Determination of the Molecular Transformations and Pathways that Occur During the Winemaking Process

Mark Dreyer[1], Paul Tarr[2], Mona Shahgholi[2], Michael Athanas[1], Tonya Second[1]
[1]Thermo Fisher Scientific, San Jose, Ca; [2]California Institute of Technology, Pasadena, Ca;

Overview

Purpose: To determine the molecular transformations and pathways that occur during the winemaking process.

Methods: Samples were analyze by injection onto an LC-MS and using post acquisition label-free analysis software.

Results: The results of this study provide a catalogue of the components present in the samples and a statistical analysis reflecting the changes in the component levels observed across the samples. This data set may serve as a first approximation of how certain influences during the winemaking process from vineyard to vineyard are potentially reflected in the organoleptic properties of the final wine product. Efficient verification of known constituents as well as detection of unexpected components is an important area of research for the agricultural industry.

Introduction

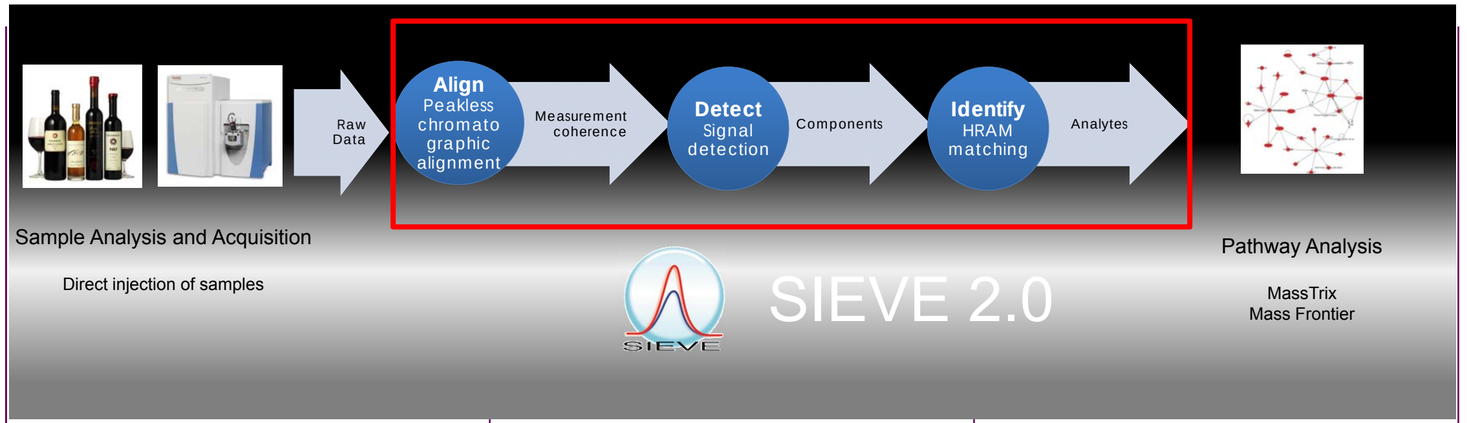
Wine is a very complex mixture and a rich source of beneficial anti-oxidants. Identification and quantitation of the components is challenging process. High Pressure Liquid Chromatography coupled to a Thermo Scientific Q Exactive mass spectrometer was used for analysis of several wines from the Ballard Canyon and Santa Ynez Valley in Santa Barbara county California. This enabled the simultaneous detection and relative quantitation of the wine's components. The chemical composition of wine defines its organoleptic properties. The final marketed product is a complex solution that is primarily the product of two metabolomes, *Vitis vinifera L.* and *Saccharomyces cerevisiae*. Furthermore, the *V.v.* genome and metabolomes is influenced by *terroir*— the characteristics that the geography, geology and climate of a certain vineyard bestow upon the final product during the growing season. Together with the yeast strain chosen for fermentation, and whether the fermented must is aged in oak, all these factors contribute a unique organoleptic footprint on the final wine product.

Elucidating these molecular transformations that occur during the winemaking process has the potential to greatly increase our understanding and control over the winemaking process. A complete chemical analysis will serve as a first step to map the metabolites back to the biochemical pathways at work in both *Vitis vinifera L.* and *Saccharomyces cerevisiae.*

Direct wine analysis was used to determine the molecular weight of the components and their variable content across the 5 samples in this study was observed. This work demonstrates the utility of analyzing complex mixtures without prior sample preparation by making use of the separation of HPLC combined with the High Resolution Accurate Mass (HRAM) of the Q Exactive quadrupole Orbitrap™ mass analyzer. The combination of these techniques provides a robust and confident means of profiling complex mixtures as well as successful identification of detected components.

Samples

5 *must* (freshly pressed fruit juice) samples were obtained from the Saarloos and Sons Winery in Los Olivos California. We have chosen to address the questions of terrior as a first past analysis of the data set we have generated. We have chooses to look at intra- and intervineyard terrior. Three distinct blocks of Syrah (all clone 877) within the Windmill Ranch vineyard located in Ballard Canyon that differ geographically by ~50 yards by their position on the face of canyon slope were compared for our intravineyard analysis. In our intervineyard analysis we compared the chemical composition of the same Grenache clone grown in the Windmill Ranch vineyard and grown in the El Camino Real vineyard located in Foxen Canyon. Approximately 1.5 miles geographically separates these two vineyards



Methods

High Performance Liquid Chromatography

The wine was stored at room temperature until analysis. Chromatography was performed using a Thermo Scientific Dionex Ultimate 3000 HPLC system, injecting 10 ul directly onto a Thermo Scientific Accucore column (50 x 2.1 mm, 2.3 um particles, equilibrated in 95% solvent A (0.1% aqueous solution of formic acid), 5% solvent B (methanol containing 0.1% formic acid). The compounds were eluted using flow rate 300 µL/min by linearly increasing solvent B concentration from 5% to final 95% over 15 min. The column was then washed with 95% solvent B (2 min) and re-equilibrated in 95% solvent A, 5% solvent B. The total run time, including column wash and equilibration, was 20 min.)

Time	%A	%B	Flow (ul/min)
0.0	98	2	300
0.5	98	2	300
15.0	2	98	300
18.0	2	98	300
18.1	98	2	300
20.0	98	2	300

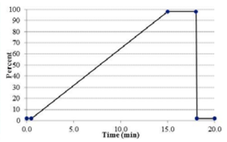


FIGURE 1. The gradient used for the direct wine analysis

Mass Spectrometry

Q Exactive benchtop Orbitrap mass spectrometer was operated in positive ion mode at 70,000 resolving power (defined as FWHM @ *m/z* 200) for full scan analysis (mass range 70 – 1000 amu).The measurements were done in duplicate with an internal standard added to each sample

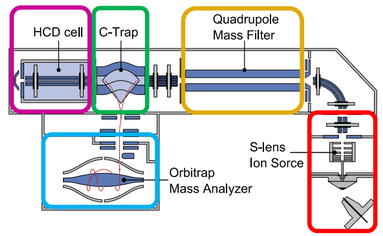


FIGURE 2. The Q Exactive benchtop Orbitrap LC-MS/MS highlighting the front-end ion optics which enhance sensitivity, the quadrupole mass filter for selectivity, the HCD cell used for Higher Energy Collisional Dissociation and the Orbitrap High Resolution Accurate Mass Analyzer.

SIEVE Workflow

Thermo Scientific SIEVE 2.0 software was used for comparative analysis. SIEVE™ software provides an easy-to-use automated solution for evaluating the large amounts of LC-MS data generated by the label-free sample analyses we used. It effectively locates compounds with statistically significant abundance differences between sample populations. SIEVE software was used to perform trend analyses to locate compounds that vary by geographical location or year within this sample set. SIEVE software is a statistically-rigorous data mining tool that has the power to compare hundreds of LC-MS data files at a time. The SIEVE software workflow first performs a chromatographic alignment using full scan spectra shape (no peaks).

Frames (a.k.a. features) are constructed based upon prominent peaks throughout the full data set. Results were then filtered using CV% < 10 for each sample type and at the same time requiring a minimum 2-fold change in peak height. It pre-filters complex data (associating adducts), greatly reducing the number of compounds that need to be evaluated and significantly decreasing time spent on identification. The resulting set of components were searched against a custom database of flavonoids and related compounds to identify potential compounds of interest which can be later confirmed by the analysis of pure standards. The results from SIEVE software were further subjected to clustering analysis and PCA analysis using R.

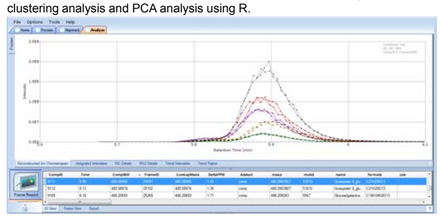


FIGURE 3 SIEVE 2.0 software display

Component Clustering

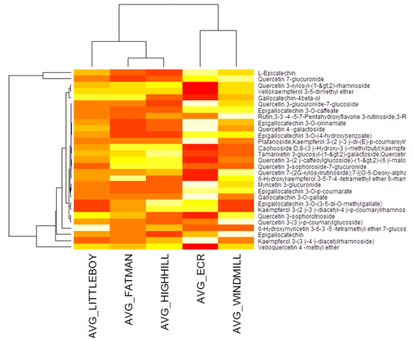


FIGURE 4. Hierarchical clustering shows that the components identified as L-Epicatechin and Quercetin 7-Glucuronide are the primary discriminators, while Velloquercetin-4-methyl ether is a secondary discriminator. Keep in mind that confirmation of the precise structure of the compounds needs to be done using standards.

PCA Analysis

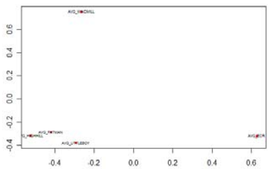


Figure 5. Principal component Analysis clearly shows that the Grenache ECR and the Grenache HighHill are different from the Fatman, Little Boy and WindMill as well as different from each other.

Conclusions

Despite the complexity of the wine samples, exhibiting numerous chromatographic peaks each of which contain multiple components, our differential analysis of the Grenache and Syrah unfiltered samples, contrasted wine from several different geographical areas. We observed at least 178 individual compounds that showed at least 2-fold higher content in the ECR than in the other samples tested. The compounds identified as L-Epicatechin and Quercetin-7-glucoronide (confirmation of the precise isomer is necessary) appear to be the most differentially present components that can classify the 5 wines characterized. The compounds identified as Epigallocatechin, Kaempferol-diacetylhamnoside and Velloquercetin 4-methyl ether (again confirmation is needed) appear to have a secondary effect on the classification.