinguish MM from its parental lines based on metabolite profiling. The application of untargeted metabolomics to reveal MM’s potential as a nutraceutical agent might assist in the development of future breeding programs. The findings of the study will be of interest to viticulturists, food scientists, and food business operators alike.

2. Materials and methods

2.1. Chemicals

Water (LC-MS grade) and methanol [99.8%, (HPLC grade)] were obtained from Fisher Chemicals (NJ, US) and NICE Chemicals (P) (Kochi, India), respectively. J.T. Baker (NJ, USA) supplied formic acid (90% p.a.), and methanol (99.9%, LC-MS grade). The accurate mass calibration mixture was procured from Thermo Fisher Scientific Co. (Mumbai, India).

2.2. Sample preparation

Mature grape berries of MM, FS, and PN were harvested from the vineyards of ICAR-NRCG, located at an altitude of 559 m above mean sea level (latitude: 18.32 °N, longitude: 7.51°E). The composite samples of each variety, comprising 1 kg of whole berries, were cryogenically ground with dry ice to generate a fine powder. These samples were stored at -80
7°C in airtight polypropylene screw-capped containers until further analysis. Similarly, 12 samples were obtained for MM, while the number of composite samples for FS and PN was 6 each. This ground berry sample (100 mg) was weighed into polypropylene micro-centrifuge (Eppendorf) tubes (2 mL) and extracted with 1.5 mL of a mixture of acidified methanol/water (7:3, v/v) in an ultrasonic water bath (Spectralab UCB 100, Mumbai, India) at 40 kHz for 30 min at room temperature. This was followed by its centrifugation at 10000 ×g for 5 min. Post-centrifugation, the resulting supernatant was diluted with water (1:1) and vortexed for 30 s. This diluted extract was filtered through a 0.22 μm hydrophilic polytetrafluoroethylene (PTFE) filter, and transferred into a glass vial. The blank sample contained a mixture of water/methanol (1:1, v/v). Aliquots (100 μL) from each vial were combined to make the pooled sample to be used as a quality control (QC) sample.

2.3. Data acquisition

The prepared samples were injected into a Dionex UltiMate 3000 Ultra High Performance (UHP) LC that was coupled with a Q-Exactive Orbitrap MS and an electrospray ionization (ESI) source (ThermoFisher Scientific, Mumbai, India). An autosampler, a rapid separation column compartment, and a quaternary pump were the features of the UHPLC system. The chromatographic separation was carried out using a reverse phase (RP) C18 column (100 × 2.1 mm, 1.8 μm) (Waters Corp., Bangalore, India). The mobile phase consisted of the following solvents: A: a mixture of acidified (0.1%, v/v formic acid) water/methanol (90:10, v/v), and B: a mixture of acidified (0.1%, v/v formic acid) water/methanol (10:90, v/v). Thereafter, the samples were acquired in the positive ESI mode with a total run time of 25 min. The optimized gradient flow included 1.0 min of 95% (A), followed by a linear decrease from 95% to 5% (A) over 10 min, isocratic cleaning step for 6 min at 95% B, a linear increase of A to the initial conditions (95% A) over 1 min, and a column equilibration step at 95% A for 7 min. The
solvent flow rate was set at 0.4 mL/min, while the injection volume was 5 μL. Throughout, the column temperature was maintained at 40 °C.

The ESI source conditions were optimized as follows: spray voltage (3.5 kV); nitrogen (>95%) as sheath gas at a flow rate of 45 (a.u.), auxiliary gas with a flow rate of 8 (a.u.) and sweep gas with a flow rate of 1 (a.u.); s-lens RF level (50); auxiliary gas heater temperature (300 °C); and capillary temperature (320 °C). The data were acquired in full scan-ddMS2 (Top N) mode with no inclusion list. Firstly, the full scan MS was performed without fragmentation in positive polarity (ESI+) within the m/z range from 100 to 1200 m/z using a maximum injection time (IT) of 100 microscan (ms), automatic gain control (AGC) target of $1 \times 10^6$, and a resolution power (RP) of 70,000 full width at half maximum (FWHM) (at m/z 200). The data-dependent MS/MS parameters were set as follows: 1 ms, RP of 17,500 FWHM (at m/z 200); AGC target of $1 \times 10^5$; maximum IT of 50 ms. An isolation window of 4 m/z was used with a stepped collision energy of 10 electron volt (eV), 30 eV, and 70 eV. To process chromatograms, XCalibur™ version 4.1 software (Qual Browser from ThermoFisher Scientific) was used.

Initially, the pooled QC samples were acquired through a series of 10 injections to stabilize the chromatographic and mass spectrometric outputs. Random injections were given to the biological samples of each type, with one pooled injection administered for every six individual samples. At the end of the sample acquisition list, three additional injections from the pooled QC sample were given.

2.4. Data processing

For data processing, the raw files were imported into Compound Discoverer version 3.3 (CD 3.3) software (ThermoFisher Scientific). The untargeted metabolomics workflow was used to align retention times (Rt), find peaks, predict elemental make-up, and remove background noise (blank subtraction). Additionally, it mapped molecules to biological pathways using the
Metabolika™ database. To remove any noise level peaks, the following detection parameters were optimized: a <5 ppm mass error for both precursor and fragment ions; a minimum intensity of 100,000; an Rt tolerance of ±0.1 min; and the consideration of selected adducts, including [M+H]+, M−, and [M+H+MeOH]+. This optimization was necessary to ensure that the recognized signals were authentic and no significant features were overlooked. The filtered features were annotated using mzCloud (https://www.mzcloud.org/), Predicted Composition, ChemSpider (http://www.chemspider.com/), and the in-house curated database of Masslist.

**Chemometric analysis**

In the statistical analysis platform, SIMCA 14 (https://www.sartorius.com/en/products/process-analytical-technology/data-analytics-software/mvda-software/simca), the processed data from CD 3.3 were first row-centered and pareto-scaled. On datasets, PCA was performed to provide insight into how similar or distinct the examined samples were. Afterwards, OPLS-DA and PLSDA were used to classify the data based on varietal identification. For identifying the most prominent distinguishing molecular features, pairwise OPLS-DA models were generated. Thereafter, the corresponding S-plots were used to identify the distinguishing metabolites. Furthermore, the significance of these identified biomarker metabolites was confirmed using a \( p \)-value >0.05, a log fold change >2, and VIP values >1 for distinguishing metabolites.

### 3. Results and Discussion

#### 3.1. Phenotypic characters and chromatographic patterns of the studied varieties

All three of the studied cultivars exhibited a distinct set of phenotypic characteristics. The PN berry was medium-sized, seeded, blue-skinned, brittle in texture, and had a bright reddish-blue pulp (Fig. S1). In contrast, the seedless cultivar FS produced large, red-skinned, firm berries with colorless pulp. As intended, MM inherited some superior traits from both of
its parents; its clusters contained large and firm berries with a soft pulp and no seeds. Interestingly, PN and MM had bluish-red skin and red pulp, whereas FS had red skin and translucent pulp. In contrast to the loose clustering pattern of PN berries in grape bunches, FS berries were loosely compact, whereas their hybrid, MM, had a compact clustering habit.

The chromatographic patterns are studied to determine the differences in metabolite composition and abundance among the three varieties. The TICs (Fig. 1) of all three varieties displayed a distinct pattern, especially for the early eluting compounds (Rt, between 1 and 2.2 min) and the mid-eluting compounds (Rt, between 8 and 15 min). While the early eluting compounds included predominantly free amino acids, the later ones comprised anthocyanins and polyphenols, mostly derived from berry skin and pulp. The relative intensity of these compounds across the TICs indicates variations in the metabolite profiles of MM, FS, and PN.

3.2. Data acquisition, processing, and compound identification

The untargeted metabolomics experiments generated large and complex data sets in the data independent acquisition (DIA) mode. The full scan-ddMS2 data acquisition mode offered the combined advantages of DIA and data dependent tandem mass fragmentation of the top four high intensity ions in a scan cycle. This required high performance computing tools for peak alignment, feature detection (based on adduct integration), and metabolite identification based on accurate mass matching of parent and fragment ions. Along with the workflow of Compound Discoverer 3.3 software, our own experiment-specific customizations were also used. Using one of the pooled QC injections, the Rt alignment was performed in the middle of the sequence, where the features were aligned according to the m/z of the detected adducts and Rt. Alignment scores of the samples that ranged from 78.6 to 99.5% indicate excellent data reproducibility across the LC-HRMS experiment.

To avoid detection of artifacts or noise as metabolite features, appropriate peak detection parameters, such as mass error <5 ppm for precursor and fragment ions, minimum
intensity of 100,000 (arbitrary unit), and Rt tolerance of 0.5 min, along with preferred adducts, for example, \([M+H]^+, M^-, \text{ and } [M+H+MeOH]^+\), were set. The filtered features were annotated using mzCloud, Predicted composition, Chemspider, and Masslist (Arita Laboratory Flavonoid Database). After initial processing, the software generated more than 2000 metabolite features, including both known and unknown ones. Further to these, the features were filtered to include only those metabolites associated with MS2 fragmentation. By analyzing and comparing the mass error, the fragment ion match score, and the mzlogic score, the compounds were shortlisted. Furthermore, the relevance of those metabolites to grapes was evaluated based on literature and previous in-house data. A final data matrix of 480 putatively identified compounds (m/z feature) was exported as an MS-Excel sheet. In multivariate analyses, compound identification and differentiation were performed using the SIMCA 14 software.

3.3. Identification of discriminating biomarkers through chemometric evaluation

The chemometric analysis determined and correlated the biomarkers with specific varieties. The following models were developed: Model 1- PCA, Model 2- OPLS-DA (all three varieties), Model 3- OPLS-DA for pairwise comparison of MM and PN, and Model 4- OPLS-DA for comparison between MM and FS. The statistical parameters of these models are presented in Table I.

The PCA model (Model 1, Table I) was based on a total of 36 samples, out of which 12 were MM, 6 were PN, 6 were FS, and 12 were pooled QCs (quality control). The Pareto scaling of the model divided each variable by the square root of the standard deviation and yielded a score of 92.6% for \(R^2_X\) (cum) and 92.1% for \(Q^2\) (cum), indicating good repeatability of the model. It was observed that the tested varieties were clearly differentiated from one another, while the quality control sample clustered densely in the vicinity of the plot’s center, suggesting that the method was successful in its application.
The model produced five principal components that together explained the total variation. Among these varieties, the first two components, PC1 and PC2 (Fig. 2A), contributed the most to the variations. Particularly, PC1 was responsible for 75% of all variations. PN and MM were clearly separated from FS along PC1. Furthermore, PC1 discriminated against them based on their skin color: purple for PN and MM, and red for FS. The distribution of metabolites along PC1 influenced this property, as shown in the loadings plot (Fig. 2B), where anthocyanins, specifically Peonidin-3-glucoside (Peo-3-glu) and Petunidin-3-glu (Pet-3-glu), had a strong positive influence on PC1. This was demonstrated by their greater relative abundance in PN and MM when compared to FS. Moreover, two amino acids, DL-Arginine (DL-Arg) and L-Glutamine (L-Glu), had a significant positive effect on PC1. Proline (Pro), another amino acid, had a negative effect on PC1 because it was more abundant in FS than the other two tested varieties. Whereas, PC2 explained 89% of the remaining variations and distinguished MM from its parents. According to the loadings plot (Fig. 2B), DL-Arg and L-Glu had a strong positive influence on PC2, whereas Cyn-3-glu and Pet-3-glu had a negative influence. When compared to the parents, DL-Arg and L-Glu were more prevalent in the progeny, while Cyanidin-3-glucoside (Cyn-3-glu) and Pet-3-glu were less abundant.

Since there was a clear separation among these studied varieties in PCA, an OPLS-DA model to predict class membership was developed, with the help of two variables (X and Y), and a total of 24 samples, consisting of 12 MM, 6 PN, and 6 FS. The model (Model 2, Table I) was auto-fitted and Pareto-scaled, which resulted in a score of 97% for $R^2_X$ (cum), 99% for $R^2_Y$ (cum), and 97% for $Q^2$ (cum), along with an RMSE$_E$ of 0.036 and RMSE$_{CV}$ of 0.058. This suggested that Model 2 was reliable and could discriminate between all of the tested varieties with high accuracy. From the six components that were calculated, the first two orthogonal multivariate components, PC1 and PC2, explained the majority of the variations that occurred among all the varieties. The score plots for PC1 and PC2 are depicted in Fig. 2C. On the basis
of amino acid contents, PC1 explained 47% of the total variation and discriminated FS from MM and PN. The loadings plot (Fig. 2D) displayed that Pro had a strong positive influence, while DL-Arg and L-Glu had a negative influence on PC1. Compared to PN and MM, higher levels of Pro and lower relative abundances of DL-Arg and L-Glu in FS contributed to their separation along PC1. On the other hand, PC2 explained 99% of the remaining variations and discriminated FS and PN from MM. The loadings plot (Fig. 2D) reveals that the metabolites, for example, Oenin (Oen), Peo-3-glu, Pet-3-glu, and Cyn-3-glu, showed a strong positive influence on PC2 and a higher relative abundance in PN and FS than MM. Hence, depending on the anthocyanin profile, PC2 discriminated between the parents and the descendant.

When a pairwise comparison between MM and PN with FS was done through OPLS-DA Models 3 and 4, it was possible to identify discriminating compounds. As mentioned above, Model 3, Table I was developed for comparing PN with MM. With a total of 18 samples (12 MM and 6 PN), this model displayed a score of 94% for $R^2_X$ (cum), 99% for $R^2_Y$ (cum), and 98% for $Q^2$ (cum), along with an RMSE$^E$ of 0.024 and an RMSE$^CV$ of 0.05, indicating its robustness, accuracy, and strong discriminatory power. The score plot of Model 3 (Fig. 3A) demonstrates a distinct separation between PN and MM, implying that their metabolomic profiles are distinct. Among the four PCs that were computed by autofitting the model, only PC1 displayed the highest variation between the two target varieties. Furthermore, to identify the marker compounds, the covariance of the metabolites was plotted against their correlation factor, yielding a scatter plot (S-plot). The metabolites with the highest covariance and correlation were selected as the marker metabolites for each pair. Following this, the VIP scores of the metabolites were computed. The details of the identified marker metabolites are presented in Table II. The metabolites (Fig. 3B), such as Oen, Peo-3-glu, and DL-Glu, were identified as the marker metabolites. Oen and Peo-3-glu had a strong positive correlation. This implies the possibility that their relative abundance was noticeably higher in PN compared to
The newly bred variety had a higher level of its biosynthetic pathway, as evidenced by the higher level of DL-Glu. Because DL-Glu is involved in the biosynthesis process, it should have a shelf life (subsequently explained in the metabolic pathway section).

In a similar vein, Model 4 was computed, but this time for comparing MM with FS and identifying the discriminating metabolites, using X and Y components and 18 samples (comprising 12 MM and 6 FS). Four orthogonal components were computed and auto-fitted in this case. Furthermore, PC1 could explain the greatest variation between MM and FS. The score plot of Model 4 (Fig. 3C) reveals a clear separation between these two, suggesting variation in their metabolite profiles. This model displayed a score of 96% for $R^2_X$ (cum), 99% for $R^2_Y$ (cum), and 98% for $Q^2$ (cum). This suggests the model has satisfactory repeatability, accuracy, and discriminatory power predictability. Similarly, an S-plot (Fig. 3D) was generated to identify the discriminatory metabolites for this pair of cultivars. To differentiate MM and FS, metabolites–Oen, Pet-3-glu, DL-Arg, Pro, and Cyn-3-glu–were selected as the discriminatory compounds.

In Fig. 4A, the relative abundance of the identified discriminating metabolites in MM and its parental lines is shown, and the Venn diagram presents the identified metabolites shared by PN, FS, and MM. The findings revealed 410 shared metabolites, indicating a strong metabolic correlation between the parental line and its descendant; among them, 73 metabolites were distinct. Compounds, namely Orsellinic acid and 3’-hydroxyacetophenone, were unique to PN. Whereas, MM showed clear evidence of the presence of polyphenols, particularly taxifolin, piceatannol, and luteolin 3’,3-acetylg glucuronide. Sphinganine and sphinganine-2-amion-1-3-octadecanediol were two examples of compounds that could only be found in FS.

Metabolomics fingerprinting was developed, which revealed the role of FS and PN in MM. There was a correlation observed between the genotypes and phenotypes, as well as the metabolite profiles of the cultivars that were investigated. In a similar way, Uttl et al. (2019)
successfully carried out varietal authentication of wines using untargeted metabolomics coupled with chemometrics in three white wine varieties, namely Pinot Gris, Riesling, and Roter Traminer.

In this study, several bioactive anthocyanins and flavonoids, namely quercetin glucoside, delphinidin, cyanidin glucoside, kaempferol, quercetin, and genistin, were upregulated in MM when compared to its parents. The elevated relative abundance of bioactive secondary metabolites in MM suggests its suitability as a potential functional food. Also, Crupi et al. (2015) characterized the polyphenolic profile of seven grape cultivars and differentiated them using PCA to determine their suitability for the production of grape juice with higher nutraceutical properties but using HPLC-DAD-MS/MS.

It is a well-known scientific fact that every species is a product of the combination of their genetics and response to the environment. Since FS, PN, and MM were cultivated in the same environment, the variations in their metabolic levels could be due to genetic differences.

3.4. Significance of the identified compounds by studying metabolic pathways

Metabolites contribute to the phenotypes of the varieties and play a significant physiological role, which could be understood from pathway analysis. The KEGG encyclopedia contains the information about the biosynthesis of compounds of numerous organisms, including *Vitis vinifera* (Kanehisa et al. 2017). Returning to the study, the aromatic amino acids, for instance, Phenyl alanine, Tryptophan, and Tyrosine, the precursors to flavonoids and anthocyanins, were identified. The anthocyanin biosynthetic pathway (Fig. S2) shows the formation of marker compounds. Cyn-3-glu is produced from Cyn and Peo-3-glu from Cyn-3-glu, also depicted in Fig. S2. As observed, the pathway leading to the synthesis of Peo-3-glu is more active or elevated in MM and PN. The berries of FS have a reddish color because they have a higher quantity of Cyn-3-glu and a small amount of Peo-3-glu. In contrast,
the berries of MM and PN have a purple skin and a reddish-blue pulp because they contain a higher concentration of Peo-3-glu. When compared to MM and PN, the relative concentration of Del in FS was significantly higher. The relative concentrations of Del-3-O-6-pGlu, Oen (Mal-3-glu), and Pet-3-glu were higher in MM and PN than they were in FS. This suggests that the biosynthetic pathway that converts Del into Del-3-o-6-p Glu, Oenin (Mal-3-glu), and Pet-3-glu was more active in MM and PN than in FS.

The amino acids, including L-Arg and L-Glu, participate in the biosynthesis of L-Pro and polyamines, such as putrescine and spermidine, via a cascade of reactions (Fig. S3). Pro is known to accumulate as a response to drought stress in plants and could be a great indicator for drought tolerance while characterizing the phenotype of a cultivar (Stines et al. 2000). Polyamines are shown by various researchers to influence the post-harvest shelf life of grapes (Colin et al. 2002; Rahimi and Mirdehghan, 2015). Higher amounts of L-Arg might result in increased levels of polyamines, which in turn, might help predict the shelf life of grapes.

4. Conclusion

A teinturier variety MM was authenticated from its parents, PN and FS, grown in the same geographic source areas with an untargeted metabolomics and advanced LC-HRMS approach. The application of metabolomic fingerprinting revealed a strong correlation between MM and its parental lines. The mass spectrometric experiments identified the major metabolites contributing to differentiation as belonging to the following classes of compounds: Alkaloids, Amino acids, Anthocyanins, Flavan-3-ols, Flavanols, Isoflavones and Stilbenoids. Some superior qualities of each parent’s metabolites (PN and FS) have been expressed in MM, along with some distinct ones, such as DL-Arginine, Epicatechin, Trigonellin, and Tryptophan, contributing significantly to the variations. The increased relative abundance of these bioactive secondary metabolites in the hybrid suggests that it could be used as a functional food.
Metabolic fingerprinting reveals the metabolic contribution of each parent to the progeny, and metabolite accumulation can be tracked across generations. Additionally, untargeted metabolomics as an effective tool linked the phenotype to the genotype of MM and its parental lines, identified and mapped important metabolites (and their biosynthetic pathways), and laid the groundwork for selecting the essential traits in a breeding program. In conjunction with other omics technologies, the metabolomics study can serve as a model for developing novel grape varieties with a heightened bioactive profile and climate-resilient traits.

References


González-Neves, G., Franco, J. and Barreiro, L. et al. (2007). Varietal differentiation of Tannat, Cabernet-Sauvignon and Merlot grapes and wines according to their anthocyanin
composition. European Food Research and Technology, 225, 111–117.

https://doi.org/10.1007/s00217-006-0388-8


Hepatic Inflammation: Involvement of NF-κB/NLRP3 and Nrf2/HO-1 Signaling System.


http://dx.doi.org/10.1016/j.chroma.2016.05.081


Declaration of competing interest

The authors declare that they have no competing interests.

Funding

Funding was received from ICAR’s NASF project.

Abbreviations

PN- Pusa Navrang, MM- Manjari Medika, FS- Flame Seedless, Peo-3-glu- Peonidin-3-glucoside, Pet-3-glu- Petunidin-3-glucoside, DL-Arg- Arginine, L-Glu- Glutamine, Cyn-3-glu- Cyanidin-3-glucoside, Oen- Oenin (Malvidin-3-glucoside, Peo-3-glu- Peonidin-3-glucoside, Pro- Proline, Del-3-o-6-pGlu- Delphinidin-3-O-6-p coumaroyl glucoside, Del- Delphinidin, KEGG- Kyoto Encyclopedia of Genes and Genomes