NMR spectroscopy and chemometrics-based analysis of grapevine

Ali, K.

CHAPTER 6

NMR spectroscopy and chemometrics as a tool for anti-TNFα activity screening in grapes

Kashif Ali[a], Federica Maltese[a], Ana Margarida Fortes[b], Maria Salomé Pais[b], Young Hae Choi[a], and Robert Verpoorte[a]

[a] Natural Products Laboratory, Institute of Biology, Leiden University, The Netherlands
[b] Plant Systems Biology Lab, ICAT, Center for Biodiversity, Functional and Integrative Genomics, FCUL, 1749-016 Lisboa, Portugal

Abstract

Grapes from 'Trincadeira', 'Touriga Nacional', and 'Aragonês', at four developmental stages were classified using NMR spectroscopy based metabolomics. Solid phase extraction was applied and the resulting water, methanol:water (1:1), and methanol fractions were tested for *in vitro* anti-TNFα production. The methanol:water fraction contained most of the phenolics and showed significantly higher activity than the other two fractions. The initial stages of grape development, green and veraison, were found more active against TNFα production as compared to the later ripe and harvest stages. Among the cultivars, 'Touriga Nacional' was found to be the most potent inhibitor. Different multivariate data analyses algorithms based on projections to latent structures (PLS) were applied to correlate the NMR and anti-TNFα production data. The VIP (variable importance in projections) plot showed that phenolics like quercetin, myricetin, (+)-catechin, (-)-epicatechin, caftarate, and coutarate, were positively correlated with high activity. This work demonstrates the great potential of NMR spectroscopy in combination with chemometrics for identifying compounds with anti-TNFα activity in complex mixtures of compounds such as plant extracts.

Introduction

Inflammation is a complex process and various mediators, like interleukins (IL) and tumor necrosis factor α (TNFα), are involved in the development of inflammatory diseases (Choy and Panayi 2001). TNFα, an inflammatory mediator, is one of the most important pro-inflammatory cytokines. It was discovered in 1975 as having an anti-tumor activity, but is now recognized as a host defense factor in immunological and inflammatory responses (Tracey et al. 1994). TNFα is known to be secreted during early stages of acute and chronic inflammatory diseases such as rheumatoid arthritis, asthma, septic shock and other allergic diseases (Herath et al. 2003; Cho et al. 2001). Low production of TNFα is advantageous for the host but overproduction of TNFα during infection plays a pivotal role in the development of diseases like disseminated intravascular coagulation, death in septic shock, cerebral malaria, along with wide range of other inflammatory diseases including asthma, dermatitis, multiple sclerosis, inflammatory bowl disease, cystic fibrosis, rheumatoid arthritis, and immunological

disorders (Björnsdottir and Cypcar 1999; Murphy et al. 1998; Medana et al. 1997). Therefore it is evident that the suppression of TNFα or anti-TNFα therapy could be beneficial for the treatment of these acute and chronic diseases.

The use of grapes for multiple purposes like juice, fresh and dried fruit, and most importantly in wine production, make them one of the most economically important and widely cultivated fruit crops across the world. In addition to their economic importance, an increasing number of medicinal advantages have been attributed to grapes. Grapes phytochemistry is known to have relatively high concentrations of phenolics which in turn resulted in many health effecting properties, for instance, cardioprotective, anti-oxidant, anti-inflammatory, and anti-cancer activities (Ali et al. 2010). Studies using human (Zern et al. 2005), and animal (Fuhrman et al. 2005; Seymour et al. 2008) models have shown that due to abundance of polyphenols, dried grapes possessed anti-oxidative and anti-inflammatory properties.

Chemical phenotyping has become the focal point in recent years, as the analysis of the low molecular weight compounds reflect the physiological activities of an organism or tissue under certain conditions. The observable chemical profile or fingerprint, referred as 'metabolome', is highly complex consisting of a variety of compounds of very different nature. Considering the great chemical diversity it is unlikely that a single analytical method could provide information about all the metabolites, and at the same time be unbiased, rapid, reproducible, and stable over time, while requiring only simple sample preparation.

An accurate snap shot of the metabolome is highly important in metabolomics which requires a reliable metabolite extraction (Colquhoun 2007; Becknort et al. 2007). NMR has a unique place not only in structure elucidation, and characterization of molecules but is now also considered as a major tool in metabolomics studies. Many reports have been published on correlating the NMR and bioactivity data using various multivariate data analysis methods (Roos et al. 2004; Bailey et al. 2004; Cho et al. 2009; Cardoso-Taketa el. 2008).

The present study describes the *in vitro* anti-TNFα assay to measure the inhibitory activity in three Portuguese grape varieties at different development stages. Two different vintages of 'Trincadeira' cultivar are also analyzed and compared. Several primary and secondary metabolites using 1D and 2D NMR techniques are identified.

The correlation of activity data and NMR data using different multivariate data analyses methods in order to identify the active ingredients in grapes is also presented.

Materials and Methods

Grape cultivars and sampling

The grape cultivars and their sampling have been already explained in Chapter 4 and 5.

Solid phase extraction (SPE)

A sample of 100 mg of lyophilized grape berries was extracted with 2 ml of the mixture of water and methanol (2:8), with ultrasonication for 20 minutes at 25 ^{o}C. The suspension was then centrifuged at 3500 rpm and the supernatant was transferred to a round bottomed flask. The same procedure was repeated two more times and the supernatants were pooled together in the flask and taken to dryness with a rotary evaporator. This grape berries extract was subjected to solid phase extraction (SPE) on SPE-C18 cartridges (Waters, Milford, MA, USA). Prior to its use, the SPE cartridge was prepared by elution of 10 mL of methanol followed by 10 mL of water. Then, the redissolved grape berry extract (1 mL of deionized water) was applied to the cartridge and eluted successively with 5 mL of water and then 5 mL of methanol:water (1:1) and finally with 5 mL of methanol. All three fractions were collected in round bottomed flasks and evaporated under vacuum and were used for further NMR analysis. All the solvents were purchased from Biosolve B.V. (Valkenswaard, the Netherlands).

Growth of cells and lipopolysaccharides stimulation

Human monocyte-like histiocytic lymphoma cells U937 obtained from the ATCC (CRL-1593.2) were grown in RPMI-1640 medium, supplemented with 10% (v/v) fetal calf serum and 2 mM L-glutamine (Life technologies, Breda, The Netherlands) at 37 °C, 5% CO_2 in a humidified atmosphere. U937 monocytic cells ($5x10^5$ cells per well) were plated in 96-well culture plate and then differentiated into macrophages using phorbol 12-myristate 13-acetate (PMA, 10 ng mL^{-1}, overnight, Omnilabo, Breda, The Netherlands). The PMA-differentiated macrophages were allowed to recover from PMA treatment for 48 h, during which the culture medium was replaced daily.

Lipopolysaccharides stimulation of cells was performed as described by (Sajjadi et al. 1996).

Cells treatment with plant extracts

Immediately after stimulation cells were treated with plant extracts at the concentration of 100µg mL^{-1} and then incubated at 37ºC for 4 hours. Supernatant were then collected and measured for TNFα content using the Human TNFα enzyme linked immunosorbent assay (ELISA) kit (R&D systems, Europe Ltd).

Enzyme-linked immunosorbent assay for TNFα

TNFα in culture supernatants were determined by quantitative ‘‘sandwich’’ enzyme-linked immunosorbent assay using paired antibodies purchased from (Biosource International, Inc., Camarillo, CA, USA). In brief, all wells of high-binding Immulon-plates (96 well NUNC MaxiSorp microplates) were coated with 100µL of the capture antibody (anti-Human TNFα) (0.250 mg 0.125 mL^{-1}). After overnight incubation at 4ºC, plates were washed with washing buffer and blocked for 1 hour with 1% bovine serum albumin in phosphate-buffered saline. Plates were aspirated and inverted on absorbent paper to remove excess liquid. Samples and standards were diluted with assay buffer. 100µL of diluted standards (recombinant Human TNFα protein) were filled in sixteen wells of first two columns of plates. Rests of the wells were filled with 100µl of samples in 100 µg mL^{-1} concentration. Only DMSO was contained in control samples. Immediately after 50 µL of working detection antibody (0.025 mg 0.125 mL^{-1}) was plated in every well and then plates were incubated for 2 hours at room temperature with continuous shaking at (700 rpm). The wells were washed again 5 times with washing buffer before addition of 100 µL of streptavidin-HRP to the wells and incubated at room temperature further for 30 minutes with continuous shaking at (700rpm). Again wells were aspirated and washed 5 times before addition of 100µL of TMB substrate. Plates were incubated for 30 minutes at room temperature with continuous shaking at (700 rpm). After 30 minutes the reactions were terminated by addition of 100 µL of 2M H_2SO_4, and absorbance was determined using a microtiter plate reader (Bio-Tek Instruments Inc., Winooski, VT, USA) at 450 nm. The concentration of TNFα in the samples was calculated by comparison of the absorbance

of the samples to the standard curve. The ratio (%) of TNFα inhibition release was calculated by the equation, i.e. Inhibition (%) = 100 x (1 – T/C), where 'T' represents the concentration of TNFα with grape extract while 'C' was the concentration of TNFα with only DMSO.

Cell viability assay

Cell viability after treatment with different plant extracts was determined by using MTT assay. Briefly, U937 cells having concentration of ($5x10^5$cells mL^{-1}) were placed in a 96 wells plate. The culture media also contains different plant extracts (100 µg mL^{-1}) in the presence or absence of 200 µg mL^{-1} LPS at 37°C. After 2.5 hours of incubation at 37°C, the medium was discarded and the formazan blue, which formed by reacting MTT with mitochondrial dehydrogenase in the living cells, was dissolved with 100 µL DMSO. The optical density (OD) was measured at 540 nm. The background signal inherent to the plates when no cell was present was subtracted from the absorbance obtained from each sample.

^{1}H NMR Spectroscopy

The three fractions eluted from SPE were redissolved in 1 mL of methanol-d_4. An aliquot of 800 µL of sample was transferred to the 5-mm NMR tube and used for the NMR analysis as explained in Chapter 4.

Data analysis and statistics

The ^{1}H NMR spectra (from all SPE fractions) were automatically reduced to ASCII files. Spectral intensities were scaled to methanol signal (δ 3.30) and reduced to integrated regions of equal width (0.04) corresponding to the region of δ 0.0–10.0. The regions of δ 4.85–4.95 and δ 3.2–3.4 were excluded from the analysis because of the residual signal of D_2O and CD_3OD, respectively. Bucketing was performed by AMIX software (Bruker) with scaling on total intensity. Principal component analysis (PCA) with scaling based on Pareto while projections to latent structures (PLS), PLS-discriminant analysis (PLS-DA), bidirectional orthogonal PLS (O2PLS), and O2PLS-discriminant analysis (O2PLS-DA), with scaling based on Unit Variance were performed with the SIMCA-P+ software (v. 12.0, Umetrics, Umeå, Sweden). The TNFα

content was arbitrarily set as 100 in the positive control and all the other values are normalized to this (% activity) and shown in results. Means and standard deviations were calculated and means comparisons were made with ANOVA at a significance level <0.01.

Results

^{1}H NMR spectra visualization

The ^{1}H NMR spectra of three SPE fractions are shown in Figure 1A. It is evident from the figure that all the three SPE fractions are quite different from each other in terms of contained metabolites. The water fraction shows mostly sugars and organic acids while the methanol fraction shows mostly amino acids and fatty acids with some resonance in phenolic region Figure 1B. The methanol:water fraction shows the presence of maximum amount of phenolics with relatively few sugars and amino acids. The phenolic regions of ^{1}H NMR spectra from all three cultivars are shown in Figure 1C. Among the cultivars, 'Touriga Nacional' is found to have highest phenolic content. It can also be observed that, in general, each developmental stage has a unique metabolic profile. As shown by NMR, the initial stage in berry growth is characterized by high levels of phenolics with fewer sugars. As the berry grows, the level of sugars seems to increase with a decrease in phenolics content. The distribution of metabolites according to grape cultivars and developmental stages is explained in detail in the following sections.

Metabolite identification

Using our in-house library of NMR data of common metabolites, we identified some flavonoids including both flavonols and flavan-3-ols. Flavonols like quercetin, and myricetin were identified in the aromatic region along with (+)-catechin and (-)-epicatechin of the flavan-3-ols group. The aromatic part of the ^{1}H NMR spectra shows some signals of hydroxybenzoates like gallic acid, syringic acid, and vanillic acid. Tartaric esters of hydroxycinnamic acid were also identified which include caftaric acid (caffeic acid conjugated with tartaric acid), fertaric acid (ferulic acid conjugated with tartaric acid), and *p*-coutaric acid (coumaric acid conjugated with tartaric acid). Along

with the *trans*- forms, the *cis*- forms of these conjugated cinnamic acids, i.e. *cis*-caftaric acid and *cis*-coutaric acid, were also detected.

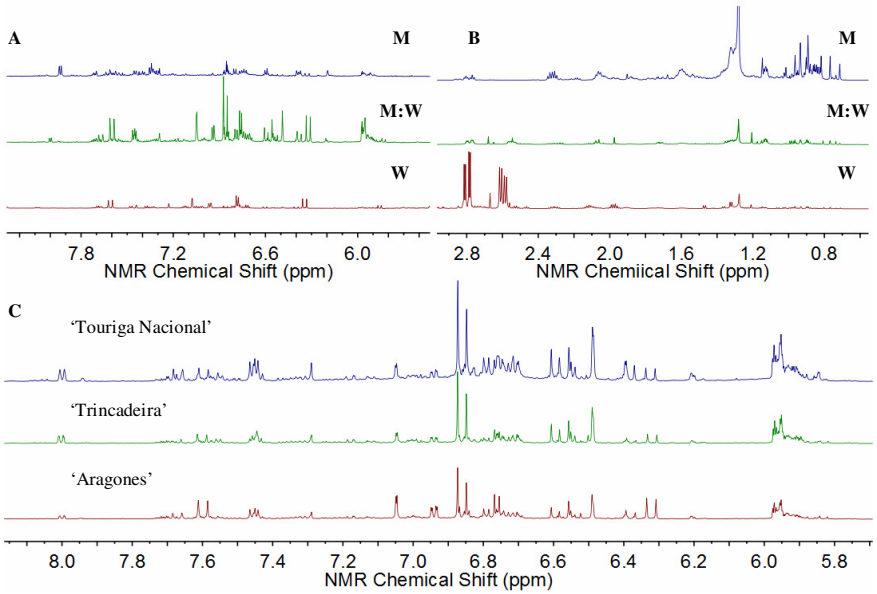

Figure 1. Comparison of ^{1}H NMR spectra of phenolic (A) and aliphatic (B) regions of SPE fractions, and phenolic region of three cultivars at green stage (C).

A number of amino and organic acids were identified due to the high signal intensities in the amino acid region. Amino acids like alanine, leucine, threonine, valine, proline, methionine, and glutamate, were identified by comparison with the reference spectra of these compounds. The signals in the carbohydrate regions were highly clustered and overlapped. This region showed the signals of the anomeric protons of β-glucose, α-glucose, fructose, and sucrose. Resonances for some other compounds like GABA, choline, and 2,3-butanediol were also identified in the same region. A number of signals have been elucidated as organic acids like α-linolenic acid, acetic acid, succinic acid, fumaric acid, formic acid, citric acid, malic acid, and tartaric acid. All these assignment of metabolites are based on our in house data base NMR data of standards measured under identical conditions.

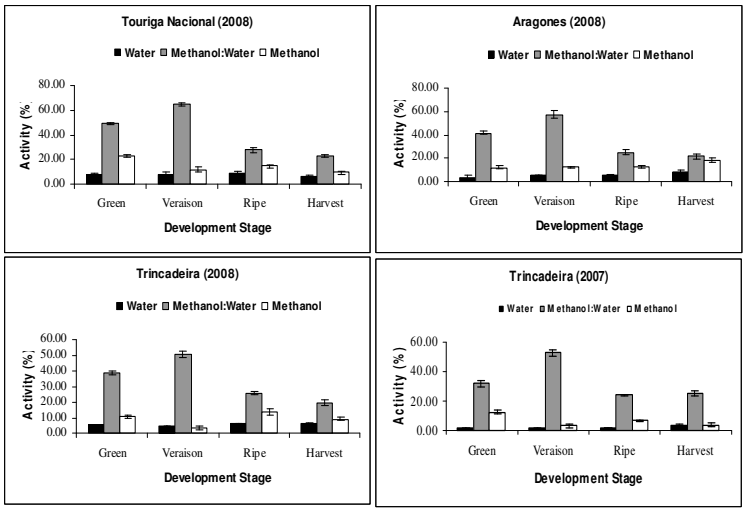

Figure 2. TNFα inhibition ratio (%) exhibit by grape cultivars and vintages at different developmental stages. Bars represent the Mean ± S.D. (n=3) at p<0.01.

Grapes and anti-TNFα activity

The potential of three grape cultivars at four developmental stages to inhibit the production of TNFα is evaluated. All three fractions from SPE of grape extracts were tested for anti-TNFα activity at 100 μg mL^{-1}. The methanol:water fractions show significantly higher activity than the water and methanol fractions. It has been shown in the previous section that the metabolic composition of these fractions are quite different from each other and the methanol:water fraction contained most of the grape phenolics. The water and methanol, fractions also showed some activity, though mostly not significantly different from each other.

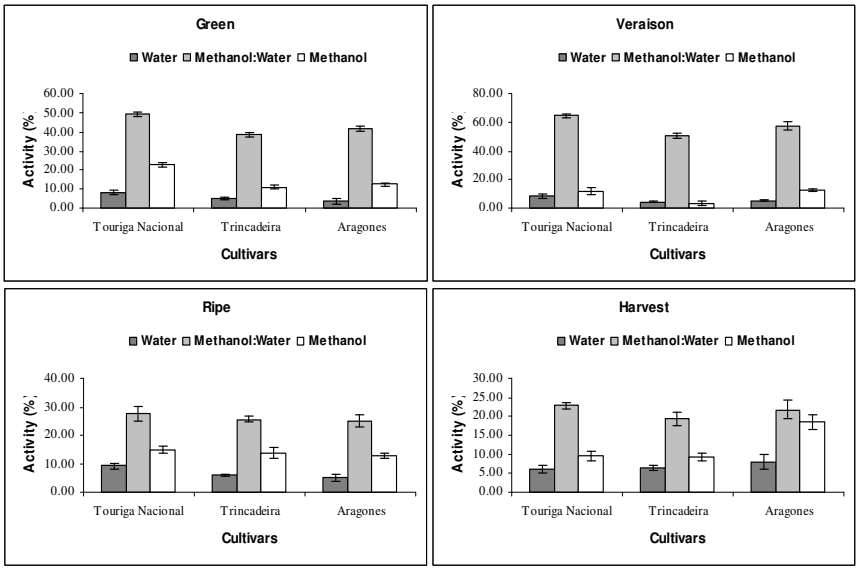

Figure 3. TNFα inhibition ratio (%) exhibit by different grape cultivars at same developmental stages. Bars represent the Mean ± S.D. (n=3) at p<0.01.

All the methanol:water extracts of the three cultivars show variable activity at different developmental stages (Figure 2). The veraison stage is found to have maximum anti-TNFα activity in every cultivar followed by the green stage. For two cultivars, 'Touriga Nacional' and 'Aragonês', the ripe and harvest stages are not significantly different in anti-TNFα activity. The ripe stage of 'Trincadeira' is significantly higher than the harvest stage in inhibiting the TNFα production. Among the green stages of all three cultivars, the green 'Touriga Nacional' grapes are found more active than the green grapes of 'Trincadeira' and 'Aragonês' (Figure 3). At veraison, the 'Aragonês' and 'Touriga Nacional' grapes are not different but both are significantly more active than 'Trincadeira'. All the three cultivars show similar potency of inhibiting TNFα production at the ripe and harvest stages.

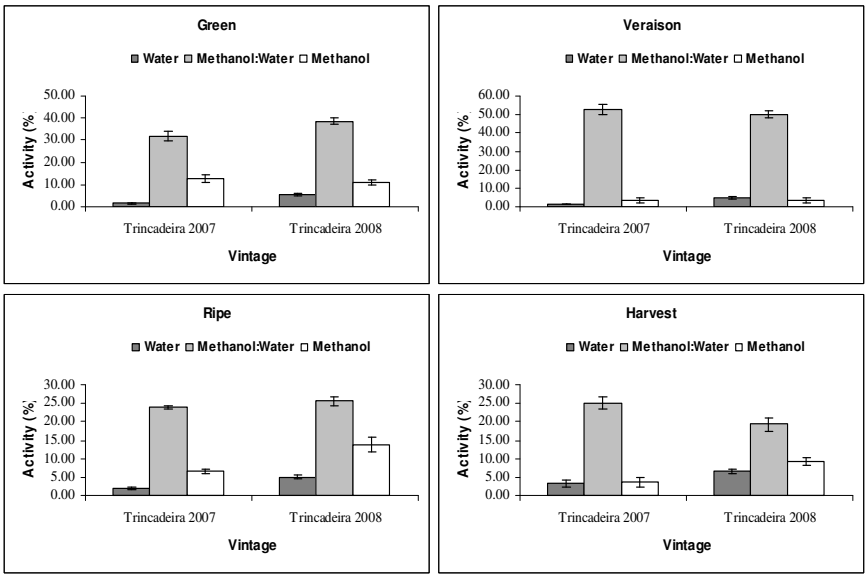

Figure 4. TNFα inhibition ratio (%) exhibit by 'Trincadeira' 2007 and 'Trincadeira' 2008 at similar developmental stages. Bars represent the Mean ± S.D. (n=3) at p<0.01.

In order to highlight the vintage effect on anti-TNFα activity, the 2007 and 2008 vintage of 'Trincadeira' cultivar are compared. 'Trincadeira' 2007 also shows highest anti-TNFα activity at veraison stage followed by green stage but unlike 'Trincadeira' 2008 (see above), the 2007 vintage shows no significant difference in TNFα inhibition at later stages of development i.e. ripe and harvest (Figure 2). Comparing every developmental stage of these two vintages, only green and harvest are different. Green 'Trincadeira' 2008 grapes show higher activity while at harvest 'Trincadeira' 2007 grapes show significant inhibition of TNFα production (Figure 4). The ^{1}H NMR spectra analysis shows that 'Trincadeira' 2007 has more phenolics as compared to 'Trincadeira' 2008, thus suggesting a relationship between phenolics and activity.

Cultivars, development stages, and SPE fractions differentiation

The NMR data from the SPE fractions of all the samples have been subjected to PCA in order to identify possible markers for the different cultivars, developmental stages, and SPE fractions. Figure 5 shows the score plots of PCA where samples are colored according to SPE fractions, cultivars, and developmental stages.

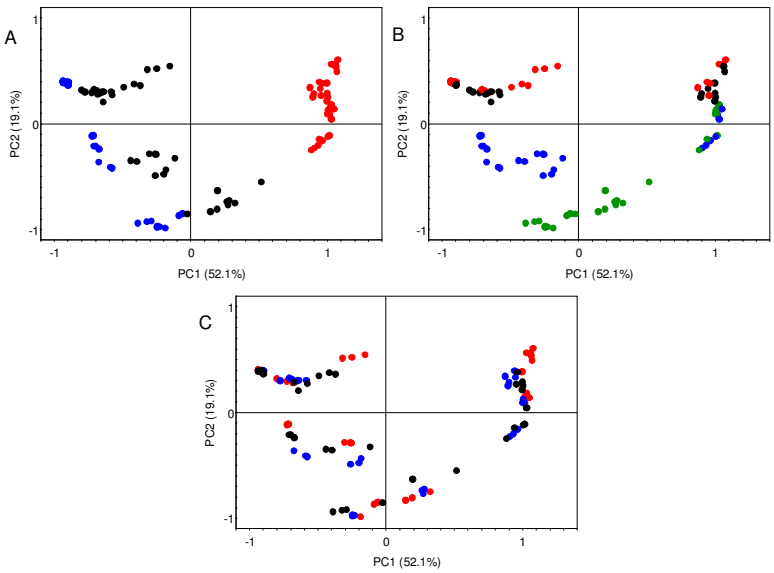

Figure 5. Principal component analysis score plots. In 'A', samples are colored according to SPE fractions where red represents methanol, black represents methanol:water (1:1), and blue represents water fraction. In 'B', samples are colored according to developmental stages where green represents green stage, blue represents veraison, black represent ripe, and red represents harvest stage. In 'C', samples are colored according to grape cultivars where red represents samples from Aragone, blue represents samples from 'Touriga Nacional', and black represents samples from 'Trincadeira'.

Figure 5A represents the PCA score plot where samples are colored according to SPE fractions. It is clear from the figure that PC1 (52.1%) is responsible for the separation of all three SPE fractions. The water fractions are clustered on the negative side of PC1 while the methanol fractions are grouped on the positive side of PC1. The methanol:water fractions are located in between the methanol and water fractions, mostly having negative PC1 values. To highlight the differences based on developmental stages, samples from the same PCA are colored according to developmental stages in Figure 5B. It is obvious from the score plot that while PC1 is responsible for the separation of SPE fractions, PC2 (19.1%) is quite effective in discriminating the developmental stages of grapes. The initial stages, like green and veraison, are on the negative side of PC2 whereas the remaining stages, like ripe and harvest, mostly have positive PC2 scores. The same PCA score plot is shown in Figure 5C but this time the samples are colored according to grape cultivar. It is evident from this figure that PCA is not very effective as the samples are not clustered based on the grape cultivars.

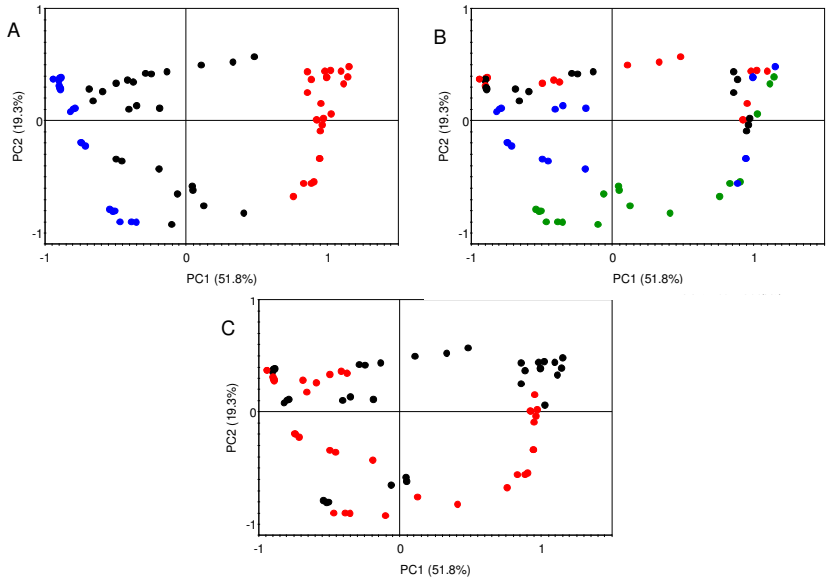

Figure 6. Principal component analysis score plots. In 'A', samples are colored according to SPE fractions where red represents methanol, black represents methanol:water (1:1), and blue represents water fraction. In 'B', samples are colored according to developmental stages where green represents green stage, blue represents veraison, black represent ripe, and red represents harvest stage. In 'C', samples are colored according to 'Trincadeira' vintages where red represents samples from 'Trincadeira' 2008, and black represents samples from 'Trincadeira' 2007.

The corresponding loading plots with the respective NMR spectra reveal the information regarding the metabolites responsible for the differentiation of samples on the score plots. As also shown by Figure 1A, the SPE fractions are very much distinct in their metabolic contents. The water fraction is relatively higher in sugars and some amino acids and organic acids. Most of the phenolics are found to elute in the methanol:water fraction while the remaining amino acid and phenolics come out with the last methanol fraction. The methanol fraction is also found higher in fatty acids. The PCA is also found very effective in discriminating the developmental stages and by examining the loading plot the metabolites involved in differentiation could be identified. The grapes in green and veraison stages have higher levels of phenolics with relatively less sugar. As the berries grow the level of phenolics start to decrease whereas sugars concentrations increase. A detailed account of the distribution of metabolites based on these cultivars at these development stages is given in Chapter 5.

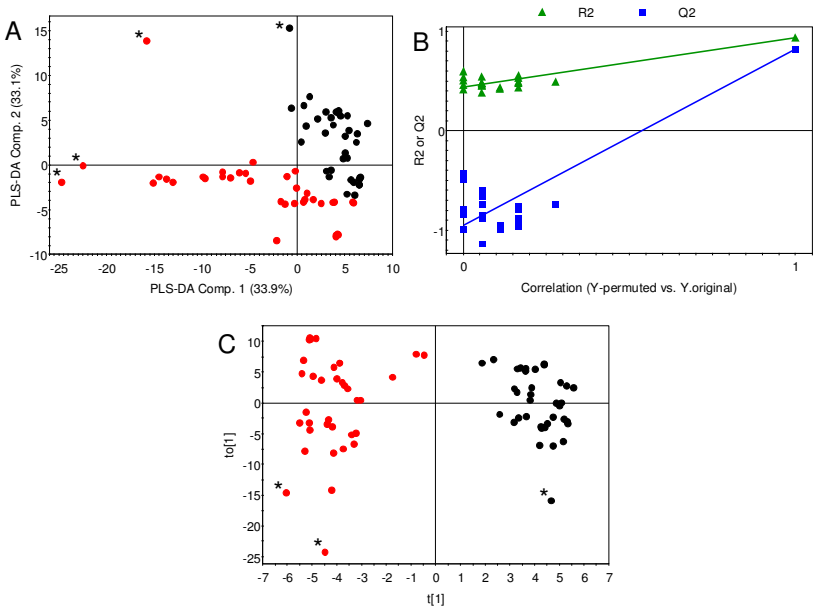

Figure 7. Multivariate data analysis of 'Trincadeira' 2007 and 2008. The PLS-DA score plot (A), permutation test for PLS-DA (B), and score plot of O2PLS-DA (C) are shown. Samples with black color are from 'Trincadeira' 2007 while samples with red color are from 'Trincadeira' 2008. Samples with '*' represents an outlier.

Vintages, development stages, and SPE fractions differentiation

SPE fractions of grapes from 2007 and 2008 vintages of the 'Trincadeira' cultivar at four developmental stages were analyzed and compared for metabolic differences. Principal component analysis, also in this case, is found effective in highlighting the metabolic differentiation among the samples based on developmental stages and SPE fractions and responsible metabolites are identified (Figure 6). As discussed above, phenolics are the main discrimination factor in SPE fractions while a similar metabolic behavior of developmental stages was observed in all grape cultivars. However, to analyze specifically the vintage effects on the grape metabolic profile, supervised multivariate data analysis was applied.

First, projections to latent structures-discriminant analysis (PLS-DA) was used in which samples are classified in to two classes based on samples from 2007 and 2008 vintages. The score plot (Figure 7A) shows good separation among the samples belonging to the two different classes but none of the components is found totally effective. The PLS-

DA model was validated using permutation test with 20 applications (Figure 7B). To draw clear conclusions, bidirectional orthogonal projections to latent structures-discriminant analysis (O2PLS-DA) was applied. The score plot (Figure 7C) shows very clear distinction among the different vintages. Component 1 is responsible for the separation as samples from 2007 and 2008 vintages are on the positive and negative side, respectively. The loading plot shows that the 2007 vintage has higher levels of phenolics than the 2008 vintage. The 2008 vintage shows elevated levels of organic acids like malate and citrate with some sugars like glucose and fructose.

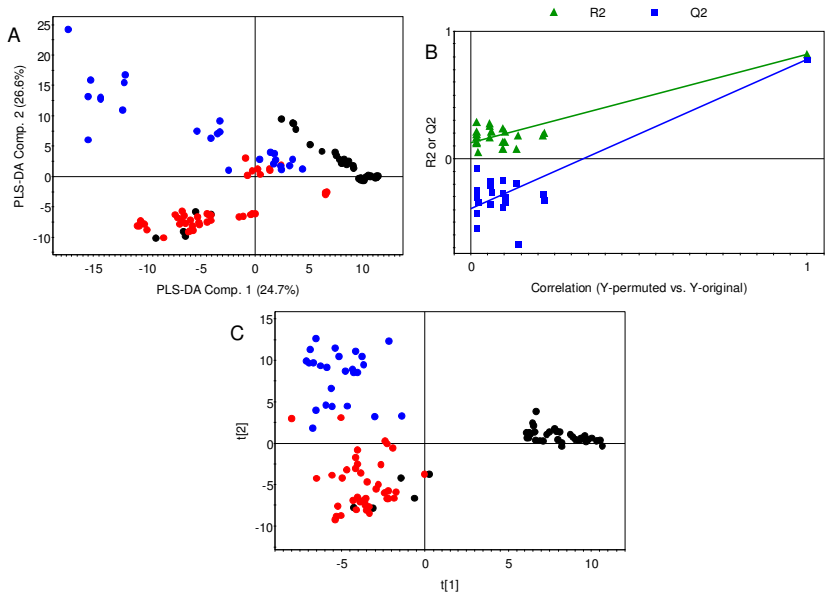

Figure 8. Score plots of PLS-DA (A) and O2PLS-DA (C) are represented. Samples with black color are of low anti-TNFα activity while samples with red and blue colors are of medium and high anti-TNFα activity, respectively. The permutation test for PLS-DA (C) is also presented.

Multivariate data analysis for metabolome and activity correlation

Projections to latent structures-discriminant analysis (PLS-DA) is considered as the second step of metabolomics studies. It is a supervised method in which samples are classified into different groups on the basis of creating a set of dummy *Y*-variables. In this study, based on activity data, we classify our samples into low (<10%), medium (≥10% and <25%), and high (≥25%) activity classes as *Y*-variables, and used these in a PLS-DA. Figure 8A shows that this gives a clear separation especially the samples with

the lowest and medium activity are grouped separately. Samples with high activity are scattered and some are mixed with the samples with medium activity.

Validation of the supervised regression model is one of the key aspects in such analysis. For supervised algorithms, like PLS and PLS-DA, permutation test is used for validation which is a calculation of goodness of fit and the predictive ability of the model, represented by R2 and Q2, respectively. Generally R2 describes how well the data in the training set are mathematically reproducible. The R2 value can be vary from 0 to 1, where 1 means a model with a perfect fit. If the Q2 value is more than 0.5, the model is considered to have good predictability and if it is higher than 0.9 and less than 1.0, then the model is considered to have an excellent predictability. It is suggested that if more than five PLS components are included in the model the training set data generally reproduce excellently. The R2 and Q2 values of PLS-DA using six components were calculated. For anti-TNFα activity the figures were 0.82 and 0.78, respectively. This PLS-DA model was validated by the permutation method through 20 applications in which all Q2 values of permuted Y vectors were lower than original ones and the regression of Q2 lines intersect at below zero (Figure 8B).

In order to get better separation, especially for the samples with high activity, bidirectional orthogonal projections to latent structures-discriminant analysis (O2PLS-DA) is applied. The score plot of O2PLS-DA (Figure 8C) shows much better separation among the samples based on anti-TNFα activity. Samples with the low activity are grouped separately on the positive side of component 1 while the samples with medium and high activity are having negative component 1 scores. Component 2 was found effective in separating the high and medium activity samples as they clustered on positive and negative side of component 2, respectively. Few samples from low activity and medium activity classes and mixed with the medium activity and high activity classes, respectively, as their anti-TNFα activity values are on the border line of their classes. The O2PLS-DA model is validated by cross validation-analysis of variance (CV-ANOVA) with a *p*-value of 8.35×10^{-38}. By examining the corresponding loadings plot, the metabolites responsible for separation are identified. Samples with different activity levels mainly differ in their phenolic contents. The high anti-TNFα activity samples have higher levels of phenolics like cinnamic acids, flavonols, and flavan-3-ols while the medium and low activity samples have less or no phenolic contents.

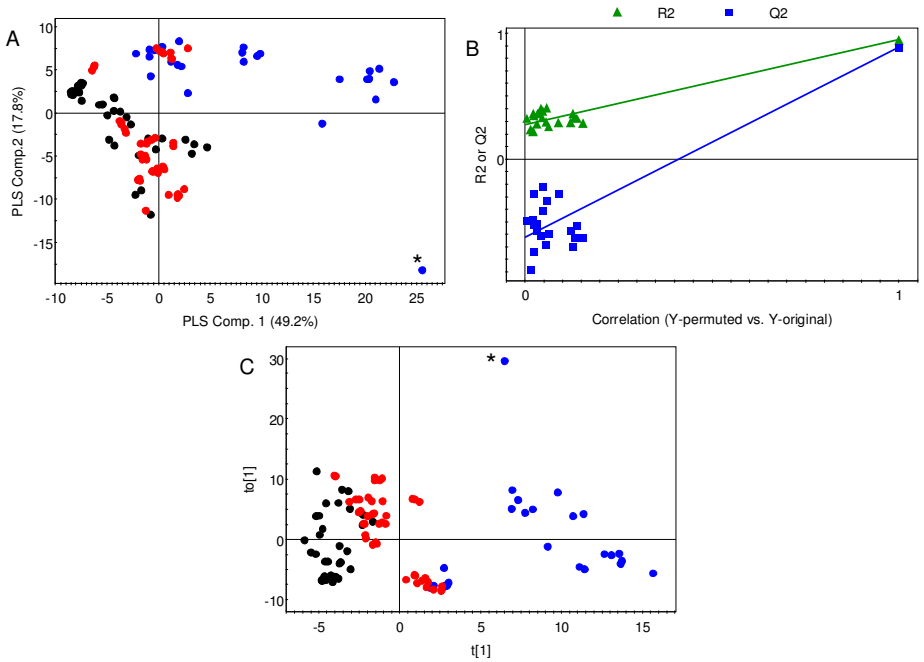

Figure 9. Score plots of PLS (A) and O2PLS (C) are represented. Samples with black color are of low anti-TNFα activity while samples with red and blue colors are of medium and high anti-TNFα activity, respectively. The permutation test for PLS (C) is also presented.

The next step is to perform the direct correlation between the activity and NMR data using original anti-TNFα assay values. Instead of classifying samples as high, medium, and low activity groups, the activity data from TNFα assay for each sample are used directly. In such approaches two different data sets, independent variable (like NMR spectral data) and dependent variable (like anti-TNFα activity data), are correlated using regression. For this purpose projections to latent structures (PLS) analysis was performed using the NMR and activity data. The PLS score plot (Figure 9A) shows relatively good separation among the samples but many are overlapped with the other groups. Component 1 is mainly responsible for the separation as the samples are arranged from low to high activity along the negative to positive side of component 1, respectively. For PLS modeling again the permutation method through 20 applications was used for validation. The regression of Q2 lines intersect at below zero with all Q2 values of permuted Y vectors were lower than original. Variance (R2) and cross-

validated variance (predictive ability of the model, Q2) values of PLS using seven components were calculated and for anti-TNFα activity the figures were 0.95 and 0.89, respectively (Figure 9B).

Finally for the identification of metabolites responsible for high activity in grapes, we used another multivariate data analysis method known as bidirectional orthogonal projections to latent structures (O2PLS). Analyses like PLS regression can cause systematic variation of any data block due to structured noise present in the data blocks. Other algorithms, like O2PLS-DA and O2PLS, are multivariate projection methods which remove the structured noise by extracting linear relationships from independent and dependent data blocks, in a bidirectional way, and results in the decomposition of systematic variation into two model parts: the predictive or parallel part and the orthogonal part (Trygg and Wold 2002; 2003). The score plot, Figure 9C, shows very nice separation among low, medium and high activity samples based on component 1. This O2PLS model was validated by CV-ANOVA with *p*-value of 1.4×10^{-37}. Like PLS-DA and O2PLS-DA the corresponding loadings plot show that the samples with high anti-TNFα activity contained more phenolics, such as cinnamates and flavonoids, when compared to samples with low and medium activity.

In PLS based regression, VIP (variable importance in the projections) can be defined as a weighted sum of squares of the PLS weights. It has been indicated that it is directly proportional with the influence of factor on the separation on score plot, meaning, factors have higher VIP values are more important for the samples separation. For O2PLS-DA and O2PLS analyses, VIP values for several phenolic compounds, responsible for separation on the score plot, are presented in Table 2. As indicated, in O2PLS-DA analysis, caftaric acid, and (+)-catechin are the metabolites with top two VIP scores while quercetin and myricetin have relatively less VIP values. In O2PLS model, again (+)-catechin is one of the top two followed by coutaric acid while caftaric acid shown a much lower VIP score. This high VIP scores for the identified phenolics legitimate their involvement in the separation of high activity samples and suggest a role of these compounds in inhibiting TNFα production.

Table 2. The VIP (variable importance in the projections) values of the major contributing compounds for the separation in the score plots derived from O2PLS-DA and O2PLS models.

Compounds	Chemical shift (ppm)	VIP values	
		O2PLS-DA	O2PLS
Quercetin-3-*O*-glucoside	6.27	1.38	2.06
Myricetin	6.51	1.17	1.62
(+)- Catechin	6.75	1.44	2.16
(-)- Epicatechin	6.96	1.40	2.09
Caftaric acid	7.62	1.55	1.80
Fertaric acid	6.32	1.43	2.08
p-Coutaric acid	7.65	1.41	2.10

Discussion

The ^{1}H NMR spectra clearly shows distinction among different cultivars and vintages, their developmental stages, and the SPE fractions. This advocates the enormous analytical potential of NMR spectroscopy as compared to other platforms for metabolomics studies (Verpoorte et al. 2008). Multivariate data analysis in combination with NMR is very popular in metabolic phenotyping studies of plants. Many reports have been published regarding grape berries using the same approach (Son et al. 2009b; Pereira et al. 2006b). In this study, metabolic profiling of different grapes at different stages of ripening has been successfully performed. The initial stages, green and veraison, have been characterized with high phenolics, whereas high sugar content is observed in the later stages i.e. ripe and harvest, as also reported previously (Ali et al. 2011). This metabolic distinction among the developmental stages reflect in the associated anti-TNFα activity as green and veraison are found more active than ripe and harvest.

The vintage effect on grape metabolome is quite obvious now as it is widely accepted that the several climatic factors are involved in the biosynthesis of several key metabolites in grapes (Pereia et al. 2006). The green and harvest stages of these two vintages present significant metabolic differences, characterized by higher and lower phenolic contents in 2008 vintage, respectively, as compared to 2007 vintage. Since vintage has shown its effects on the phenolic contents of 'Trincadeira', the anti-TNFα activity shown by these vintages is also affected. The green and harvest stages from

2008 and 2007 vintages, respectively, showed significantly different anti-TNFα activity. As shown by the NMR spectra, this is due to difference in phenolic contents. It has been reported that different factors like hot and dry climate can result in higher phenolic contents in grapes (Pereia et al. 2006). For instance, the insolation totals were higher in July and August of 2007 and differences in rain totals and average temperature were also observed in between seasons and may influence the fine tuning of phenolics´ biosynthesis (unpublished data). It is interesting to note that transcriptomic analysis using Affymetrix GrapeGen® genome array showed that a gene coding for anthocyanidin reductase which is involved in proanthocyanidins biosynthesis such as catechin was more expressed in 2007 samples. Since (+)-catechin seems to present high anti-TNFα activity as suggested by the results hereby presented this may constitute a good example of positive integration of transcriptomic and metabolomic data, and medicinal properties that deserves further attention.

Data correlation using different multivariate data analysis tools is now increasingly popular and found efficient in predicting the unknown NMR signals (metabolites) by using the resulting training model (Eriksson et al. 2006). Many reports have been published targeting to develop the predictive models for certain pharmacological activities in plants. Plants like St. John Wort (*Hypericum perforatum*) (Roos et al. 2004), *Artmesia annua* (Bailey et al. 2004), *Citrus grandis* (Cho et al. 2009), and *Galphimia glauca* (Cardoso-Taketa et al. 2008), have been efficiently studied for the prediction of different medicinal properties, using this approach. Such chemometrics based approach can provide first hand knowledge regarding the plant extracts and any related bioactivity without any tedious chromatographic separations.

Since grapes are one of the richest sources of polyphenolics, many studies (Chuang et al. 2011; Seymour et al. 2008) have shown their potency against TNFα production as grape polyphenolics are widely acclaimed and accepted to have as anti-oxidative and anti-inflammatory properties (Baur et al. 2006; Breksa et al. 2010). Phenolics in grapes, like resveratrol (Stewart et al. 2008) and quercetin (Rivera et al. 2008; Teissedre et al. 1996) are known to reduce inflammation, while others like cinnamates, benzoates, flavonols, flavan-3-ols, and anthocyanins, are well known antioxidants (Meyer et al. 1996; Lee et al. 2009). The present study is the only known attempt to analyze different grape cultivars, their developmental stages, and vintages for TNFα inhibition. The identified

NMR signals, responsible for the activity, are related to quercetin, myricetin, (+)-catechin, (-)- epicatechin, coutaric acid, fertaric acid, and caftaric acid, which are found relatively higher in the samples with high activity using different chemometrics methods.

Conclusion

Nuclear magnetic resonance spectroscopy (1D and 2D) has been applied for the phenotyping of three grape cultivars from Portugal at different stages of their development. Solid phase extraction was used in order to separate the phenolics from the other components of grape metabolome which resulted in water, methanol:water (1:1), and methanol fractions and have been tested for TNFα inhibition. The assay showed that grapes from all three cultivars at veraison and green stages strongly inhibit TNFα production. Various multivariate data analysis methods showed good correlation between the NMR resonances for phenolics and anti-TNFα activity. Algorithms like PLS and PLS-DA showed good separation among the samples classified as high, medium, and low activity with high model validity. The application of bidirectional orthogonality, i.e. O2PLS-DA and O2PLS, showed even better distinction among the classes. The VIP plot showed that NMR signals related to metabolites like quercetin, myricetin, (+)-catechin, (-)-epicatechin, caftaric acid, and coutaric acid, were statistically significantly correlated with high activity. Using the presented approach, the analysis of NMR shifts in relation to pharmacological activity can provide information about what part of the NMR spectrum (aromatic or aliphatic regions) correlates with the activity which gives information about the active ingredients in crude extracts of medicinal plants. In the approach shown here the compounds related to activity can be identified without extensive and elaborate chromatographic separation, and thus allows rapid identification of extracts with biological activity.