RESEARCH ARTICLE

Proton Nuclear Magnetic Resonance-Spectroscopic Discrimination of Wines Reflects Genetic Homology of Several Different Grape (*V. vinifera* L.) Cultivars

Boran Hu[1☯]*, Yaqing Yue[1☯¤], Yong Zhu[2], Wen Wen[2], Fengmin Zhang[3], Jim W. Hardie[4]

1 College of Food Science and Engineering, Yangzhou University, Yangzhou, Jiangsu, China, 2 College of Tourism and Gastronomy, Yangzhou University, Yangzhou, Jiangsu, China, 3 Testing Center of Yangzhou University, Yangzhou, Jiangsu, China, 4 Northern Melbourne Institute of TAFE, Epping, Victoria, Australia

☯ These authors contributed equally to this work.
¤ Current address: College of Tourism and Gastronomy, Yangzhou University, Yangzhou, Jiangsu, China.
* huboran@yzu.edu.cn

Abstract

Background and Aims

Proton nuclear magnetic resonance spectroscopy coupled multivariate analysis (^{1}H NMR-PCA/PLS-DA) is an important tool for the discrimination of wine products. Although ^{1}H NMR has been shown to discriminate wines of different cultivars, a grape genetic component of the discrimination has been inferred only from discrimination of cultivars of undefined genetic homology and in the presence of many confounding environmental factors. We aimed to confirm the influence of grape genotypes in the absence of those factors.

Methods and Results

We applied ^{1}H NMR-PCA/PLS-DA and hierarchical cluster analysis (HCA) to wines from five, variously genetically-related grapevine (*V. vinifera*) cultivars; all grown similarly on the same site and vinified similarly. We also compared the semi-quantitative profiles of the discriminant metabolites of each cultivar with previously reported chemical analyses. The cultivars were clearly distinguishable and there was a general correlation between their grouping and their genetic homology as revealed by recent genomic studies. Between cultivars, the relative amounts of several of the cultivar-related discriminant metabolites conformed closely with reported chemical analyses.

Conclusions

Differences in grape-derived metabolites associated with genetic differences alone are a major source of ^{1}H NMR-based discrimination of wines and ^{1}H NMR has the capacity to discriminate between very closely related cultivars.

Significance of the Study

The study confirms that genetic variation among grape cultivars alone can account for the discrimination of wine by ^{1}H NMR-PCA/PLS and indicates that ^{1}H NMR spectra of wine of single grape cultivars may in future be used in tandem with hierarchical cluster analysis to elucidate genetic lineages and metabolomic relations of grapevine cultivars. In the absence of genetic information, for example, where predecessor varieties are no longer extant, this may be a particularly useful approach.

Introduction

Grape wine is a complex mixture of several hundred components present in various concentrations. It contains many metabolites from grapes but most chemicals in wine are products of alcoholic and malolactic fermentations. The main compounds are water, ethanol, glycerol, sugar and organic acids. Other compounds, such as phenols, amino acids and so-called secondary metabolites, are present in much lower concentrations. Many factors, including microbial ecology of the ferment, geographical origin, grape cultivar and winemaking technologies contribute to the composition and content of metabolites in wines and affect the quality of wines [1].

Due to the increasing interest in the qualities and provenance of grape cultivars and wines by consumers, regulatory authorities and producers, it is of commercial importance to develop fast and accurate methods to distinguish wines according to grape cultivar, geographic origin and other product-defining features.

NMR has been used increasingly in the characterization and quality control of wine and food [2,3]. Although the sensitivity of NMR is dependant on magnetic strength of the instrument and some NMR experiments do not yield more or better information than alternative analytical techniques, NMR still has many advantages over other methods. NMR allows rapid, simultaneous, measurement of a large number of organic compounds in complex mixtures and requires minimal pre-treatment of samples; thus reducing errors introduced by pre-treatment and preservation of the samples ([4–6]). Multivariate data analyses including principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA) are used to decrease the dimensionality of multivariate data and classify items according to inherent patterns in the data. Hierarchical cluster analysis (HCA) allows segregation of groups of items according to the degree of similarity between them.

Proton nuclear magnetic resonance (^{1}H NMR) spectroscopy coupled with PCA or PLS-DA has been used to discriminate between wines produced from the same cultivar grown in different geographic regions [7–9], two or more cultivars grown in the same geographic region [10,11]—including under different seasonal ('vintage') conditions [12,13], several cultivars of different *Vitis* species grown in different geographic regions [14] and binary blends of single cultivar wines [15]. Notably, Pereira et al. found that within a season, across three contrasting soil types within a vineyard, genotypic discrimination of cultivars by ^{1}H HMR spectra was dominated by variation in soil type and that, between seasons, variability induced by the climate was most dominant.

Recent genomic studies (Robinson et al. [16] and references cited therein) have revealed for the first time the genetic relations of some commercially important cultivars. This allowed us to determine the extent to which ^{1}H NMR spectra, in the absence of confounding environmental and vinicultural factors, reflect inherent genotypic differences between cultivars. He found

that three of the cultivars, Cabernet Sauvignon, Merlot and Ruby Cabernet, are closely related genetically by sharing Cabernet Franc either as a parent (Cabernet Sauvignon and Merlot) or a grandparent (Ruby Cabernet). The other two cultivars, Zinfandel (syn. Tribidrag, Crljenak Kastelanski, Primativo) and Syrah (syn. Shiraz), are more distantly related, both to each other and the 'Cabernet Franc' group [16]. But wines of Ruby Cabernet and Zinfandel have not been included in previously reported NMR studies. Accordingly we applied ^{1}H NMR with pattern recognition analyses to discriminate wines vinified similarly from five different grape cultivars of the same species, *Vitis vinifera* L., grown similarly on the same soil type in the same vineyard and harvested at similar ripeness. To corroborate our evidence of genetic discrimination, we compared the cultivar-related profiles in the levels of discriminant grape metabolites with reported chemical analyses of those cultivars.

We show the capacity of ^{1}H NMR, coupled with PCA, PLS-DA and HCA. To discriminate between wines on the basis of the genetic relatedness of the cultivars, we also report the discriminant compounds and their general relative, cultivar-based profiles.

Materials and Methods

We state clearly that no specific permissions were required for these activities, because the brewing company issued the permission. We confirm that the field studies did not involve endangered or protected species.

All wine samples are produced in Shacheng Manor Wine Co. Ltd., Hebei province, China. After fermentation, we got 3 parallel samples of each wine from the sampling mouth of a 30 ton fermenting tank, each replicate sample was funnelled, using a 750 mL funnel, into a brown glass bottle which was then sealed with a cork and transported to the laboratory, and then stored at -4°C.

Sample origin and chemical analysis

The wine samples were of Cabernet Sauvignon, Merlot, Ruby Cabernet, Syrah and Zinfandel, each of the 2010 vintage and grown, non-grafted, in a single vineyard of uniform soil type in the Shacheng region of Hebei Province, China. The grapes of each cultivar were harvested at similar concentrations of reducing sugar (Table 1). Four replicate wines of each cultivar were vinified with the same yeast (Lalvin CY 3079), without chemical adjustment other than addition of potassium metabisulfite (50 mg/L) and they were not matured in contact with wood. The physical and chemical features of the wines, analyzed according to the Chinese national standard (GB/T 15038–2006), are presented in Table 2.

NMR sample preparation

10mL of wine of each replicate was centrifuged at 4,000rpm for 10min. Supernatants (3 mL) were frozen at -70°C for 10h and then lyophilized for 48h. The lyophilized product was dissolved in 99.9% deuterium oxide (630 μL, D_2O) and 0.75% 4, 4-dimethy-l-4-silapentane-

Table 1. Grape cultivars and fruit composition at harvest.

Cultivar	Harvest Date	Reducing Sugar (g/L)	Titratable Acidity (g/L)	pH
Cabernet Sauvignon	Oct. 6	226.2	6.8	3.3
Merlot	Sept. 25	215.6	6.4	3.4
Ruby Cabernet	Oct. 3	211.5	6.2	3.4
Syrah	Oct. 4	225.1	6.7	3.3
Zinfandel	Sept. 21	213.4	6.3	3.4

doi:10.1371/journal.pone.0142840.t001

Table 2. Physical and chemical features of the wines *.

| | Cultivars | | | | |
	Cabernet Sauvignon	Ruby Cabernet	Zinfandel	Merlot	Syrah
Alcohol content %vol	12.8	12.1	11.8	12.4	12.8
Residual sugar(glucose) g/L	2.11	2.07	2.03	2.11	2.23
Total acid g/L	5.7	5.8	5.5	6	6.1
Volatile acid g/L	0.48	0.45	0.51	0.48	0.48
Dry extract g/L	24.8	25.8	23.8	24.9	25.8
pH	3.55	3.56	3.54	3.51	3.46
Total SO_2 mg/L	72	82	81	79	76
Free SO_2 mg/L	31	28	34	30	34
Methanol mg/L	221	191	201	214	199
Fe^{3+} mg/L	1.2	2.1	1.8	2.2	1.8
Cu^{2+} mg/L	0.05	0.05	0.08	0.06	0.07
K^+ mg/L	882	931	999	988	972
Ca^{2+} mg/L	85	86	97	95	85
Tartaric acid g/L	2.42	2.37	2.28	2.24	2.41
Citric acid g/l	0.33	0.28	0.27	0.35	0.2
Lactic acid g/L	2.09	2.98	2.64	2.78	2.14
Colour tone	11.7	11.8	10.8	12.7	12.9
Colour tint	0.78	0.82	0.82	0.81	0.78

*Methods of determination of physical and chemical features accorded with China National Standard GB/T 15038–2006

doi:10.1371/journal.pone.0142840.t002

1-sulphonic acid (70 μL, DSS) and centrifuged at 13,000rpm for 10min. The supernatant (500 μL) was placed in a 5mm NMR tube. DSS provided a chemical shift reference (δ = 0) and the internal standard for quantitative analysis. D_2O provided a field frequency lock signal.

To confirm the influence of pre - ^{1}H NMR sample lyophilization the following procedure was applied: 3mL of wine (Cabernet Sauvignon) was centrifuged at 12,000rpm for 20min. The supernatant (450 μL) was mixed with 50 μL of a solution of 0.75% 4, 4-dimethy-l-4-silapentane-1-sulfonic acid (DSS) in 99.9% deuterium oxide (D_2O) in a 5mm NMR tube. Four replicate samples of wine treated this way were subsequently spectrally compared with lyophilized samples after suppression of the water peaks.

NMR spectroscopy

^{1}H NMR spectra were recorded on a Bruker AVANCE 600 spectrometer, operating at 600.13MHz 1H frequency and a temperature of 298K, using a [1]H {13C/15N} probe. A NOESY-PRESAT pulse sequence was used to suppress the residual water signal. A total of 256 transients were collected into 32,000 complex data points with a spectral width of 7183.9Hz, with a relaxation delay of 2s, an acquisition time of 2.3s, and a mixing time of 100ms. The NMR spectra were processed with a line-broadening factor of 0.3Hz prior to Fourier transformation. [1]H-^{1}H correlation spectroscopy (COSY) data were acquired with a time domain data matrix of 2048 by 512 and a mixing time of 80ms.

NMR data reduction

All of the NMR spectra were phase and baseline corrected by AMIX and then the NMR spectral data were reduced into 0.005ppm spectral buckets. The regions corresponding to water

(4.6–4.8ppm), incompletely removed ethanol (1.18–1.22 ppm and 3.57–3.72 ppm) and DSS (-0.5–0.5 ppm, 1.74–1.84 ppm and 2.90–2.95 ppm) were removed. The datasets were then imported into SIMCA-P version 12.0 for multivariate statistical analysis.

Multivariate data analysis

Principal component analysis (PCA), an unsupervised pattern recognition method, i.e. without preliminary hypotheses, was used to examine the intrinsic variation in the dataset. To maximize the separation between samples partial least squares discriminant analysis (PLS-DA) was applied. PLS-DA can be regarded as a supervised derivative of PCA that provides the maximum covariance between measured data (X variable, metabolites in NMR spectra) and the response variable (Y variable, NMR spectral intensities), it can selectively extract the Y-relevant variation of variables, but also ensure the maximum correlation between them. The Hotelling's T^2 region, shown as an ellipse in score plots of the models, defines the 95% confidence interval of the modeled variation [17]. After orthogonal signal correction (r = 4) was applied to eliminate information that did not contribute to the discrimination, PLS-DA score plots from the ^{1}H NMR spectra of wines of the five cultivars were generated in pairwise comparisons. The quality of the models is described by the R^2X and Q^2 values. R^2X is defined as the proportion of variance in the data explained by the models and indicates goodness of fit. Q^2 is defined as the proportion of variance in the data predictable by the model and indicates predictability.

Hierarchical cluster analysis

^{1}H NMR data for four replicate wine samples of each cultivar was imported into IBM SPSS 19.0 for hierarchical cluster analysis to characterize the degree of metabolic similarity between the cultivars.

Assignment of metabolites

From the ^{1}H NMR and ^{1}H-^{1}H COSY spectra, e.g. Fig 1, and according to chemical shifts reported by Larsen et al. [18]; Lee et al. [19]; Pereira et al. [20]; and Son et al. [7], we identified and assigned the major resonances for 17 components in the wines (Table 3). Identification of corresponding loading plot signals indicated the PLS-DA discriminant metabolites.

Relative amounts of discriminant metabolites between cultivars

To provide an indication of the relative amounts of each of the discriminant metabolites between the cultivars and to allow some general hierarchical trends in those metabolites
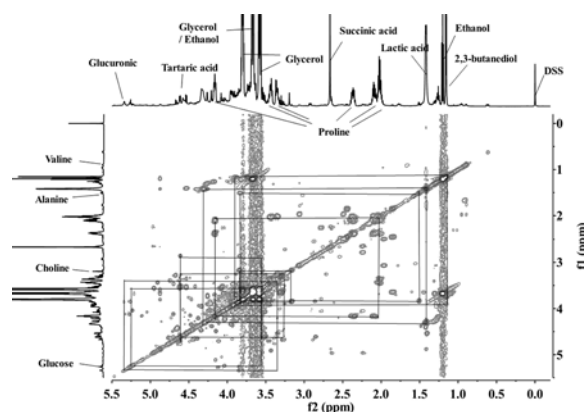

Fig 1. The ^{1}H-^{1}H COSY NMR spectrum of lyophilized Cabernet Sauvignon dry red wine.

according to cultivar, a matrix of cultivar-discriminating compounds was constructed using the relative peak heights of compounds represented in the pairwise comparison loading plots. Ranking within each hierarchy was established on a semi-quantitative basis with regard only to the comparative magnitude (i.e. greater or lesser) of the peak heights in each possible pairwise comparison.

Results

Comparison of wine samples with and without lyophilization

Fig 2 presents the comparison of NOESYPRESAT water peak—suppressed ^{1}H NMR spectra of Cabernet Sauvignon wine with and without lyophilization to remove water. In the spectra of samples without lyophilization the signals of ethanol and glycerol were so intense that they obscured the signal of other less abundant components. That made the identification and assignment of minor components difficult. In the spectra of lyophilized samples, even though a small amount of water and ethanol remained, the minor components were not obscured and were more readily identified.

Multivariate analysis

All of the NMR spectra were pre-treated by AMIX and then imported into SIMCA-P version 12.0 for PCA and PLS-DA. PCA of ^{1}H NMR spectra revealed a general clustering of the replicate wines by cultivar. The score plot of the first two principal components (PCs) is shown in Fig 3A. The first two PCs accounted for 65.9% of the total variance. On this analysis wines of

Table 3. Metabolites and their ^{1}H chemical shifts identified by 600 MHz ^{1}H NMR[*].

No.	Compound	Molecular formula	^{1}H NMR chemical shift
1	Ethanol	C_2H_5OH	1.19(t,C2H$_3$),3.66(q,C1H$_2$)
2	Methanol	CH_3OH	3.36(s,CH$_3$)
3	2,3—butanediol	$C_4H_{10}O_2$	1.15(d,C1H$_3$),3.88(q,C4H$_3$)
4	Glycerol	$C_3H_8O_3$	3.56(q,C2H$_2$),3.65(q,C3H$_2$),3.78(m,C1H)
5	Acetic acid	CH_3COOH	2.07(s,C3H$_3$)
6	Lactic acid	$C_3H_6O_3$	1.38(d,C3H$_3$),4.29(m,C2H)
7	Tartaric acid	$C_4H_6O_6$	4.53(s,C2H+C3H)
8	Succinic acid	$C_4H_6O_4$	2.67(s,C2H$_2$+C3H$_2$)
9	β-glucose	$C_6H_{12}O_6$	4.64(d,α-C1H,ring),3.24(dd,u,C2H,ring)3.54(dd,u,C3H,ring)
10	α-glucose	$C_6H_{12}O_6$	5.26(d,β-C1H,ring),3.55(dd,u,C2H,ring)
11	α-D-glucuronic	$C_6H_{10}O_7$	5.34(d,C1H,ring)
12	β-D-glucuronic	$C_6H_{10}O_7$	4.58(d,C1H,ring)
13	Valine	$C_5H_{11}NO_2$	0.89(d,C4H$_3$),0.95(d,C5H$_3$)
14	Alanine	$C_3H_7NO_2$	1.52(d, β-C3H$_3$),3.88(q, u, α-CH)
15	Proline	$C_5H_9NO_2$	2.02(m,u,γ-CH$_2$),2.07(m,u,β-CH),2.36(m,u,β'-CH),3.36(m,u,δ-CH),3.42(m,u,δ-CH),4.17(m,u,α-CH)
16	Choline	$C_5H_{15}NO_2$	3.19(s,N-CH$_3$),3.49(t,u, α-CH$_2$),3.83(t, β-CH$_2$)
17	Gallic acid	$C_7H_6O_5$	7.16(s,C2H+C6H)

[*]Letters in parentheses indicate the peak multiplicities

s, singlet

d, doublet

t, triplet

dd, doublet of doublet

tt, triplet of triplets

q, quartet; and

m, multiplet.

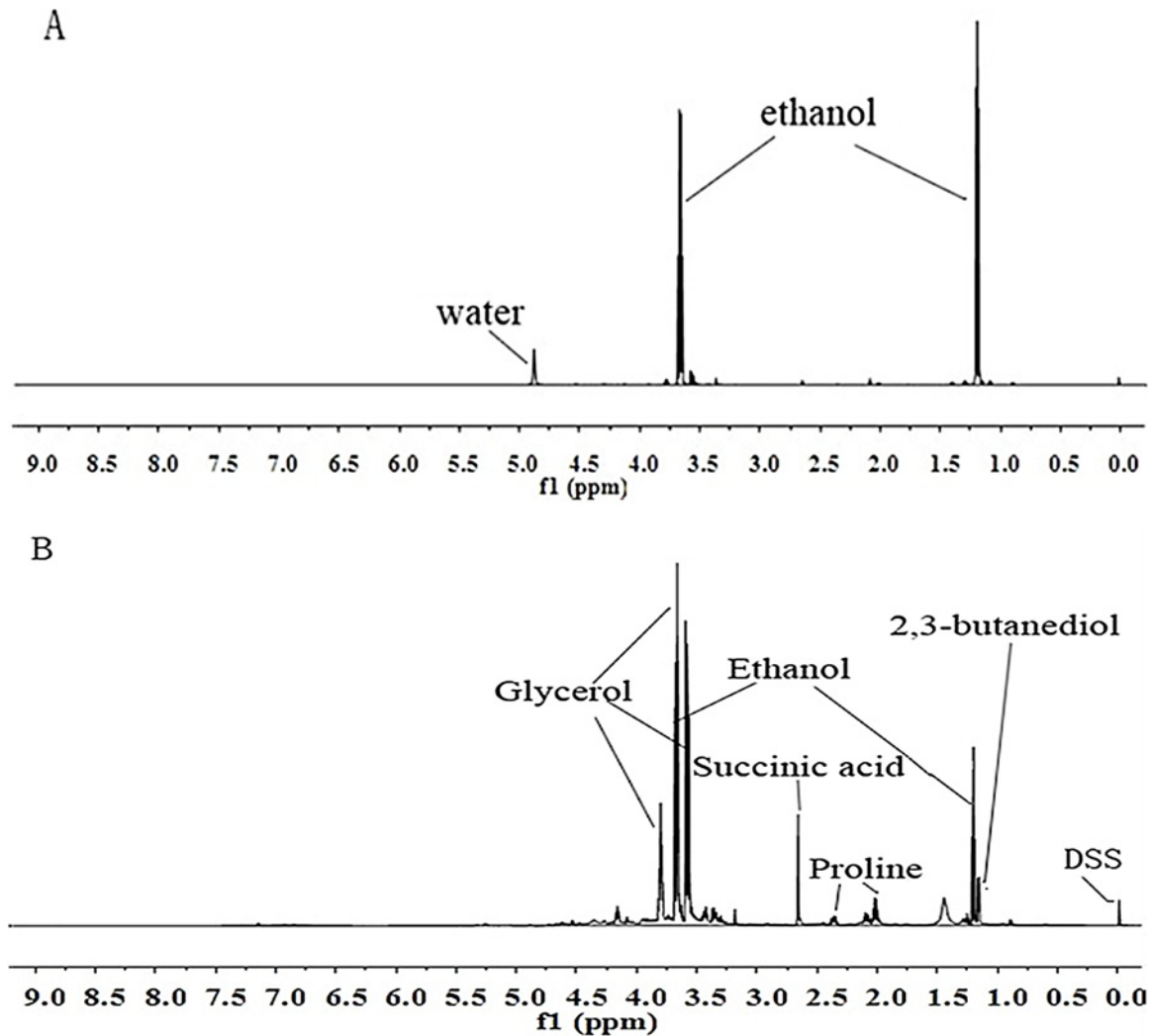doi:10.1371/journal.pone.0142840.t003

Fig 2. ¹H NMR spectrum of Cabernet Sauvignon wine with NOESYPRESAT water peak suppression, (A) sample not lyophilized, (B) sample lyophilized.

both Zinfandel and Syrah were characterized by positive score clusters on PC1 in contrast to the other cultivars that were generally characterized by negative scores on that PC. Wines of the other cultivars were less strongly discriminated but on PC2 the scores for Ruby Cabernet were positive while those for Merlot were mostly negative and those of Cabernet Sauvignon were both positive and negative.

PLS-DA of ¹H NMR spectra (Fig 3B) provided similar but slightly greater discrimination than PCA, PC1/PC2 mapping accounted for 65.5% of total variability. With the exception of Cabernet Sauvignon wines, the scores of each cultivar cluster were more exclusively positive or negative on PC2 than in the PCA analysis.

Hierarchical cluster analysis

Following aggregation of the replicates of each cultivar, the successive nodes of the dendrogram aggregated Merlot and Cabernet Sauvignon, the 'Cabernet Franc' group, and Syrah and Zinfandel respectively (Fig 4). There was clear separation between each of the groups.
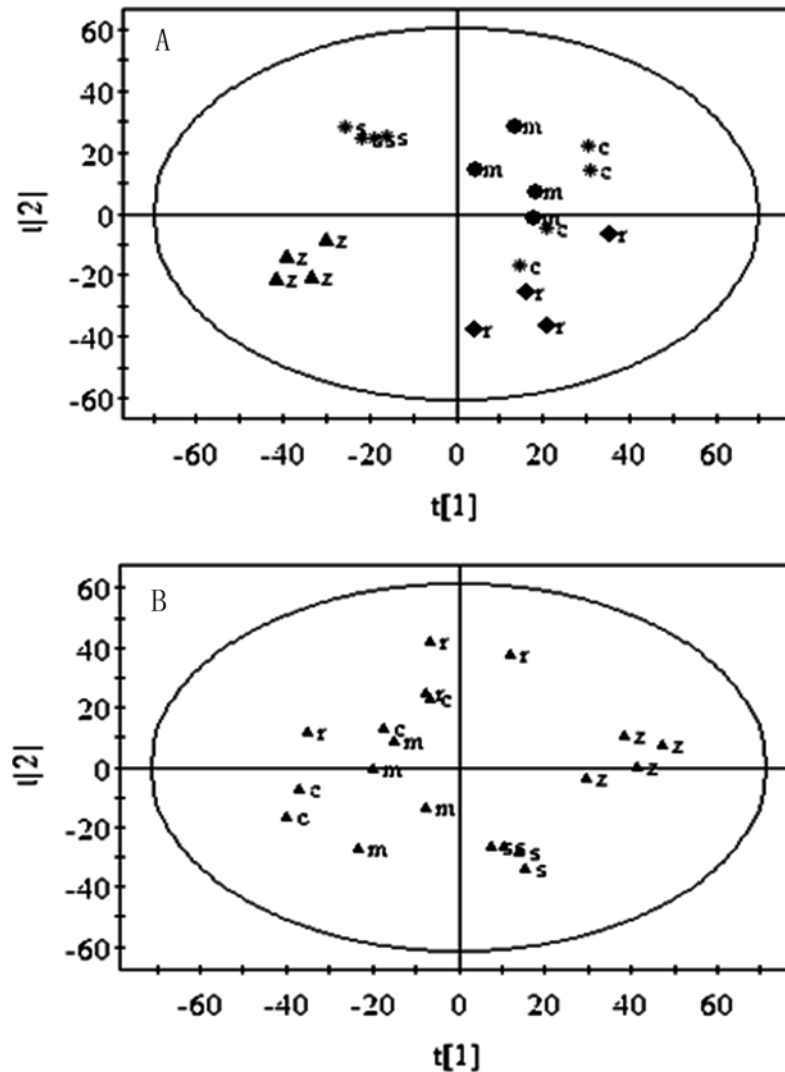
Fig 3. Comparison of wine discrimination derived from the ^{1}H NMR spectra. (A) PCA Scores Plot (t1/t2), R^2X = 0.959, Q^2 = 0.872. PC1/PC2 accounted for 65.9% of the total variance. (B) PLS-DA Scores Plot (t1/t2), R^2X = 0.959, R^2Y = 0.991, Q^2 = 0.954. PC1/PC2 accounted for 65.5% of the total variance. Grape cultivars: m, Merlot; s, Syrah; z, Zinfandel; r, Ruby Cabernet; c, Cabernet Sauvignon.

doi:10.1371/journal.pone.0142840.g003

Pairwise comparison of wines and discriminating metabolites

Pairwise PLS-DA score plots derived from the ^{1}H NMR spectra of wines of the five cultivars showed clear discrimination (i.e. high values of the regression coefficients R^2X, R^2Y and cumulative Q^2; the proportion of the variation of Y predicted by the model) between wines of each cultivar by the first component, and the loading plots showed the metabolites that contributed to the discrimination (Fig 5). In the loading plot each peak represents an integral region of the NMR spectra and the height of the peak corresponds to a contribution coefficient of the region to the separation shown in the score plot. The discriminating compounds were grape metabolites, viz. proline, valine, total phenols (mostly Gallic acid), tartaric acid and glucose; and vinification products, viz. glycerol, 2, 3—butanediol, succinic acid, lactic acid, acetic acid and valine.

Table 4 presents the matrix of cultivar-discriminating compounds derived from the pairwise comparison of peak heights from the loading plots. The matrix provides an indication of the
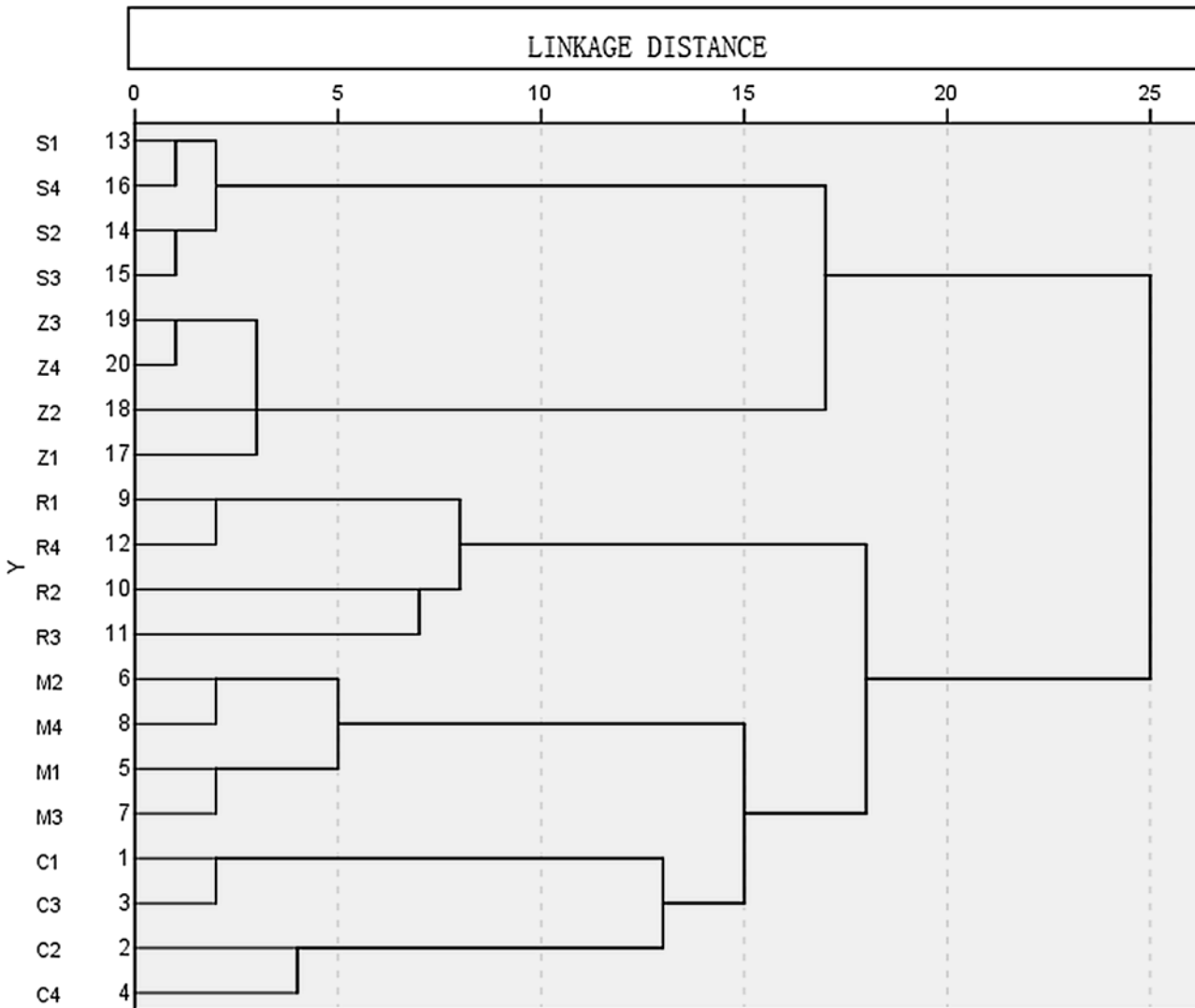
Fig 4. Dendrogram of wines of five cultivars (M, Merlot; C, Cabernet Sauvignon; R, Ruby Cabernet; Z, Zinfandel; S, Syrah), (4 replicates of each), based on multidimensional analysis of metabolites detected by ^{1}H NMR spectroscopy.

relative levels of many of those compounds and allows some general hierarchical classifications by cultivar. Some compounds were insufficiently represented within the matrix to allow hierarchical classification.

Discussion

The high resolution of ^{1}H NMR spectral data coupled with either PCA or PLS-DA allowed discrimination between wines of Syrah and Zinfandel and between each of those two cultivars and each of the representatives of the Cabernet Franc group, viz. Cabernet Sauvignon, Merlot and Ruby Cabernet. Among those three cultivars there was clear discrimination between Merlot and Ruby Cabernet, and Cabernet Sauvignon clustered between them. While several previous NMR studies of wine — including wine from Merlot, Cabernet Sauvignon and Cabernet Franc grown on three different soil types within a single vineyard in Saint Emilion, southwestern France [13] — have shown clear discrimination between the cultivars, their genetic relations
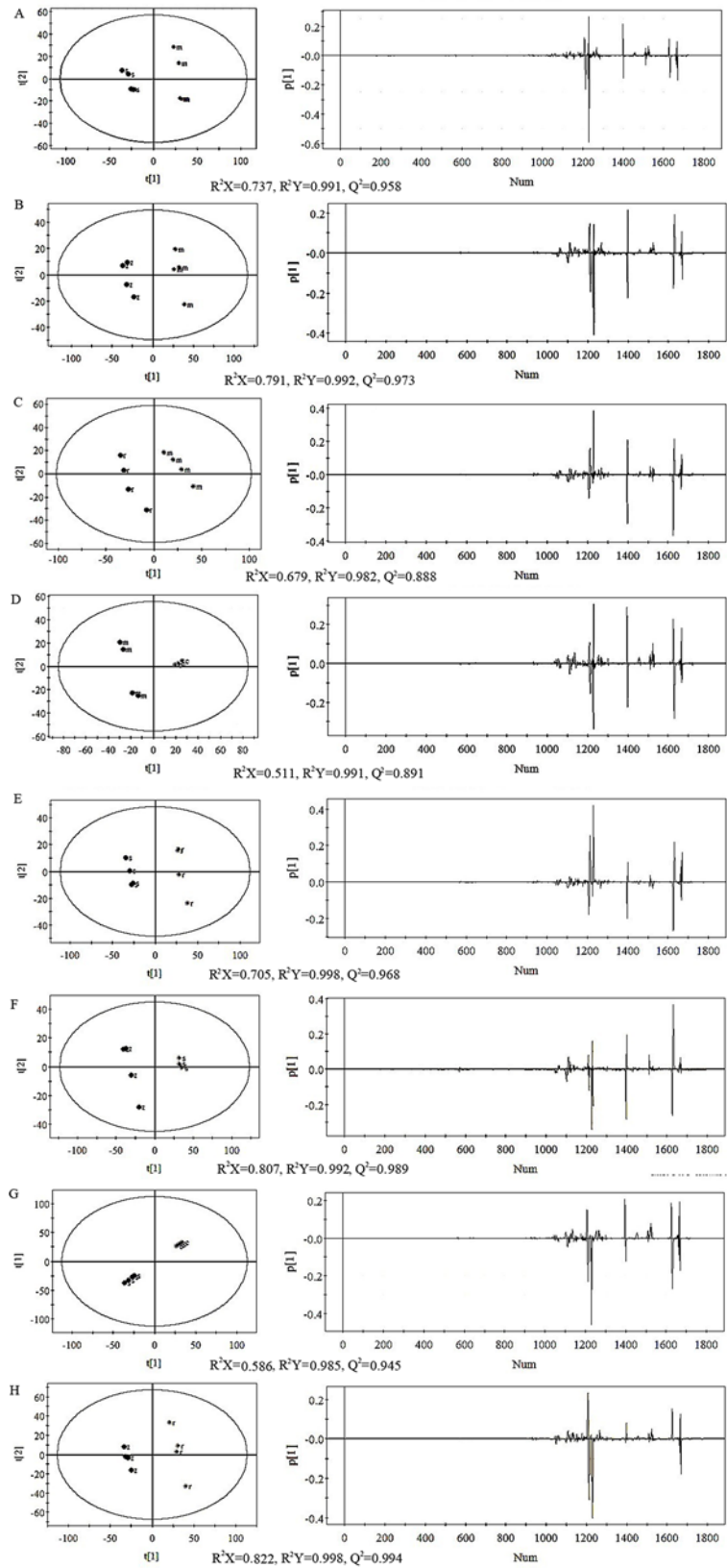
Fig 5. PLS-DA score plots, loading plots and correlation parameters derived from the ^{1}H NMR spectra of wines as pair wise comparisons (A) Merlot (m) and Syrah (s), $R^{2}X = 0.737$, $R^{2}Y = 0.991$, $Q^{2} = 0.958$. PC1/PC2 variance accounted for 72.1%. (B) Merlot (m) and Zinfandel (z), $R^{2}X = 0.791$, $R^{2}Y = 0.992$, $Q^{2} = 0.973$. PC1/PC2 accounted for 79.1%. (C) Merlot (m) and Ruby Cabernet (r), $R^{2}X = 0.679$, $R^{2}Y = 0.982$, $Q^{2} = 0.888$. PC1/PC2 variance accounted for 67.9%. (D) Merlot (m) and Cabernet Sauvignon (c), $R^{2}X = 0.511$, $R^{2}Y = 0.991$, $Q^{2} = 0.891$. PC1/PC2 variance accounted for 51.4% (E) Syrah (s) and Ruby Cabernet (r), $R^{2}X = 0.705$, $R^{2}Y = 0.998$, $Q^{2} = 0.968$. PC1/PC2 variance accounted for 70.5%. (F) Syrah (s) and Zinfandel (z), $R^{2}X = 0.807$, $R^{2}Y = 0.992$, $Q^{2} = 0.989$. PC1/PC2 accounted for; 80.7%. (G) Syrah (s) and Cabernet Sauvignon (c), $R^{2}X = 0.586$, $R^{2}Y = 0.985$, $Q^{2} = 0.945$. PC1/PC2 variance accounted for 56.7%. (H) Ruby Cabernet (r) and Zinfandel (z), $R^{2}X = 0.822$, $R^{2}Y = 0.998$, $Q^{2} = 0.994$. PC1/PC2 variance accounted for 80.2%. (I) Cabernet Sauvignon (c) and Zinfandel (z), $R^{2}X = 0.897$, $R^{2}Y = 0.994$, $Q^{2} = 0.983$. PC1/PC2 variance accounted for 89.7%. (J) Cabernet Sauvignon (c) and Ruby Cabernet (r), $R^{2}X = 0.775$, $R^{2}Y = 0.995$, $Q^{2} = 0.976$. PC1/PC2 variance accounted for 74.8%.

were not considered. Significantly, in the present study the degree of discrimination between all cultivars reflected a broad range of genetic homology. Ruby Cabernet, bred by H. P. Olmo in California, is well known as the progeny of Cabernet Sauvignon and the distantly related Carignan (syn. Cariñena) [21]. Recent genomic analyses have shown that Cabernet Sauvignon and Merlot have a common parent; Cabernet Franc,a cultivar grown in southwestern France from the seventeenth century [22,23]. The other two wines were from more distantly related cultivars, viz. Zinfandel, grown in Croatia as Tribidrag or Crljenak Kastelanski from the seventeenth century [16], and Syrah, a cultivar from southeastern France [24]. The discrimination of Ruby Cabernet samples from the other representatives of the 'Cabernet Franc' group almost certainly reflects metabolic differences arising from the intermixing of the genes of 'Cabernet Franc' group of south-western France and Carignan, a cultivar believed to have originated in Spain [16]. Significantly Pereira et al. (2007) found that for similarly vinified wines both seasonal climate and soil type influenced the degree of discrimination between cultivars [13]. Thus on the basis of differences in those factors we would expect differences evident in the degree of discrimination of Cabernet Sauvignon and Merlot in our study and theirs.

Hierarchical cluster analysis revealed groupings that not only reflect the genetic heritage of the cultivars but also represented the degree of similarity between them. The cultivars segregated into two groups, viz. the 'Cabernet Franc' group comprising Cabernet Sauvignon, Merlot and Ruby Cabernet, and the 'others'; Zinfandel and Syrah. Notably, within the 'Cabernet Franc' group Merlot and Cabernet Sauvignon which have the common parent Cabernet Franc, were more closely linked to each other than to Ruby Cabernet; despite Cabernet Sauvignon being a parent of Ruby Cabernet. As we have suggested, this is possibly the result of human-mediated breeding which introduced the more distantly related cultivar, Carignan, into the group. Within the 'other' group it appears that, in contrast to the 'Cabernet Franc group, Zinfandel and Syrah do not share close relatives.

Most significantly, the general correlation between degree of discrimination and genetic relatedness across a wide range of genetic homology represented by the cultivars in this study, including discrimination between very closely related cultivars, establishes that genetic variation among grape cultivars can account for the discrimination of wine by ^{1}H NMR. The grape metabolites that most represented that variation were proline, valine, total phenols (mostly Gallic acid), tartaric acid and glucose (see in S1 File.). Proline, which unlike other amino acids, usually (but not always) remains unconsumed by yeast during fermentation, was the predominant amino acid found in the NMR spectrogram. It is well known that among the red wine cultivars Cabernet Sauvignon and Merlot have particularly high levels of proline [25,26]. Ough (1968) showed that the general order of abundance (% of total juice N) of proline among the 'Cabernet Franc group' is Cabernet Sauvignon > Merlot = Cabernet Franc > Ruby Cabernet. Similarly in our study, the content of proline was highest in wines of the 'Cabernet Franc

Table 4. Matrix of cultivar-discriminating compounds[*].

	Cabernet Sauvignon	Merlot	Ruby Cabernet	Syrah	Zinfandel	
	Compounds that quantitatively exceed those of the cultivar in left hand column					
Cabernet Sauvignon		lactic acid, glucose,valine	lactic acid, tartaric acid, glucose, valin, phenols	2,3-butanediol,glycerol, lactic acid,tartaric acid, glucose	glycero,tartaric acid,valine,gallic acid	
Merlot	2,3-butanedio,glycero, tartaric acid,succinic acid, proline		lactic acid, succinic acid, phenols	ethanol,2,3-butanediol, glycerol	2,3-butanediol,glycerol,succinic acid,glucose	
Ruby Cabernet	2,3-butanediol,acetic acid, succinic acid,glucose, proline	2,3-butanediol, glycerol,acetic acid,tartaric acid,proline		2,3-butanediol,glycerol, tartaric acid	2,3-butanediol,glycero	
Syrah	acetic acid,succinic acid, proline	tartaric acid, glucose, proline,phenols	lactic acid, succinic acid, proline, phenols		glycerol,tartaric acid,succinic acid,glucose	
Zinfandel	2,3-butanediol,acetic acid, succinic acid,glucose, proline,phenols	acetic acid, tartaric acid, valine,proline, phenols	lactic acid, tartaric acid, succinic acid, glucose, valine, proline, phenols	2,3-butanediol,acetic acid		

[*]Compounds determined from loading variables from pair wise PLS-DA. Contents of compounds within cultivar columns exceed those of the cultivar shown in the column to the left.

doi:10.1371/journal.pone.0142840.t004

group' and the hierarchical order of abundance from highest to lowest was Cabernet Sauvignon > Merlot > Ruby Cabernet > Syrah = Zinfandel. This finding accords with that of Son et al. (2008) who found that proline was an important variable in discriminating Australian Cabernet Sauvignon from Australian Shiraz wines by ^{1}H NMR[7]. Although valine was a discriminant compound in our study, levels of this amino acid are determined by both grape and microbial metabolism [5] and there was insufficient evidence to identify it as a useful discriminating variable in terms of genetic relatedness of cultivars. Variation in glucose content among grape cultivars at similar total sugar content is well known but variation among yeast strains in the capacity to ferment glucose and fructose is also known [27]. Thus variation in the glucose content of wine in our study may not be attributable solely to genetic differences among the grape cultivars.

The levels of phenols (mono and polymeric) in wine are affected, not only by grape cultivar but also by region, cultural practices, fermentation processes, maturation with wood contact and wine age [1]. Polyphenolic compounds, in particular, are not easily assigned by NMR due to their complex s tructures [2]. Nevertheless total phenols (mostly Gallic acid) contributed to

the discrimination of wines by cultivar in this study. The order of abundance, viz. Ruby Cabernet > Cabernet Sauvignon > Merlot > Syrah > Zinfandel, generally accord with analytical reports which indicate relative levels of total phenols in wines of those cultivars as follows: Ruby Cabernet > Cabernet Sauvignon ≥ Merlot ≥ Cabernet Franc > Syrah ≥ Zinfandel; although the position of Syrah from those reports is not completely clear [28–34].

Tartaric acid is the major grape acid in wine. The concentration of tartaric acid in grapes depends on cultivar, ripening conditions and fruit maturity. Although tartaric acid is also commonly added prior to fermentation in order to adjust the pH and acidity, the wines examined in this study were not treated this way. In our study, tartaric acid was a discriminant compound however there was insufficient evidence to establish a hierarchy among the cultivars based on tartaric acid. It is well known that the content of tartaric acid is not constant during wine fermentation because of the formation of potassium bitartrate and calcium tartrate which tend to precipitate in wines. Generally, as Son et al. [14] noted, tartaric acid is unlikely to be a useful biomarker to characterize wines of hetereogenous vinification processes but as we have shown, without addition during vinification, it may be a useful component for discrimination wines by genetic origin.

Although compounds of grape origin obviously provide the foundation for discrimination between cultivars, the PCA and PLS-DA approach relies on discrimination on the basis of maximizing differences in chemical profiles in both qualitative and quantitative terms regardless of source. The discriminating compounds arising during alcoholic fermentation were 2, 3—butanediol, glycerol, valine and succinic acid (see in S2 File.).

In the wine of each cultivar the amount of ethanol, although quantitatively the most abundant metabolite, was similar and ethanol was not a major discriminant compound,in afct in my study,there is no enanol through procrsses. 2, 3—butanediol and glycerol are major constituents of wine. In this study the content of 2, 3—butanediol in Syrah wine was higher than in other wines and that compound appeared to be an important discriminant factor. The concentration of 2, 3—butanediol in the five wines in order from high to low was: Syrah > Cabernet Sauvignon > Zinfandel > Merlot > Ruby Cabernet. Glycerol was also an important discriminant compound. Its concentration in the five wines in order from high to low was: Zinfandel > Syrah > Cabernet Sauvignon > Merlot > Ruby Cabernet.

Most succinic acid in wine is a product of nitrogen metabolism by yeast but it is also formed during malolactic fermentation (MLF) in which lactic acid bacteria convert malate and citrate into lactate and other components [1]. Levels of succinic acid are thus related to the malic acid content of the grapes. The high level of lactic acid in each of the wines indicates MLF. The concentration of succinic acid is also generally related to the concentration of ethanol, glycerol and 2, 3—butanediol [9]. In a survey of the succinic acid content of Australian red wines made from Cabernet Sauvignon, Merlot, Ruby Cabernet and Shiraz between 1991 and 2003, Coulter et al. [35] found no influence of cultivar. In our study the range in concentration of succinic acid in wines was low and the order from high to low: Zinfandel > Ruby Cabernet > Cabernet Sauvignon > Merlot > Syrah. Succinic acid is very stable and its concentration changes little during aging [1] thus its close association with malic acid, a cultivar-related trait [36] indicates that it is likely to be a useful discriminant of wines according to cultivar. Interestingly Kliewer et al. reported that the level of malic acid in Zinfandel was about 70% greater than that in Merlot and Cabernet Sauvignon. That may explain the relatively high levels of succinic acid in Zinfandel wines in our study.

Having demonstrated the impact of plant genotype alone in discriminant analysis of ^{1}H NMR spectra at one geographic site, we note that due to the biological interaction of genotype with environment both the discriminant factors and the degrees of discrimination obtained using the ^{1}H NMR-PCA/PLS-DA approach are subject to specific environmental (edaphic,

climatic, biotic and phyto-cultural) conditions at other sites as previous cross-environment studies have shown. It is this biological interaction that makes ^{1}H NMR-PCA/PLS-DA a particularly powerful tool in the geographic discrimination of wine and other plant-based products.

Finally, our study indicates that ^{1}H NMR-derived spectra of wine of single grape cultivars grown under the same conditions may be used in tandem with hierarchical cluster analysis to elucidate presently unknown genetic lineages and metabolomic relations of grapevine cultivars. In the absence of genetic information, for example, where predecessor varieties are no longer extant, this may be a particularly useful approach. Further confirmation of this prospect will require validation with more fully genetically-characterized cultivars.

Conclusions

It may be concluded from this study that differences in grape-derived metabolites associated with genetic differences alone are a major source of 1H NMR-based discrimination of wines and that ^{1}H NMR has the capacity to discriminate between very closely related cultivars.

The study also indicates that ^{1}H NMR-derived spectra of wine of single grape cultivars grown under the same conditions coupled with hierarchical cluster analysis may be used to elucidate genetic lineages and metabolomic relations of grapevine cultivars.

Supporting Information

S1 File. Metabolites for PDO Lambrusco wine of Modena (Lambrusco Salamino di Santa Croce) DOI: 10.1021/jf302728b.
(PDF)

S2 File. Figures of compounds, PCA and PLS-DA to show discrimination between cultivars. DOI: 10.1016/j.foodres/2009.08.006.
(PDF)

Acknowledgments

Author Contributions

Conceived and designed the experiments: BH YY. Performed the experiments: YY YZ WW. Analyzed the data: BH YY. Contributed reagents/materials/analysis tools: FZ. Wrote the paper: BH YY JH.

References

1. Jackson R.S. (2000) Wine Science. Academic Press, California, USA. S

2. Hong Y.S. (2011) NMR-based metabolomics in wine science. Magnetic Resonance Chemistry 49, S13–S2.[2]

3. Košir I. J. and Kidrič J. (2002) Use of modern nuclear magnetic resonance spectroscopy in wine analysis: determination of minor compounds. Analytica Chimica Acta 458, 77–84.

4. Anastasiadi M., Zira A., Magiatis P., Haroutounian A., Skaltsounis A. L. and Mikros E. (2009) ^{1}H NMR-based metabonomics for the classification of Greek wines according to variety, region and vintage. Comparison with HPLC data. Journal of Agricultural and Food Chemistry 57 11067–11074. doi: 10.1021/jf902137e PMID: 19904930

5. Brescia M. A., Caldarola V., De Giglio A., Benedetti D., Fanizzi F. P. and Sacco A. (2002) Characterization of the geographical origin of Italian red wines based on traditional and nuclear magnetic resonance spectrometric determinations. Analytica Chimica Acta 458, 177–186.

6. Du Y., Bai G., Zhang X. and Liu M. (2007) Classification of wines based on combination of 1H NMR spectroscopy and principal component analysis. Chinese Journal of Chemistry 25, 930–936.

7. Son H.S., Kim K. M., Van Den Berg F., Hwang G.S., Park W.M., Lee C.H. et al. (2008) 1H nuclear magnetic resonance-based metabolomic characterization of wines by grape varieties and production areas. Journal of Agricultural and Food Chemistry 56, 8007–8016. doi: 10.1021/jf801424u PMID: 18707121

8. Son H.S., Hwang G.S., Kim K.M., Ahn H.J., Park W.M., Van Den Berg F., et al. (2009b). Metabolomic studies on geographical grapes and their wines using 1H NMR analysis coupled with multivariate statistics. Journal of Agricultural and Food Chemistry 57, 1481–1490.

9. Viggiani L. and Morelli M.A.C. (2008) Characterization of wines by nuclear magnetic resonance: A work study on wines from the Basilicata region in Italy. Journal of Agricultural and Food Chemistry 56, 8273–8279. doi: 10.1021/jf801513u PMID: 18693739

10. Papotti G., Bertelli D., Graziosi R., Silvestri M., Bertacchini L., Durante C. et al. (2012) Application of one and two-dimensional NMR spectroscopy for the characterization of protected designation of origin Lambrusco wines of Modena. Journal of Agricultural and Food Chemistry 61, 1741–1746. doi: 10.1021/jf302728b PMID: 22967082

11. Rochfort S., Ezernieks V., Bastian S.E.P. and Downey M. O. (2010) Sensory attributes of wine influenced by variety and berry shading discriminated by NMR metabolomics. Food Chemistry 121, 1296–1304.

12. Ali K., Maltese F., Toepfer R., Choi Y. H. and Verpoorte R. (2011) Metabolic characterization of Palatinate German white wines according to sensory attributes, varieties, and vintages using NMR spectroscopy and multivariate data analyses. Journal of Biomolecular NMR 49, 255–266. doi: 10.1007/s10858-011-9487-3 PMID: 21359513

13. Pereira G. E., Gaudillere J., Van Leeuwen C., Hilbert G., Marcourt M., Deborde C., et al. (2007) 1H NMR metabolic profiling of wines from three cultivars, three soil types and two contrasting vintages. Journal International des Sciences de la Vigne et du Vin 41, 103–109.

14. Son H.S, Hwang G.S., Ahn H.J., Park W.M., Lee C.H. and Hong Y. S. (2009a) Characterization of wines from grape varieties through multivariate statistical analysis of 1H NMR spectroscopic data. Food Research International 42, 1483–1491.

15. Imparato G., Di Paolo E., Braca A. and Lamanna R. (2011) Nuclear magnetic resonance profiling of wine blends. Journal of Agricultural and Food Chemistry 59, 4429–4434. doi: 10.1021/jf200587n PMID: 21425876

16. Robinson J., Harding J. and Vouillamoz J. (2012) Wine Grapes: A complete guide to 1,368 vine varieties, including their origins and flavours. Allen Lane,426–953.

17. Hotelling H. (1931) The generalization of student's ratio. The Annals of Mathematical Statistic, 2, 360–378.

18. Larsen F. H., Van Den Berg F. and Engelsen S. B. (2006) An exploratory chemometric study of 1H NMR spectra of table wines. Journal of Chemometrics 20, 198–208.

19. Lee J., Hong Y. and Lee C. (2009) Characterization of fermentative behaviors of lactic acid bacteria in grape wines through 1H NMR- and GC-based metabolic profiling. Journal of Agricultural and Food Chemistry 57, 4810–4817. doi: 10.1021/jf900502a PMID: 19441818

20. Pereira G. E., Gaudillere J., Van Leeuwen C., Hilbert G., Lavialle O., Marcourt M., et al. (2005) 1H NMR and chemometrics to characterize mature grape berries in four wine-growing areas in Bordeaux, France. Journal of Agricultural and Food Chemistry 53, 6382–6389. PMID: 16076122

21. Olmo H. P. (1948) Ruby Cabernet and Emerald Riesling. California Agriculture Experiment Station Bulletin 704, 1–9.

22. Boursiquot J-M., Lacombe T., Laucou V., Julliard S., Perrin F.-X., Lanier N., et al. (2009) Parentage of Merlot and related winegrape cultivars of south western France: discovery of the missing link. Australian Journal of Grape and Wine Research 15, 144–155.

23. Bowers J. E. and Meredith C. P. (1997) The parentage of a classic wine grape, Cabernet Sauvignon. Nature Genetics 16, 84–87. PMID: 9140400

24. Bowers J. E., Siret R., Meredith C.P., This P. and Boursiquot J-M. (2000) A single pair of parents proposed for a group of grapevine varieties in northeastern France. Acta Horticulturae 528, 129–132.

25. Lafon-Lafourcade S. and Peynaud E. (1959) Dosage microbiologique des acides aminés des moûts de raisin et des vins. Vitis 2, 45–56.

26. Ough C.S. (1968) Proline content of grapes and wines. Vitis, 7, 321–331.

27. Amerine M. A. and Thoukis G. (1958) The glucose-fructose ratio of California grapes. Vitis 1 224–229.

28. Arozarena I., Casp A., Marin R. and Navara M. (2000) Multivariate differentiation of Spanish red wines according to region and variety. Journal of the Science of Food and Agriculture 80, 1909–1917.

29. Bourzeix M., Heredia N. and Kovac V. (1983) Richesse de différents cépages en composés phénoli-ques totaux et en anthocyanes. Progrès Agricole et Viticole 100, 421–428.

30. Cliff M. A., King M. C. and Schlosser J. (2007) Anthocyanin, phenolic composition, colour measurement and sensory analysis of BC commercial red wines. Food Research International 40, 92–100.

31. Frankel E. N., Waterhouse A. L. and Teissedre P-L. (1995) Principal phenolic phytochemicals in selected California wines and their anti oxidant activity in inhibiting oxidation of human low density lipo-proteins. Journal of Agricultural and Food Chemistry 43, 890–894.

32. Jiang B. and Zhang Z. (2012) Comparison on phenolic compounds and antioxidant properties of Caber-net Sauvignon and Merlot wines from four wine grape growing regions in China. Molecules 17, 8804–8821. doi: 10.3390/molecules17088804 PMID: 22832882

33. Landrault N., Poucheret P., Ravel P., Gasc F., Cros G. and Teissedre P. (2001) Antioxidant capacities and phenolics levels of French wines from different varieties and vintages. Journal of Agricultural and Food Chemistry, 49, 3341–3348 PMID: 11453773

34. Woraratphoka J., Intarapichet K. and Indrapichate K. (2007) Phenolic compounds and antioxidative properties of selected wines from north east Thailand. Food Chemistry 104, 1485–1490.

35. Coulter A. D., Godden P. W. and Pretorius I. S. (2004) Succinic acid. Australian and New Zealand Wine Industry Journal 19, 16–25.

36. Kliewer W. M., Howarth L. and Omori M. (1967) Concentrations of tartaric and malic acids and their salts in Vitis vinifera grapes. American Journal of Enology and Viticulture 18, 42–54.