Aberystwyth University

An Analytical Pipeline for Quantitative Characterization of Dietary Intake
Garcia-Perez, Isabel; Posma, Joram M; Chambers, Edward S.; Nicholson, Jeremy K.; Mathers, John C.; Beckmann, Manfred; Draper, John; Holmes, Elaine; Frost, Gary

An Analytical Pipeline for Quantitative Characterization of Dietary Intake: Application To Assess Grape Intake

Isabel Garcia-Perez†, Joram M. Posma‡, Edward S. Chambers†, Jeremy K. Nicholson‡, John C. Mathers§, Manfred Beckmann‖, John Draper‖, Elaine Holmes‡, and Gary Frost*†

†Nutrition and Dietetic Research Group, Division of Endocrinology and Metabolism, Imperial College London, London W12 0NN, U.K.
‡Computational and Systems Medicine, Department of Surgery and Cancer, Imperial College London, London SW7 2AZ, U.K.
§Human Nutrition Research Centre, Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne NE4 5PL, U.K.
‖Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth SY23 3DA, U.K.

Abstract

Lack of accurate dietary assessment in free-living populations requires discovery of new biomarkers reflecting food intake qualitatively and quantitatively to objectively evaluate effects of diet on health. We provide a proof-of-principle for an analytical pipeline to identify quantitative dietary biomarkers. Tartaric acid was identified by nuclear magnetic resonance spectroscopy as a dose-responsive urinary biomarker of grape intake and subsequently quantified in volunteers following a series of 4-day dietary interventions incorporating 0 g/day, 50 g/day, 100 g/day, and 150 g/day of grapes in standardized diets from a randomized controlled clinical trial. Most accurate quantitative predictions of grape intake were obtained in 24 h urine samples which have the strongest linear relationship between grape intake and tartaric acid excretion ($r^2 = 0.90$). This new methodological pipeline for estimating nutritional intake based on coupling dietary intake information and quantified nutritional biomarkers was developed and validated in a controlled dietary intervention study, showing that this approach can improve the accuracy of estimating nutritional intakes.

INTRODUCTION

Higher consumption of vegetables and fruits is associated with lower all-cause mortality, and recent studies suggest that daily intakes of 5[1] or 7 or more portions of vegetables and fruits[2] lowers the risk of death, particularly cardiovascular mortality. Such evidence is the basis for government recommendations to improve eating patterns. However, monitoring compliance to dietary advice at the population level is extremely difficult because existing dietary assessment tools based on self- reporting methods are inherently inaccurate.[3,4] Under-reporting was found to be 34% in men and 33% in women, with the highest occurrence of under-reporting in obese and overweight individuals.[5,6] There is an unmet need for quantifiable dietary biomarkers that accurately reflect consumption of foods and nutrients.[4] Dietary intake biomarkers are based on the concept that excretion levels of food-related metabolites are highly correlated with food intake over a given period of time. These biomarkers can be components of the food itself, excreted unchanged, or compounds that have undergone metabolic conversion by the human or by the resident gut bacteria. Global metabolic profiling using spectroscopic technology has been applied to detect food-derived compounds[7] including polyphenols (berries),[8] alkyl resorcinols (wheat),[9] proline betaine (citrus fruit),[10] (N-acetyl-)S-methyl-L-cysteine sulfoxide (cruciferous vegetables), [11] and trimethyl-amine-N-oxide and methylhistidine (oily fish),[8] which are potential biomarkers of original food intake. Metabolic profiling allows an independent and objective assessment of food intake from which energy and nutrients can be calculated. While much of this metabolic profiling research was conducted using panels of metabolites or excretion patterns, few studies have established quantitative relationships between amounts of specific foods consumed and food-derived metabolites excreted. Exceptions include the following: total urinary nitrogen for protein intake, urinary potassium and sodium outputs for potassium and sodium intake, [12] and proline betaine derived from citrus fruit consumption, which exhibits total clearance within 24 h. [13] However, as with many putative biomarkers of individual foods, proline betaine is not entirely specific to citrus fruits and can be found in low concentrations in alfalfa, pulses, kiwi, and pears. [10,14] Nevertheless, proline betaine has been shown to be a robust and quantifiable dietary biomarker of citrus intake. [10,13]

Specific foods may have benefits for specific diseases. For example, consumption of grapes and grape-based products has potential efficacy in cancer prevention [15–18] and is associated with decreased risk for cardiovascular disease; [19] however, there is no irrefutable evidence for the health benefits of grapes from randomized

controlled intervention studies. A methodology for generating reliable markers of specific foods such as grapes would be a valuable addition to the armory of nutritional tools. Here we develop and apply a rapid, accurate, and efficient analytical pipeline for assessing food intake using nuclear magnetic resonance (NMR) spectroscopy and demonstrate its application in measuring urinary concentrations of tartaric acid as an indicator of grape intake.

MATERIALS AND METHODS

Identification of a candidate biomarker for grape intake.
To explore potential quantifiable biomarkers of intake of grapes, we designed a grape challenge pilot study that was undertaken to characterize the urinary metabolite excretion profile following consumption of grapes. Six volunteers (3 men and 3 women, age range: 22–32, BMI range: 21.2–25.3 kg/m^2) were recruited and participated in a 3-day study. Participants were healthy, nonsmokers and did not consume drugs or food supplements regularly. The grapes were administered as part of a standard breakfast consisting of one cup of tea or coffee and a fruit salad containing apple, pineapple, and red grapes. The amount of red grapes in the consumed fruit salad was increased over the three consecutive days viz. 50 g, 100 g and 200 g of red grapes on days 1, 2 and 3, respectively, while apple and pineapple intake remained constant. For the 24 h preceding the grape challenge and throughout the remaining part of each experimental day, participants ate their habitual diet avoiding any products containing grapes, raisins or food products derived from grapes. All urine produced was collected daily into four timed aliquots per day corresponding to 0–4 h, 4–8 h, 8–12 h and 12–24 h post consumption, using single-use urine containers (International Scientific Supplies Ltd., Bradford, United Kingdom). In addition, a spot urine sample was collected on day 1 prior to consumption of grapes to provide a baseline profile. Urine samples were stored at −80 °C until analysis.

Compositional analysis of grapes study.
To characterize the chemical composition of varieties of grapes consumed in the UK, ten varieties of red grapes and two varieties of green grapes from five countries were selected according to seasonal availability. Two batches of 400 g of each variety of grapes were purchased and for each variety 50 g of grapes were picked

from random areas of different clusters to provide a representative sample. A total of three replicates of each batch were prepared. Each individual sample was homogenized using a Kenwood KMix Blender for 5 min and, approximately, 50 mL of must were obtained and then filtered using a stainless steel filter. An aliquot of 1 mL of each sample was centrifuged for 5 min at 16,000xg, and the supernatant fluid was analyzed by ^{1}H NMR spectroscopy. Quantification of the tartaric acid concentration in red and green grapes was carried out using a standard one-dimensional NMR pulse sequence ensuring fully relaxed presaturation of the water resonance.

^{1}H NMR spectroscopic analysis of urine and grape samples.
An aliquot of each urine sample (540 µL) was mixed with 60 µL of pH 7.4 phosphate buffer containing trimethylsilyl-[2,2,3,3,-^{2}H4]-propionate (TSP) as an internal chemical shift reference before being prepared for the NMR spectroscopic analysis following the protocol described by Dona et al.[20] Urine samples were analyzed in 96-well plates containing one quality control (QC) sample every ten samples. QC samples were prepared by pooling 50 µL of each urine sample. Filtered homogenates (400 µL) of representative samples of the grapes given to participants were mixed with 200 µL of pH 7.4 phosphate buffer. QC samples of grape homogenates were prepared by pooling 50 µL of each grape fluid sample and analyzed every ten samples.

^{1}H NMR spectroscopy was performed at 300 K on a Bruker 600 MHz spectrometer (Bruker BioSpin, Karlsruhe, Germany) using the following standard one-dimensional pulse sequence with saturation of the water resonance: RD–gz,1–90°–t–90°–tm–gz,2–90°–ACQ, where RD is the relaxation delay, t is a short delay typically of about 4 µs, 90° represents a 90° radio frequency (RF) pulse, t_m is the mixing time (10 ms), gz,1 and gz,2 are magnetic field z-gradients both applied for 1 ms, and ACQ is the data acquisition period (2.7s). Water suppression was achieved through continuous wave irradiation at the water resonance frequency using 25 Hz RF strength during RD and also during tm. The receiver gain was set to 90.5 for all experiments. Each urine spectrum was acquired using 4 dummy scans, 32 scans, 64K time domain points and with a spectral window set of 20 ppm. Prior to Fourier transformation, the free induction decays were multiplied by an exponential function corresponding to a line broadening of 0.3 Hz. To achieve accurate quantification of metabolites, it was necessary to ensure that both tartaric acid and TSP resonances were fully relaxed using a long RD (7 × t1) between each pulse. The interpulse delay time d1 was therefore set to 100s.[21]

Preprocessing of NMR spectra.

The ^{1}H NMR spectra were digitized over the range of δ10.0 to −0.5 and imported into MATLAB (2014a, Mathworks Inc., USA), and automatically phase- and baseline-corrected. Urine spectra were then referenced to the internal standard, TSP at δ0.0. The spectral regions occupied by water and urea (δ4.45– 6.95) and TSP (<δ0.35) were excluded. Each spectrum was normalized to the total urine volume excreted in order to correct differences in urinary dilution. Spot urine samples were normalized to the spot volume. Each spectrum of grape juice extract was phased and baseline corrected as above, and normalized to the TSP signal.

Statistical analysis.

Urinary global profiling analysis in combination with unsupervised Principal Component Analysis (PCA) [22] and supervised Partial Least Squares Discriminant Analysis (PLS-DA) [23] was applied to identify candidate biomarkers of grape intake. PCA was used to visualize any trends in the data and these trends were then further analyzed using PLS-DA and Monte Carlo cross-validation (MCCV). [24] For each component in the PCA model the percentage of explained variance (R^2_X) was calculated. The variability of the predicted scores was visualized using Kernel Density Estimation (KDE). Across all MCCV models, the mean score and variance of the score for each sample were used to calculate the KDE. Specifically, for each sample the mean and standard deviation across all MCCV models can be used to give a normal distribution of the predicted scores. Taking the sum of all individual distributions within each class then yields the KDE as shown in the figures. The stability (variance) of each regression coefficient was assessed across the MCCV models using bootstrap resampling [25] of each model. Using the variance and mean regression coefficient, a t-score, and subsequently a P-value, were calculated. The P-values were corrected for multiple testing using the Storey-Tibshirani [26] False Discovery Rate (q-value). The goodness of fit (R^2_Y) of the MCCV models was calculated across all models using the training data and the goodness of prediction (Q^2_Y) for the test data.

Hierarchical cluster analysis (HCA) was used to investigate correlations among the identified biochemical components of different variaties of grapes, specifically to determine whether the concentration of tartaric acid shows similarity with other compounds. To quantify the relationship, HCA was used to determine clustering in the data. Significance of the correlation between pairs of compounds was assessed based on a Bonferroni correction to the P-values. HCA was applied to the resulting correlation matrix and the optimal number of clusters was determined by calculating the modularity [27] of the network and choosing the highest modularity

as optimal number of clusters. The modularity is a weighted measure between the number of links (correlations) within a cluster and the number of links from one cluster to other clusters, with clusters defined by cutting at different heights of the hierarchical clustering tree. The highest modularity indicates that, relatively, there are more within-cluster links compared to between-cluster links.

Standard linear regression was used to relate grape intake to tartaric acid excretion. The variability of the regression coefficient is visualized in corresponding plots as the 95% confidence intervals (CI) estimate by bootstrap resampling of the regression coefficient. The squared correlation coefficient of the model (r^2) is a measure of linearity of the data and corresponding model. Bland–Altman [28] plots were used to show the agreement between the estimated intake of grapes versus the real intake of grapes.

PCA and PLS-DA analyses were also carried out on spectral data of grape homogenates normalized to volume in MATLAB to compare different grape varieties. Correlation of spectral variables using Statistical Total Correlation Spectroscopy[29] and Subset Optimization by Reference Matching[30] were used as data-driven approaches to aid metabolite identification. Confirmation of metabolite identities was obtained using 1D and 2D NMR experiments (spiking of chemical standards, J-Resolved spectroscopy, Total Correlation Spectroscopy, Heteronuclear Single Quantum Coherence spectroscopy).

Quantification of tartaric acid related to grape intake.
Nineteen volunteers (10 male and 9 female, age range: 25–60, BMI range: 21.1–33.3 kg/m2) attended the NIHR/Wellcome Trust Imperial CRF for four 4-day inpatient periods separated by a period of >3 days. Potential subjects were excluded if they had clinically significant illnesses, were taking prescription medication, current3 smokers, a history of substance abuse, and any abnormalities detected on physical examination, electrocardiography, or screening blood tests (measurement of complete blood count, electrolytes, fasting glucose and lipids, thyroid function and liver function). Women were ineligible if they were pregnant or breast-feeding. In a random order, participants were provided with four different diets throughout each of the 4-day inpatient periods. Each of the four diets represented 25% (diet 1), 50% (diet 2), 75% (diet 3) and 100% (diet 4) of the healthy eating targets based on UK recommendation for fruits, fats, sugars, vegetables, carbohydrates, fiber and salt. Grapes were consumed as an afternoon snack 2 h after lunch. Alcohol- and grape-derived products were not provided as part of any of the four diets.
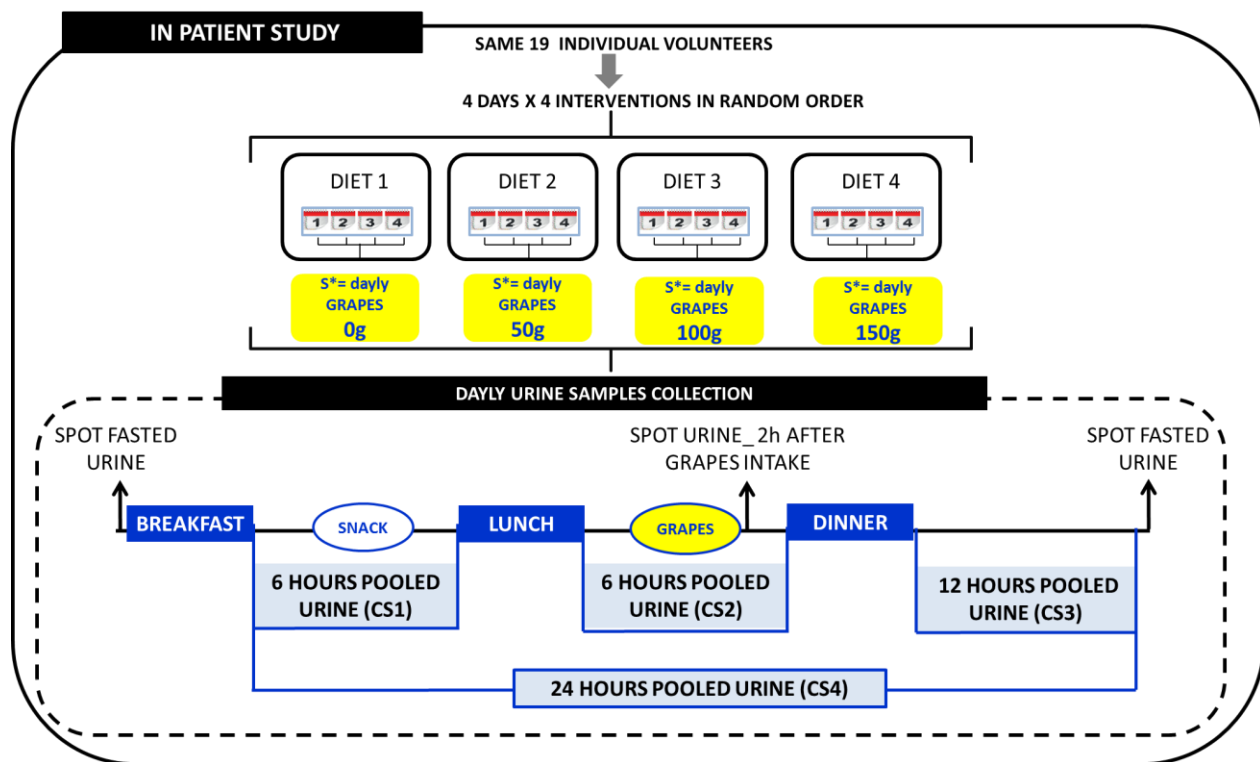
Figure 1. Study design for the controlled dietary intervention study (n=19). The study is conducted against the background of a standardized diet where incremental amounts of grapes were provided as an afternoon snack to the volunteers.

The amounts of red grapes administered daily were 0 g (diet 1), 50 g (diet 2), 100 g (diet 3), and 150 g (diet 4) (Figure 1). The randomization procedure was conducted by an investigator not directly involved in the study with the use of opaque, sealed, sequentially numbered envelopes that each contained a random order for the four dietary interventions. The envelopes were stored securely and opened in sequence by an investigator (ESC) once volunteers had been recruited. Volunteers and investigators could not be blinded during data collection; however, all investigators conducting data analysis were blinded to the randomization order. Different varieties of red grapes from different countries were provided according to seasonal availability. Wine, raisins, and any fruit juice were excluded from the experimental diets. Moreover, volunteers did not take any supplements, and minimal physical activity was imposed. Fasting spot urine samples were collected on arrival at the NIHR/Wellcome Trust Imperial Clinical Research Facility and daily thereafter for the 4 days of each of four dietary interventions. Each participant collected cumulative urine samples (CS) daily, over the four-day period of each dietary intervention, from

after breakfast to before lunch (CS1), from after lunch to before dinner (CS2), and from after dinner to next day fasting urine sample (CS3). Finally, a 24 h urine sample was obtained by pooling CS1, CS2, and CS3. In addition, a spot sample was collected daily 2 h after the afternoon snack. Aliquots of urine were transferred into Eppendorf tubes and stored at −80 °C until analysis by ^{1}H NMR spectroscopy. All subjects provided informed, written consent prior to the clinical trial (Registration No: ISRCTN 43087333), which was approved by the London Brent Research Ethics Committee (13/LO/0078). All studies were carried out in accordance with the Declaration of Helsinki.

Monitoring in-patient volunteers for 24 h tartaric acid excretion.
Samples from day 3 of each dietary intervention were chosen to monitor volunteers' excretion (Figure 1) of tartaric acid over 24 h, as well as in the fasting urine sample of day 3 of the study, spot urine sample collected 2 h after grapes consumption and the fasting urine sample on the day 4 of the study.

Tartaric acid gives rise to a single peak in the NMR spectrum at $\delta 4.34$. This signal was integrated for the urine samples using an automated algorithm. [31] The amount of tartaric acid excreted in 24 h was calculated by dividing the corresponding integral by the number of ^{1}H of tartaric acid signals ($4 \times$ [1]H) and multiply with the number of ^{1}H in TSP ($9 \times$ [1]H). Then multiplying with the concentration of TSP in the sample gives the concentration of tartaric acid. Three calibrations curves corresponding to CS2, CS3 and 24 h urine samples, were built to establish the relationship between excretion of tartaric acid and the amount of grapes consumed.

Prediction of grape intake based on tartaric acid urinary excretion in a randomized highly controlled clinical trial.
Tartaric acid was quantified for each of the spot and cumulative sample sets (corresponding to 0−3 h post consumption (CS2); 3−15 h post consumption (CS3) and the 24 h urine cumulative samples) for each of the four levels of grape intake (0, 50, 100, 150 g) to investigate how accurately grape intake could be estimated in a highly controlled clinical trial. The models were trained using calibration curves built using the tartaric acid signals from urine spectra derived from samples obtained on the third day of the study (n = 304). The model was then used to predict the quantity of grape intake using samples collected on the first and the second day of each dietary intervention (n = 608).
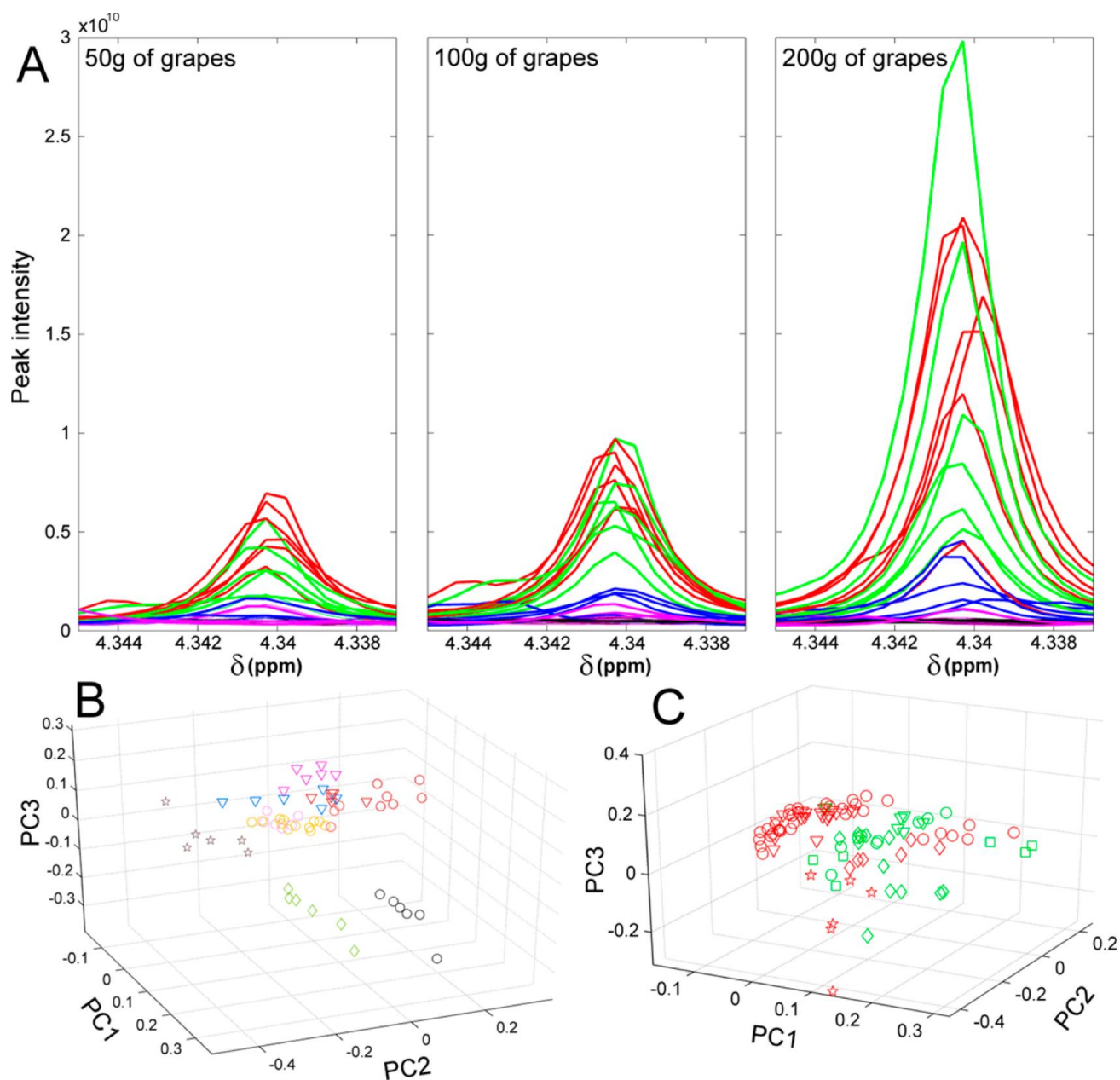
Figure 2. (A) Urinary excretion kinetics of tartaric acid in 6 volunteers after consumption of 50g, 150 g and 200g of grapes. The spectral region corresponding to the tartaric acid singlet at δ 4.34 was absent in baseline samples (0-h), visible in the 0-4-h samples (red) and 4-8-h samples (green) after grape intake, and subsequently decreased towards baseline levels in the 8-12-h samples (blue) and 12-24-h samples (magenta) after grape intake. (C) 3D-PCA scores plot of red grapes (red) vs green grapes (green) from different countries. R^2_X for principal components (PC) 1, 2, and 3 is 25%, 13%, and 11%, respectively. Key, origin: ☆ Peru; □, Namibia; □, Chile; □, India; ▽, South Africa; variety: Ralli, red circles; Magenta, magenta triangles; Jack's

Salute, blue triangles; Black Princess, black circles; Pink Muscat, pink circles; Red Globe, brown stars; Sharad, green diamonds; Flame, orange circles; Crimson, red triangles.

RESULTS AND DISCUSSION

Monitoring ^{1}H NMR urinary global profile over 24 h following grape intake.
The tartaric acid signal corresponding to a singlet at δ4.34 was identified from the global profiling analysis of the pilot human intervention study (n = 6) as a candidate marker of grape consumption. The signal for tartaric acid was absent in the baseline urine sample prior to grape consumption and showed an incremental increase in intensity as the amount of grapes consumed increased. Although other
signals from metabolites such as glucose, hippurate and 4-hydroxyhippurate were also qualitatively associated with grape intake, the tartaric acid signal was the only peak in the ^{1}H NMR global metabolic profile observed to increase proportionally with incremental grape intake (Figure 2A) over the three consecutive days in all participants. The urinary excretion kinetics of tartaric acid was calculated. In all participants, excretion of tartaric acid peaked between 4 h and 8 h postintervention and the majority of the excretion occurred in the first 12 h. Tartaric acid concentrations declined almost to baseline after the 12–24 h collection.

^{1}H NMR global profiling of red and green grapes.
To confirm the dietary origin of tartaric acid, we obtained 96 313 ^{1}H NMR global profiles of red (n = 63) and green grapes (n = 33) from different varieties and countries (as listed in Figure 2B). As expected, tartaric acid was one of the dominant compounds from the 31 metabolites identified in the global metabolic profiles of grapes (Supporting Information Figure 1). The PCA scores plot showed clustering of grapes according to variety and country (Figure 2B) with the Black Princess (Chile) and the Sharad (India) varieties being most distinctive in profile. Moreover, the PCA model including all the red and green grape samples showed a trend in clustering according to the color of the grape (Figure 2C). Systematic differences between red and green grapes were determined from the PLS-DA model (R^2_Y = 0.65, Q^2_Y = 0.52) (Figure 3A, B), including significantly higher concentrations of phenylalanine and leucine in green grapes whereas ethyl glucuronide was significantly higher in the red grapes (Figure 3C). The concentration of tartaric acid (Figure 3C, inset) was not significantly different between red and green grapes or varieties, indicating that it could have general applicability as a marker of grape intake.

Correlations between the 31 small molecules displayed in a heat-map (Supporting Information Figure 1) showed grouping of chemical components of the grape homogenate. Examples of distinct clusters, correspond to amino acids in one and for instance ethanol, acetoin and 2,3-butanediol (markers of fermentation) in another cluster. It shows that tartaric acid is an independent component in grapes as it is independent of other small molecules found in grapes.

Assessment of tartaric acid as a quantitative biomarker of grape intake in human urine samples.

As expected, the third and fourth day spot-fasting urine samples and the CS1 urine sample (cumulative sample from after breakfast to before lunch) collected during the controlled clinical trial did not contain any tartaric acid as the samples were collected before volunteers ingested red grapes (afternoon snack).

The calibration curves for tartaric acid calculated using urine samples obtained on the third day of each dietary intervention showed a linear relationship between quantity of grapes consumed and tartaric acid excreted, with the exception of the spot urine samples. Although, spot urine samples collected 2 h after grape intake showed some evidence of tartaric acid excretion, there were no significant differences in concentrations (Figure 4A) in relation to the different amounts of grapes consumed (r^2 = 0.04). The CS2 urine samples (collected from after lunch to before dinner) (Figure 4B), which contained the urine excreted in the first 3 h following consumption of grapes as an afternoon snack, showed a linear relationship with a correlation coefficient of r^2 = 0.58.

The correlation coefficient defining the relationship between the amount of grape intake and tartaric acid excretion corresponding to samples that were collected overnight (CS3:12 h collection from after dinner to next day fasting urine sample) was stronger (r^2 = 0.80) than the correlation coefficient (Figure 4C) found for the previous period (CS2). However, the 24 h calibration curve (Figure 4D) showed the highest correlation coefficient r^2 = 0.90, and therefore, it was used to quantify the total 24 h urinary tartaric acid excretion for day three. On average, 0.16, 0.30, and 0.49 mMol of tartaric acid was excreted in 24 h urine samples after eating 50, 100 and 150 g of grapes, respectively. The mean, standard deviation and 95% confidence interval of urinary tartaric acid excreted calculated for the three urine collection period: 24 h (0.161 ± 0.035 mMol (0.101–0.231)), CS3 (0.115 ± 0.040 mMol (0.026–0.195)) and CS2 (0.046 ± 0.023 mMol (0.014–0.093)) samples after eating 50 g of grapes.
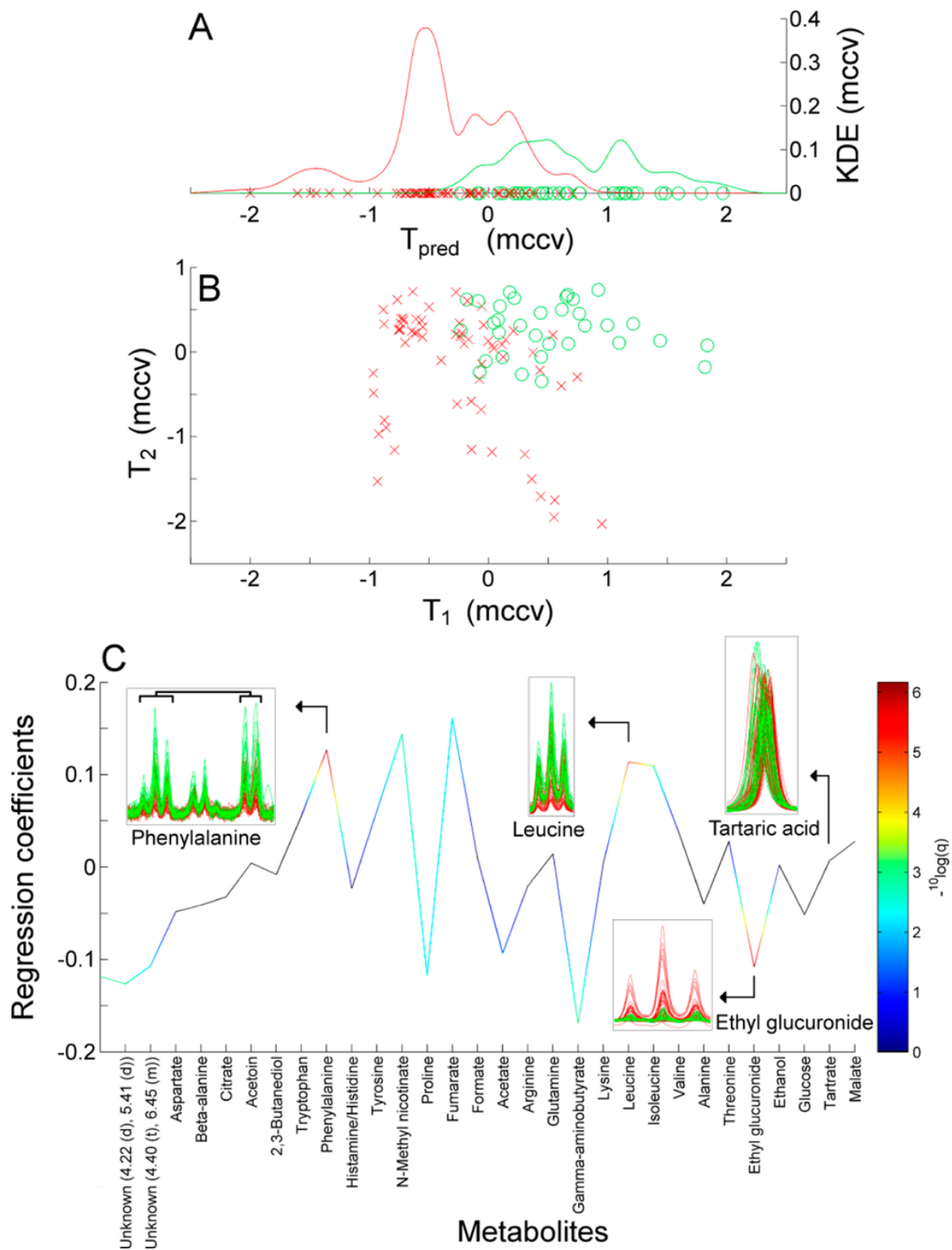
Figure 3. (A) Kernel density estimate (KDE) of the predicted PLSDA scores shows good separation between red (red crosses) and green (green circles) grapes. An R^2_Y of 0.65 shows the goodness of fit of the model, and a Q^2_Y of 0.52 shows good

capability for prediction. (B) PLS-DA scores plot. (C) PLS-DA loadings plot. The ^{1}H NMR peaks of 3 metabolites (phenylalanine, leucine, and ethyl glucuronide) significantly different between red and green grapes are shown in individual panels, as is tartaric acid, for which no significant difference in the concentration was found between red and green grapes.

Calculation of the ratio of tartaric acid consumed and tartaric acid excreted in the urine.

The mean concentration of tartaric acid in 50 g of red grapes (ten different varieties) and in 50 g of green grapes (two different varieties) was 0.84 ± 0.03 mMol for red and 0.85 ± 0.08 mMol for green grapes (Supporting Information Table 1). These results corroborated the PLS-DA results, which found no significant differences in the quantity of tartaric acid according to grape color or variety. The amount of tartaric acid excreted unchanged in the urine after consumption of 50 g of red grapes represented 19.2% of the amount consumed. Of this 5.5% was excreted in the first 3 h post consumption (CS2) and the remaining 13.7% was excreted 3–15 h post consumption (CS3).

Estimation of grape intake based on tartaric acid urinary excretion in a randomized highly controlled clinical trial.

Finally, we assessed the quantity of grapes consumed during the controlled clinical trial based on the models calculated from tartaric acid excretion in the urine samples collected on day 3. The CS2, CS3 and 24 h urine samples collected during the first and the second day of the control clinical trial were used as a test set of samples for this purpose. The amount of urinary tartaric acid excreted was interpolated from the calibration curve built with the CS2, CS3 and 24 h urine samples obtained on the third day of the study in order to estimate the amount of grapes consumed by the volunteers on the first 2 days in each intervention.

The Bland–Altman plots of the actual and estimated grape intake in CS2, CS3, and 24 h samples are shown in Figure 5A, B, and C, respectively. These plots highlight that the predictions of mean grape intakes in grams were close to the known intakes of 50 g (50.9 g), 100 g (91.4 g), and 150 g (163.0 g) of grapes and also verify that tartaric acid was not present in diet 1, where grapes were not consumed. We found that the accuracy of the predicted intake was inversely proportional to the amount of grapes consumed for all collection periods (CS2, CS3, and 24 h). In general, the most accurate prediction of intake was found in the 24 h urine sample (Figure 5),

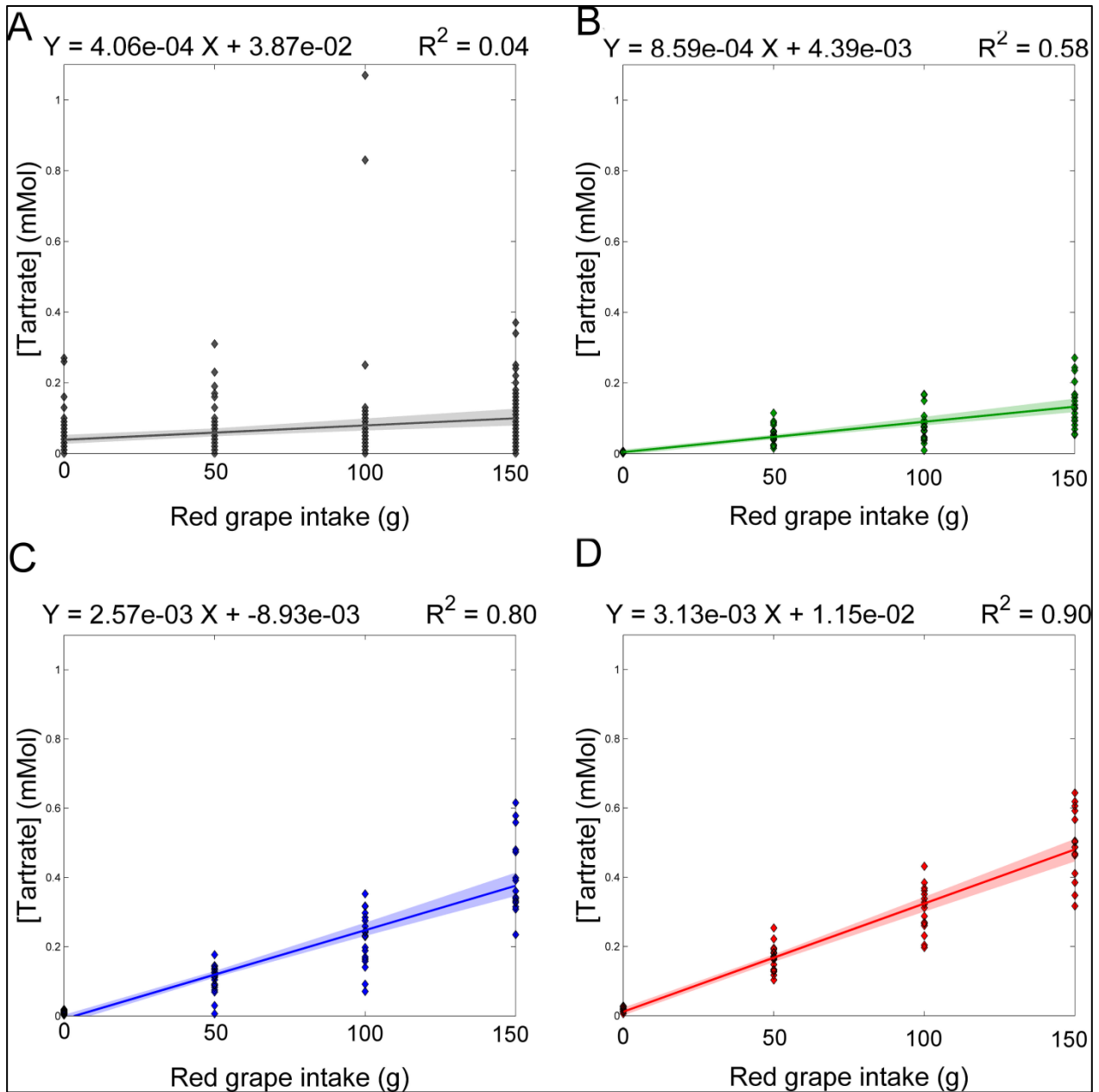indicating that collection of 24 h samples is the most appropriate strategy for estimation of grape intake.

Figure 4. Calibration curves of tartaric acid (mMol) excreted in urine after the intake of 0, 50, 100, and 150 g of grapes in (A) the first spot urine samples 2 h after grape intake, (B) the cumulative urine sample from after lunch to before dinner (CS2), (C) the cumulative urine sample from after dinner to the next day fasting urine sample (CS3), and (D) 24 h urine samples. The shaded area represents the 95% confidence interval obtained using bootstrap resampling.
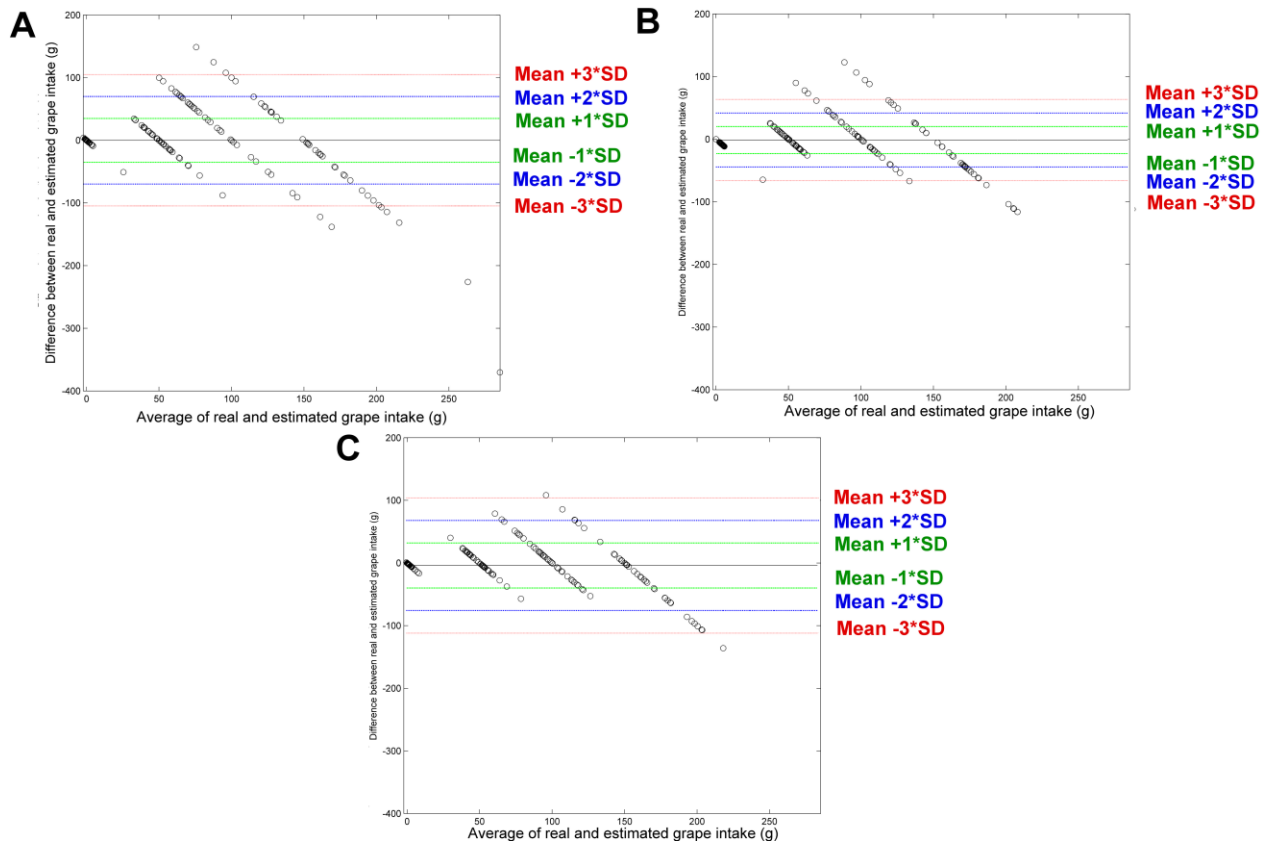
Figure 5. Bland–Altman plots of the actual and estimated grape intake in (A) CS2 samples, (B) CS3 samples, and (C) 24 h urine samples, showing that the predictions of amount of grape intake in grams are close to the real intake values of 50, 100, and 150 g of grapes. The green line indicates mean ±1 S.D.; blue, ±2 SD; and red, ±3 SD.

Development of a quantifiable biomarker for grape consumption.

There is a need to develop quantifiable dietary biomarkers for a range of food and nutrients to identify healthy and/or unhealthy eating patterns, since examples of validated quantifiable nutritional biomarkers are rare. Development of measurable biomarkers of intake of specific foods in body biofluids presents a complex challenge and is a multistage procedure involving: (i) discovery and identification of chemical(s) reflecting exposure to specific dietary components, (ii) assignment of candidate biomarkers to endogenous or exogenous origin, (iii) validation of candidate biomarker in an independent cohort, (iv) evaluation of the most appropriate biological sample for quantification of the biomarker and (v) investigation of the specificity and sensitivity of the proposed nutritional biomarker with respect to estimating exposure to specific food or food groups in an epidemiologic context.

In the current study, we developed a quantifiable biomarker for grape intake. ^{1}H NMR spectra of urine specimens from a pilot study showed tartaric acid to be a candidate biomarker reflecting exposure to incremental grape consumption. The follow-up kinetic study showed that the majority of tartaric acid was excreted between the first 4 h and 8 h postintervention in all participants, with almost complete excretion within 24 h postconsumption (Figure 2A). The dietary origin of tartaric acid as natural organic acid in grapes at high concentrations is well-known. It is also present in lower amounts in bananas, cranberries and tamarinds, but rare in most other common plants.[32,33] Traces of tartaric acid can also be found in processed foods as an acidifying agent.[34] PCA analysis of red grapes showed clustering according to variety and country (Figure 2B), which has been demonstrated due to soil, climatic region, and cultivar practices[35] and is reflected in the variation of sugars, amino acids and organic acids. However, differences in chemical composition reflected in PCA and PLS-DA models were unrelated to tartaric acid excretion. Tartaric acid was found in all varieties of grapes tested, and the concentration was similar between varieties (Supporting Information Table 1). Thus, it is feasible to suggest that tartaric acid serves as a robust, NMR-quantifiable biomarker of grape intake, independent of origin and grape variety.

Evaluation of optimal sampling strategy.
We evaluated the most appropriate sampling strategy for detection and quantification of dietary biomarkers based on 24 h total urinary biomarker excretion vs cumulative timed and spot urine collections. The evaluation was conducted by monitoring 24 h urinary excretion of tartaric acid in volunteers attending a highly controlled dietary intervention study. Findings unequivocally indicated that the cumulative 24 h collection performed best in terms of the ability to accurately estimate the quantity of grapes consumed and to detect and quantify tartaric acid. Evaluation of spot urine samples, as a more practical alternative in terms of study cost and participant compliance, indicated that it was not possible to obtain accurate estimates of grape consumption from spot urines obtained either at 7 am following an overnight fast or 2 h after grape intake. Although there was evidence of urinary excretion 2 h after grape intake in the spot urine samples, there was no strong association between dose and amount of tartaric acid excreted. This can most likely be explained by inter- and intra-individual variability in human metabolism. Tartaric acid is an exogenous compound, the majority of which is not absorbed. Our results concur with previous human metabolism studies suggesting that 15–20% of the dietary tartaric acid is excreted in the urine unchanged.[36,37] Tartrate either undergoes bacterial digestion in the large intestine, by at least 23

varieties of bacteria, or it is excreted in the stool in the form of an insoluble salt such as calcium tartrate,[36] which can affect the ratio of tartaric acid excreted in feces vs urine and explain interindividual differences in excretion.

However, cumulative urine samples CS2 and CS3 showed a linear relationship between grape intake and tartaric acid urinary excretion. These results emphasize the importance of understanding the kinetics of biomarker excretion in order to select the best time and sampling strategy, particularly if spot urine samples are to be used to detect specific dietary biomarkers. The optimal sampling time should correspond to the peak excretion time, which will be specific for each dietary biomarker. For example, proline betaine is predominantly detected between 2 and 6 h after citrus intake, [10] while the optimal window for tartaric acid was between 4 and 8 h after grape intake. With this in mind, nutritional epidemiological studies applying global metabolic profiling strategies should collect the first spot urine sample in a 3–6 h time window after the meal or food challenge. It is crucial that urine is collected at the same time for each study participant to avoid introduction of unnecessary variation. Consideration should be given as to whether the biomarker is unique to a given food: for example tartaric acid is itself added as a preservative to some processed foods and beverages. Moreover, it is important to note that sample preservatives can react with biomarkers an example being the complex tartaric acid forms with boric acid [38] altering the intensity and shape of the peak. As expected, 24 h urine samples remain the best option for total quantification of tartaric acid as a dietary biomarker since total excretion occurs in this time period regardless of interindividual differences.

Limitations in the proposed strategy for the quantification of dietary biomarkers.
The major confounder for assigning individual chemicals as biomarkers of specific foods or nutrients is the fact that these candidate biomarkers are seldom unique to a specific food. Urinary tartaric acid, in addition to being a major component of grapes, is also found in high concentrations in grape juice and wine, while lower concentrations can be found in other fruits or as additives. We found that 19.2% of all tartaric acid present in 50 g of grapes is excreted in the urine unchanged, therefore over 80% of tartaric acid is metabolized to other compounds. However, to address the lack of specificity of biomarkers, one solution to for instance differentiate grape and wine-derived tartaric acid is to utilize biomarker patterns to ascertain dietary origin. Biomarkers of wine intake identified from interventional studies include gallic acid, 4-o-methyl gallic acid, [39] caffeic acid and catechins. Of these, resveratrol is the only wine intake biomarker validated in clinical and epidemiological studies. [40,41] Therefore, resveratrol could be used to estimate

the amount of tartaric acid expected to derive directly from wine, and will allow the differentiation of urinary tartaric acid excreted as a consequence of grape intake or wine intake or both. We found the amount of tartaric acid to be comparable between red and white grapes, therefore it is a stable biomarker for grapes. However, it has been shown that resveratrol is found in higher concentrations in red and white grapes and has been validated as biomarker of red wine. [34] In the case of intake of both red grapes and red wine, assessing the ratio between resveratrol and tartaric acid excreted in urine following wine intake alone is needed in order to determine the amount tartaric acid that comes from wine and grapes. Ethanol and ethylglucuronide can also be used to assess consumption of alcoholic beverages in general to supplement using resveratrol alone.

Despite the fact that many studies have focused on the identification of wine [40,41] or grape juice[42] biomarkers, we are not aware of any studies reporting (quantitative) biomarkers of grape intake to date. The present work applies a new analytical pipeline to assess tartaric acid as a quantitative biomarker of grape intake that could be used in clinical and epidemiological studies in order to assess accurate grape intake. Grapes are rich in polyphenols, which contain antioxidants, conferring health benefits such as reduced risk of certain cancers. Dietary health policies are based on traditional dietary assessment methods, which are prone to misreporting. We propose a strategy using NMR as an alternative to current dietary reporting methods and demonstrate its applicability in a controlled nutritional trial. However, this strategy can be extendable to other analytical techniques in order to quantify other types of compounds.

In summary, we developed an analytical pipeline employing multivariate statistics for identifying a candidate food intake biomarker, followed by calculation of urinary concentration from ^{1}H NMR signal intensities and estimating grape intake using calibration curves. We describe a new dietary assessment tool that can be used to confirm intake of specific foods. We exemplified this approach using tartaric acid as a quantifiable biomarker of grape intake. Although this strategy was developed in the context of a highly controlled dietary intervention study, we provide proof-of-principle that coupling self-recorded dietary intake information with quantified nutritional biomarkers may be used to achieve a more objective measure of dietary exposure. Joining efforts from the nutritional and epidemiological researches identifying as many quantitative dietary biomarkers as possible will allow us in the future to accurately assess dietary intake both individually and in a population framework. Further studies are required to validate the application of this strategy to assess accurate dietary intake in free-living people.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jafc.5b05878. Supplementary Figure 1: Correlation of 31 identified compounds in ^{1}H NMR global profiles of red and green grapes visualized in a heat map with hierarchical clustering applied. Supplementary Table 1: Mean values of quantified tartaric acid in different varieties of red and green grapes from different countries. (PDF)

Corresponding Author

*Phone: +44(0)2075943225. E-mail: g.frost@imperial.ac.uk.

Author Contributions

IGP, JMP, EH, and GF designed research; IGP and ESC conducted the clinical trial; IGP conducted research; JMP and IGP analyzed data and performed statistical analyses; IGP, JMP, JKN, JCM, JD, EH, and GF wrote the paper. GF had final responsibility for final content. All authors read and approved the final manuscript.

Notes

The authors declare the following competing financial interest(s): Gary Frost has personal links with the food industry through Unilever, Nestl and Malaysian Palm Oil Board. No other authors declare a conflict of interest.

ABREVIATIONS USED

BMI, body mass index; CI, confidence interval; CS, cumulative (urine) sample; ^{1}H, proton; HCA, hierarchical cluster analysis; KDE, kernel density estimation; MCCV, Monte Carlo cross validation; NMR, nuclear magnetic resonance; PCA, principal component analysis; PLS-DA, partial least squares discriminant analysis; QC, quality control; Q^2_Y, goodness of prediction; R^2_Y, goodness of fit; r^2, squared correlation coefficient; RD, relaxation delay; RF, radio-frequency; t_m, mixing time; TSP, trimethylsilyl-[2,2,3,3-2H4]-propionate

REFERENCES

(1) Wang, X.; Ouyang, Y.; Liu, J.; Zhu, M.; Zhao, G.; Bao, W.; Hu, F. B. Fruit and vegetable consumption and mortality from all causes, cardiovascular disease, and cancer: systematic review and dose- response meta-analysis of prospective cohort studies. BMJ. 2014, 349, g4490.

(2) Oyebode, O.; Gordon-Dseagu, V.; Walker, A.; Mindell, J. S. Fruit and vegetable consumption and all-cause, cancer and CVD mortality: analysis of Health Survey for England data. Journal of epidemiology and community health 2014, 68, 856–62.

(3) Olendzki, B. C.; Ma, Y.; Hebert, J. R.; Pagoto, S. L.; Merriam, P. A.; Rosal, M. C.; Ockene, I. S. Underreporting of energy intake and associated factors in a Latino population at risk of developing type 2 diabetes. J. Am. Diet. Assoc. 2008, 108, 1003–8.

(4) Ismail, N. A.; Posma, J. M.; Frost, G.; Holmes, E.; Garcia-Perez, I. The role of metabonomics as a tool for augmenting nutritional information in epidemiological studies. Electrophoresis 2013, 34, 2776–86.

(5) Kelly, M. T.; Rennie, K. L.; Wallace, J. M.; Robson, P. J.; Welch, R. W.; Hannon-Fletcher, M. P.; Livingstone, M. B. Associations between the portion sizes of food groups consumed and measures of adiposity in the British National Diet and Nutrition Survey. Br. J. Nutr. 2009, 101, 1413–20.

(6) Shaneshin, M.; Rashidkhani, B.; Rabiei, S. Accuracy of energy intake reporting: comparison of energy intake and resting metabolic rate and their relation to anthropometric and sociodemographic factors among Iranian women. Archives of Iranian medicine 2012, 15, 681–7.

(7) Belton, P. S.; Delgadillo, I.; Holmes, E.; Nicholls, A.; Nicholson, J. K.; Spraul, M. Use of High-Field ^{1}H NMR Spectroscopy for the Analysis of Liquid Foods. J. Agric. Food Chem. 1996, 44, 1483−1487.

(8) Lloyd, A. J.; Fave, G.; Beckmann, M.; Lin, W. C.; Tailliart, K.; Xie, L.; Mathers, J. C.; Draper, J. Use of mass spectrometry fingerprinting to identify urinary metabolites after consumption of specific foods. Am. J. Clin. Nutr. 2011, 94, 981−991.

(9) Marklund, M.; Landberg, R.; Aman, P.; Kamal-Eldin, A. Determination of alkylresorcinol metabolites in human urine by gas chromatography-mass spectrometry. J. Chromatogr. B: Anal. Technol. Biomed. Life Sci. 2010, 878, 888−94.

(10) Heinzmann, S. S.; Brown, I. J.; Chan, Q.; Bictash, M.; Dumas, M. E.; Kochhar, S.; Stamler, J.; Holmes, E.; Elliott, P.; Nicholson, J. K. Metabolic profiling strategy for discovery of nutritional biomarkers: proline betaine as a marker of citrus consumption. Am. J. Clin. Nutr. 2010, 92, 436−443.

(11) Edmands, W. M. B.; Beckonert, O. P.; Stella, C.; Campbell, A.; Lake, B. G.; Lindon, J. C.; Holmes, E.; Gooderham, N. J. Identification of Human Urinary Biomarkers of Cruciferous Vegetable Consumption by Metabonomic Profiling. J. Proteome Res. 2011, 10, 4513−4521.

(12) Bingham, S. A. Biomarkers in nutritional epidemiology. Public health nutrition 2002, 5, 821−7.

(13) Lloyd, A. J.; Beckmann, M.; Fave, G.; Mathers, J. C.; Draper, J. Proline betaine and its biotransformation products in fasting urine samples are potential biomarkers of habitual citrus fruit consumption. Br. J. Nutr. 2011, 106, 812−24.

(14) Wood, K. V.; Stringham, K. J.; Smith, D. L.; Volenec, J. J.; Hendershot, K. L.; Jackson, K. A.; Rich, P. J.; Yang, W. J.; Rhodes, D. Betaines of alfalfa: characterization by fast atom bombardment and desorption chemical ionization mass spectrometry. Plant Physiol. 1991, 96, 892−7.

(15) Hertog, M. G.; Kromhout, D.; Aravanis, C.; Blackburn, H.; Buzina, R.; Fidanza, F.; Giampaoli, S.; Jansen, A.; Menotti, A.; Nedeljkovic, S.; et al. Flavonoid intake and long-term risk of coronary heart disease and cancer in the seven countries study. Arch. Intern. Med. 1995, 155, 381−6.

(16) Shirataki, Y.; Kawase, M.; Saito, S.; Kurihara, T.; Tanaka, W.; Satoh, K.; Sakagami, H.; Motohashi, N. Selective cytotoxic activity of grape peel and seed extracts against oral tumor cell lines. Anticancer Res. 2000, 20, 423−6.

(17) Do, M. H.; Lee, S. S.; Kim, J. Y.; Jung, P. J.; Lee, M. H. Fruits, vegetables, soy foods and breast cancer in pre- and postmenopausal Korean women: a case-control study. Int. J. Vitam. Nutr. Res. 2007, 77, 130−41.

(18) Kaliora, A. C.; Kountouri, A. M.; Karathanos, V. T.; Koumbi, L.; Papadopoulos, N. G.; Andrikopoulos, N. K. Effect of Greek raisins (Vitis vinifera L.) from different origins on gastric cancer cell growth. Nutr. Cancer 2008, 60, 792−9.

(19) Dohadwala, M. M.; Vita, J. A. Grapes and cardiovascular disease. J. Nutr. 2009, 139, 1788S−93S.

(20) Dona, A. C.; Jimenez, B.; Schafer, H.; Humpfer, E.; Spraul, M.; Lewis, M. R.; Pearce, J. T.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping. Anal. Chem. 2014, 86, 9887−94.

(21) Yang, L.; Ru, G. Y.; Tang, H. R.; Liu, C. Y. Human Urinary Metabolite Quantification in Nuclear Magnetic Resonance-based Metabonomics Using Electronic Internal Reference. Chinese J. Anal Chem. 2010, 38, 789−794.

(22) Hotelling, H. Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology 1933, 24, 417−441.

(23) De Jong, S. Simpls - an Alternative Approach to Partial Least-
Squares Regression. Chemom. Intell. Lab. Syst. 1993, 18, 251−263.

(24) Xu, Q. S.; Liang, Y. Z. Monte Carlo cross validation. Chemom. Intell. Lab. Syst. 2001, 56, 1−11.

(25) Efron, B. Nonparametric Estimates of Standard Error − the Jackknife, the Bootstrap and Other Methods. Biometrika 1981, 68, 589−599.

(26) Storey, J. D.; Tibshirani, R. Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. U. S. A. 2003, 100, 9440−5.

(27) Guimera, R.; Nunes Amaral, L. A. Functional cartography of complex metabolic networks. Nature 2005, 433, 895−900.

(28) Bland, J. M.; Altman, D. G. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986, 1, 307−10.

(29) Cloarec, O.; Dumas, M. E.; Craig, A.; Barton, R. H.; Trygg, J.; Hudson, J.; Blancher, C.; Gauguier, D.; Lindon, J. C.; Holmes, E.; Nicholson, J. Statistical total correlation spectroscopy: An exploratory approach for latent biomarker identification from metabolic H-1 NMR data sets. Anal. Chem. 2005, 77, 1282−1289.

(30) Posma, J. M.; Garcia-Perez, I.; De Iorio, M.; Lindon, J. C.; Elliott, P.; Holmes, E.; Ebbels, T. M.; Nicholson, J. K. Subset optimization by reference matching (STORM): an optimized statistical approach for recovery of metabolic biomarker structural information from ^{1}H NMR spectra of biofluids. Anal. Chem. 2012, 84, 10694−701.

(31) Hao, J.; Astle, W.; De Iorio, M.; Ebbels, T. M. BATMAN−an R
package for the automated quantification of metabolites from nuclear magnetic resonance spectra using a Bayesian model. Bioinformatics 2012, 28, 2088−90.

(32) Ribeŕeau-Gayon, P.; Glories, Y.; Maujean, A.; Dubourdieu, D. Organic Acids in Wine. In Handbook of Enology, 2nd ed.; John Wiley & Sons, Ltd: 2006; Vol. 2, pp 1–49.

(33) Velioglu, Y. S. Food acids: organic acids, volatile organic acids, and phenolic acids. In Advances in Food Biochemistry; Yildiz, F., Ed.; CRC Press: Boca Raton, 2009; p 522.

(34) Regueiro, J.; Vallverdu-Queralt, A.; Simal-Gandara, J.; Estruch, R.; Lamuela-Raventos, R. M. Urinary tartaric acid as a potential biomarker for the dietary assessment of moderate wine consumption: a randomised controlled trial. Br. J. Nutr. 2014, 111, 1680–1685.

(35) Son, H. S.; Hwang, G. S.; Kim, K. M.; Ahn, H. J.; Park, W. M.; Van Den Berg, F.; Hong, Y. S.; Lee, C. H. Metabolomic Studies on Geographical Grapes and Their Wines Using H-1 NMR Analysis Coupled with Multivariate Statistics. J. Agric. Food Chem. 2009, 57, 1481–1490.

(36) Finkle, P. The fate of tartaric acid in the human body. J. Biol. Chem. 1933, 100, 349–355.

(37) Gry, J.; Larsen, J. C. Metabolism of L(+)-and D(−)-tartaric acids in different animal species. Arch. Toxicol., Suppl. 1978, 1, 351–3.

(38) Rebstockova, M.; Bartusek, M. Reactions of Boric-Acid with Oxalic, Glycolic and Tartaric-Acids. Collect. Czech. Chem. Commun. 1977, 42, 627–636.

(39) Mennen, L. I.; Sapinho, D.; Ito, H.; Bertrais, S.; Galan, P.; Hercberg, S.; Scalbert, A. Urinary flavonoids and phenolic acids as biomarkers of intake for polyphenol-rich foods. Br. J. Nutr. 2006, 96, 191–198.

(40) Zamora-Ros, R.; Urpi-Sarda, M.; Lamuela-Raventos, R. M.; Estruch, R.; Vazquez-Agell, M.; Serrano-Martinez, M.; Jaeger, W.; Andres-Lacueva, C. Diagnostic performance of urinary resveratrol metabolites as a biomarker of moderate wine consumption. Clin. Chem. 2006, 52, 1373–1380.

(41) Zamora-Ros, R.; Urpi-Sarda, M.; Lamuela-Raventos, R. M.; Estruch, R.; Martinez-Gonzalez, M. A.; Bullo, M.; Aros, F.; Cherubini, A.; Andres-Lacueva, C. Resveratrol metabolites in urine as a biomarker of wine intake in free-living subjects: The PREDIMED Study. Free Radical Biol. Med. 2009, 46, 1562–1566.

(42) Lord, R. S.; Burdette, C. K.; Bralley, J. A. Significance of urinary tartaric acid. Clin. Chem. 2005, 51, 672–3.

Supplementary Figure 1: Correlation of 31 identified compounds in ^{1}H-NMR global profiles of red and green grapes visualized in a heat map with hierarchical clustering applied.
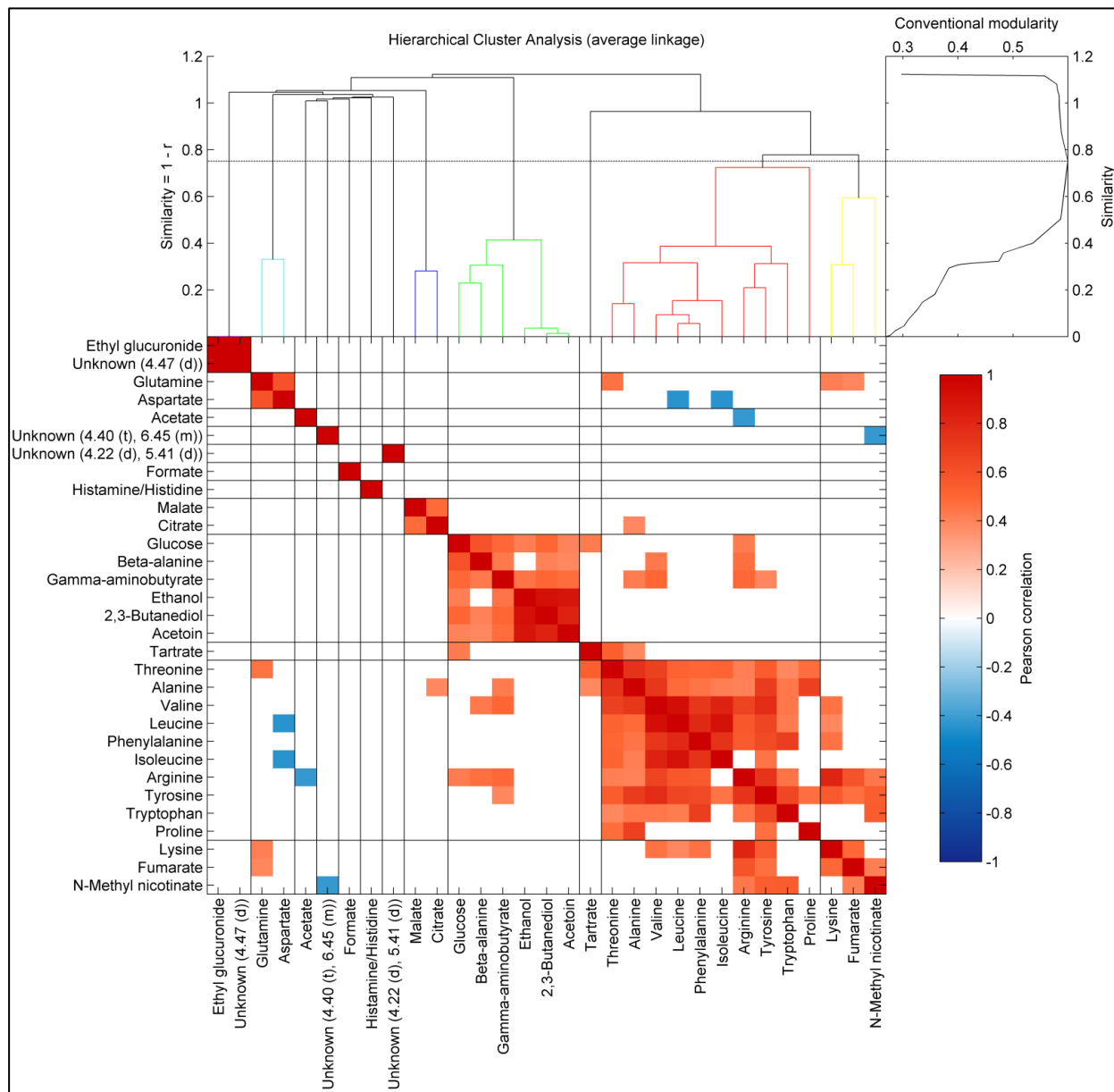
Supplementary Table 1: Mean values of quantified tartaric acid in different varieties of red and green grapes from different countries

red grape variety	mean tartaric acid (mMol)s	SD
Sharad (India)	0.88	0.03
Red Globe (Peru)	0.59	0.09
Ralli (Chile)	0.81	0.04
Pink Muscat (Chile)	0.81	0.04
Flame (Chile) (1)	0.93	0.05
Flame (Chile) (2)	0.85	0.05
Crimson (Chile)	1.11	0.09
Black Princess (Chile)	0.98	0.00
Jack's Salute (South Africa)	0.71	0.08
Magenta (South Africa)	0.79	0.03
Crimson (South Africa)	0.81	0.04
total	0.84	0.03

green grape variety	mean tartaric acid (mMol)	stdv
Thompson (India) (1)	1.11	0.14
Thompson (India) (2)	0.71	0.04
Thompson (Namibia)	0.82	0.04
Thompson (Chile)	0.76	0.03
Sugraone (Chile	0.91	0.09
Thompson (South Africa)	0.80	0.11
total	0.85	0.08