

A Predictive Compound Database Approach to the Tentative Identification and Semiquantitation of Volatile-Phenol Glycosides in Smoke-Affected Grapes from Wildfires

Authors

Andrew Caffrey, Larry Lerno,
Arran Rumbaugh,
Raul Girardello,
Anita Oberholster, and
Susan E. Ebeler, University of
California, Davis, CA;

Jerry Zweigenbaum,
Agilent Technologies, Inc.,
Wilmington, DE 19898

Abstract

This Application Note outlines the workflow and processes used to survey for volatile-phenol glycosides in smoke-affected grapes and semiquantitate their abundances throughout winemaking. An Agilent 1290 Infinity II LC with an Agilent 6545 Accurate-Mass Q-TOF LC/MS was used to identify 31 volatile-phenol glycosides in grapes, and demonstrate the first evidence for trisaccharide volatile-phenol glycosides in grapes. Analysis began with an untargeted search against a user-generated personal compound database (PCD). Potential compounds were then screened using dual-jet electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI) sources. Electrospray ionization was chosen for final analyses, and MS/MS fragmentation data of database matches were acquired and evaluated by Molecular Structure Correlator (MSC) for tentative identification. Putatively identified compounds were semiquantitated throughout winemaking using the integration process in Agilent MassHunter Quantitative Analysis using the “all ions” approach.

Introduction

The complementary article for this analysis of volatile-phenol glycosides can be found in the American Journal of Enology and Viticulture.¹ Volatile-phenol glycosides consist of an aglycone volatile phenol, such as guaiacol, bound to a glycone group that contains at least one sugar molecule. These nonvolatile glycosides do not have an aroma, and are not readily detectable by the consumer. However, volatile-phenol glycosides are of high interest to the wine industry due to the unpleasant aromas they impart on wines when they are hydrolyzed, releasing the free volatile phenol. These aromas have been described as smoky, medicinal, and ash-like, along with several other aroma characteristics that can tarnish the quality of the wines.^{1,2}

Volatile phenols are introduced into *Vitis vinifera* berries through exposure to wildfire smoke.³ After the smoke exposure, the free volatiles are glycosylated in the berries, and they remain in that bound state until hydrolysis occurs when the grapes are crushed and fermented. As a result, the glycosides act as a reservoir for smoke-related off flavors during fermentation and storage.⁴

Conventionally, samples containing volatile-phenol glycosides have been studied using gas chromatographic (GC) methods. Although GC methods have been insightful, these methods lack information on the make-up and specific abundances of the intact volatile-phenol glycosides present in the grapes or wines. The availability of commercial standards for volatile-phenol glycosides is limited in nature. A recent study shows that ultrahigh performance liquid chromatography paired with electrospray ionization quadrupole time-of-flight tandem mass spectrometry (UHPLC Q-TOF MS/MS) is capable of

characterizing volatile-phenol glycosides found in smoke-affected grapes and wine using a targeted approach with synthesized standards.⁵ In the original report of the analysis outlined here, an untargeted approach was used for a broader characterization and semiquantitation of volatile-phenol glycosides in the absence of standards.¹

This Application Note describes an in-depth overview on the software used to expand the current knowledge base of volatile-phenol glycosides in smoke-affected grapes. This approach uses Agilent PCDL Manager, Agilent MSC, Agilent MassHunter Qualitative Analysis, and MassHunter Quantitative Analysis. The software was effective in using a large database to search for, analyze, and elucidate structures of unknown compounds for semiquantitation throughout the winemaking process.

Experimental

All grape and winemaking samples were prepared for analysis using solid-phase extraction methods found in the complementary report.¹ Extracts were analyzed on an Agilent 1290 Infinity II LC coupled to an Agilent 6545 Accurate Mass Q-TOF LC/MS. The UHPLC system was equipped with an Agilent 1290 Infinity binary pump (G7120A), a temperature-controlled Agilent autosampler (G7167B), an Agilent infinity isocratic pump (G1310B), and a temperature-controlled column compartment (G1316C). Table 1 presents the UHPLC parameters.

Negative polarity ESI and APCI were considered for this analysis. Ultimately, ESI was chosen for identification and semiquantitation of samples. Table 2 shows the electrospray source, APCI source, and Q-TOF parameters. The analysis used two separate time segments. For the first 15 minutes of analysis, total ion spectra for collision

energies of 0 and 20 eV were collected during analysis for an "All Ions MS/MS" method (henceforth referred to as "all ions"). The relative quantitation was performed on the 0 eV collision energy channel. During the remaining time of the analysis, the LC stream was diverted to waste.

A reference mass solution was used during analysis for continuous mass calibration. The reference masses chosen were deprotonated purine (m/z 119.0362) and the acetate adduct of hexakis (¹H,¹H,³H-tetrafluoropropoxy) phosphazine (m/z 980.016375) to coincide with the mass range of m/z 100 to 1,000.

Salicin was introduced as an internal standard for semiquantitation. Each sample was spiked with salicin before solid-phase extraction to account for variations from the solid phase extraction process.

Table 1. Agilent 1290 Infinity II Binary LC parameters.

Instrument	Agilent 1290 Infinity II Binary LC
Mobile Phases	A) 0.1% Acetic acid in water B) 0.1% Acetic acid in acetonitrile
Gradient	Linear Time (minutes) %B 0 5 10 35 18 95 20 5 25 5
Flow Rate	0.420 mL/min
Column	Agilent Poroshell 120 Phenyl Hexyl UHPLC column 2.1 × 150 mm, 2.7 micron (p/n 693775-912) with an Agilent 2.1 × 5 mm 2.7 micron guard column (p/n 821725-914)
Temperature	40 °C
Injection Volume	10 µL

Table 2. Agilent 6545 Accurate-Mass Q-TOF LC/MS parameters.

Instrument	Agilent 6545 Accurate-Mass Q-TOF LC/MS
MS Acquisition Rate	3.0 spectra/s
MS Mass Range	100 to 1,000 <i>m/z</i>
MS/MS Mass Range	50 to 750 <i>m/z</i>
MS1 Acquisition Rate	3.00 spectra/s
MS2 Acquisition Rate	4.00 spectra/s
Precursor per MS1 Spectrum	3
Active Exclusion	On
"All Ions" MS Collision Energy(s)	0 and 20 eV
MS/MS Collision Energy(s)	0, 20, and 30 eV
Negative Electrospray with Dual Agilent Jet Stream Technology	
Drying Gas Temperature	150 °C
Drying Gas Flow Rate	10 L/min
Sheath Gas Temperature	350 °C
Sheath Gas Flow Rate	11 L/min
Nebulizer Gas	35 psig
Skimmer Voltage	65 V
Octopole RF	750 V
Fragmentor	120 V
Capillary	3.5 kV
Negative Atmospheric Pressure Chemical Ionization	
Gas Temperature	350 °C
Gas Flow Rate	8 mL/min
Vaporizer Temperature	350 °C
Nebulizer Gas	35 psig
Capillary	4,000 V
Corona	5 μA
Fragmentor	75 V

Data analysis

A more in-depth description of data analysis software can be found in the cited Agilent Technologies Application Note.⁶ Tentative identification of volatile-phenol glycosides started with a survey of smoke taint literature to identify possible compounds. Each hypothetical compound followed the basic structures outlined in Figure 1. Compounds were drawn within MarvinSketch, and imported into PCDL Manager as mol files. PCDL Manager automatically created a PCD based on the molecular structures given.

The PCD contained the compound name, structure, and exact mass for over 600 potential compounds. The PCD was imported into MassHunter Qualitative Analysis software where it was applied to an LC/MS chromatogram of a representative grape sample using the Find-by-Formula algorithm. Over 100 hits were matched to a compound from the database and scored based on the criteria of mass accuracy, isotope abundance, and isotope spacing.

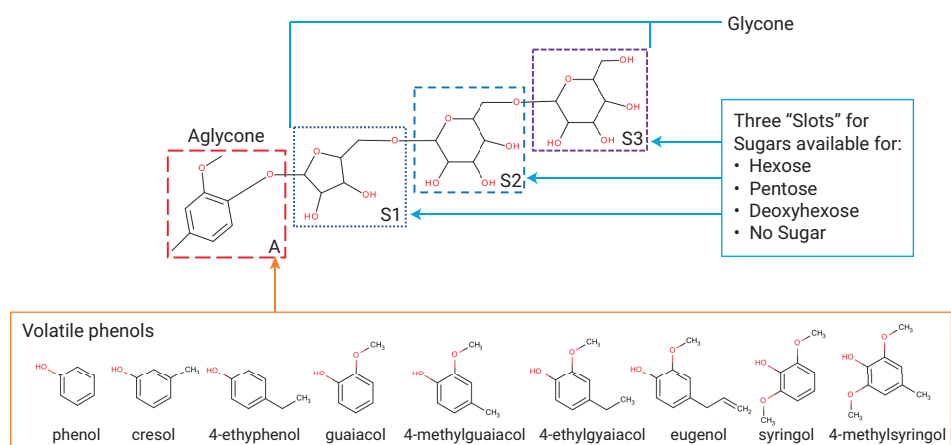


Figure 1. Shown is the general drawing scheme for the compounds listed in the PCD. All aglycones used are shown and glycone sugars are reported in Caffrey *et al.* (2019).¹ Using this scheme, over 600 compounds were included for the untargeted search.

Resulting hits with a score over 80 were exported from qualitative analysis to the data acquisition software as an inclusion list for Auto MS/MS. The inclusion list was edited to ensure that there were no redundant repetitions of compounds. Auto MS/MS was performed to generate spectra with fragmentation for structural elucidation. MS/MS results were converted to CEF files in qualitative analysis after using the Find-by-Auto MS/MS algorithm. MassHunter MSC compared all of the MS/MS spectra to the structures in the original database. This resulted in potential identifications with visualizations of the 30 most abundant fragments in each MS/MS spectrum. To ensure that fragmentation

patterns were consistent with other reports in literature, identifications were manually interpreted within MSC. MSC offered a quick and user-friendly identification of fragments from each compound by giving abundances, mass errors, and potential sources of the fragmentation. Figure 2 shows the results of this analysis.

After compounds were putatively identified in grapes, a personal compound database library (PCDL) was created in the PCDL Manager software. This library consisted of retention time, exact mass, molecular structure, and fragmentation data. The PCDL was used by MassHunter Quantitative Analysis for the auto-integration of grape and

winemaking samples. An integration method was generated to auto-identify noise regions for each compound to calculate signal-to-noise ratio (S/N), identify qualifier ions for each compound, and automatically normalize peak areas to the internal standard salicin. The method was applied to the entire batch of data, and outliers were flagged based on deviations in retention time or a S/N less than three. Data in the form of normalized peak areas were exported into excel for data preprocessing before analysis in R studio for ANOVA and *post hoc* testing. Figure 3 shows the overall outline of this process.

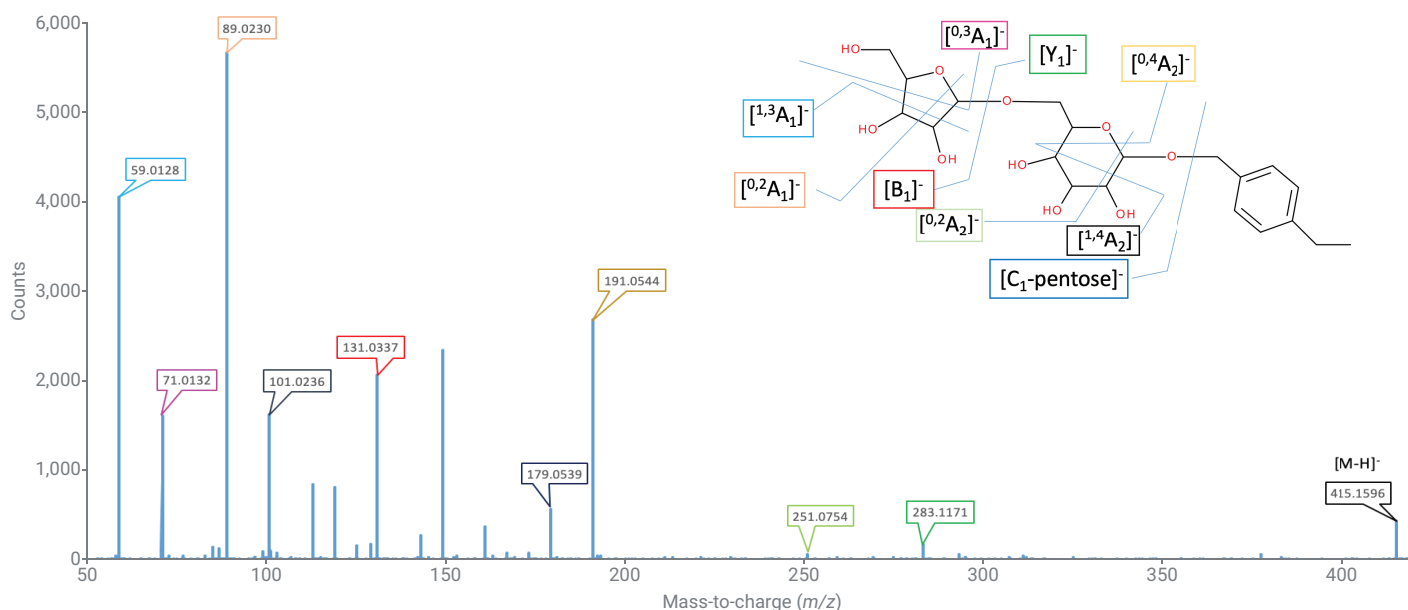


Figure 2. MSC was used to easily identify fragment ions in the MS/MS data and show their location in the compound pulled from the PCD. Displayed is a visual adaptation of the fragmentation data for pentose-hexose-4-ethylphenol obtained with the assistance of MSC.

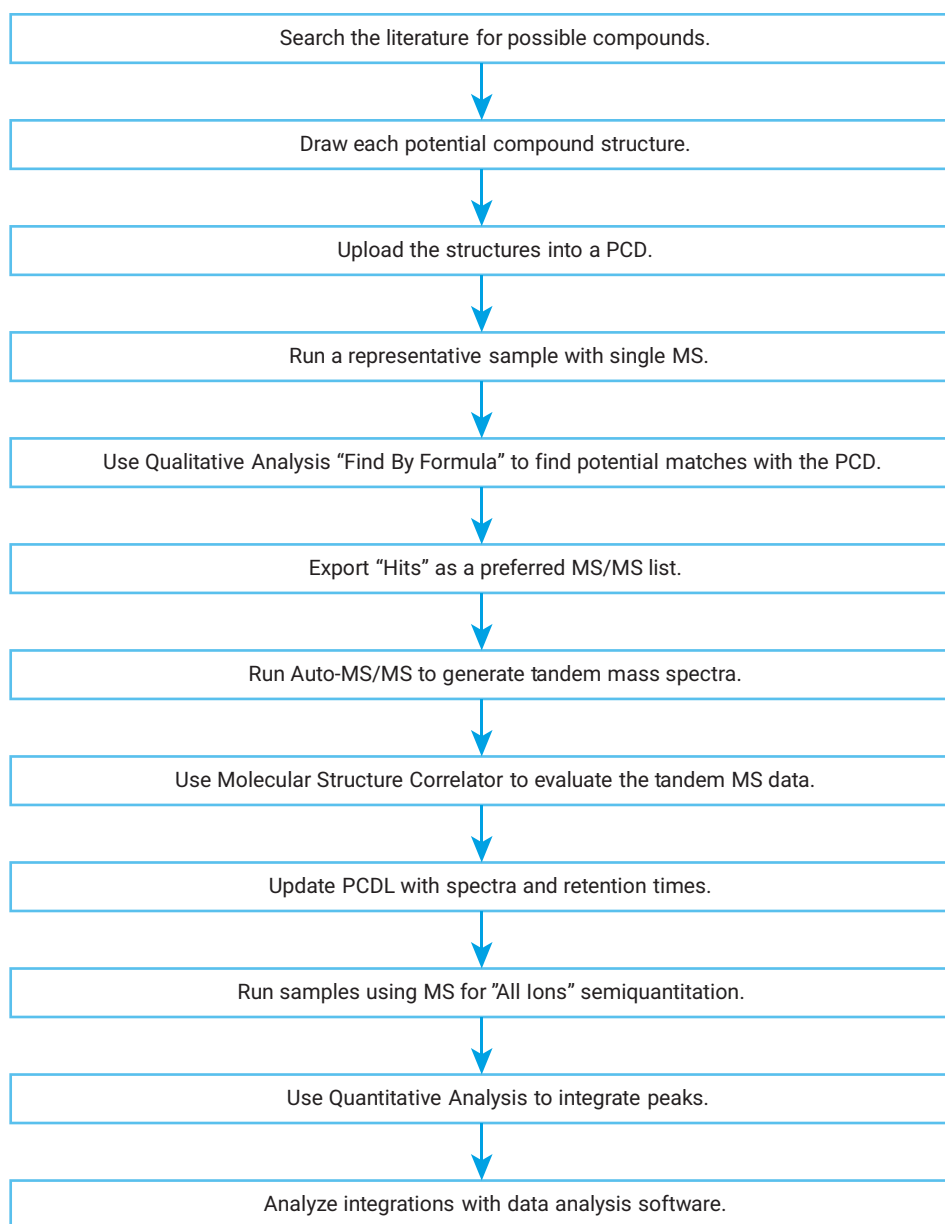


Figure 3. The overall approach for this Application Note, starting with an untargeted search for compounds and ending with data analysis in programs such as R.

Results and discussion

Within the study, 31 volatile-phenol glycosides were tentatively identified. A complete list of the compounds may be found in the complementary article.¹ Trisaccharide volatile-phenol glycosides were reported for the first time in grapes. Conventionally, glycones of more than two sugars are challenging to study due to the number of possible compounds in an untargeted search. The original PCD contained over 400 predicted trisaccharides for each possible combination of aglycone and glycone shown in Figure 1. The MassHunter software allowed for a more automated approach less prone to errors in mass calculation and MS/MS spectra interpretation. MSC supplemented the process by identifying the fragments in a spectrum, in addition to their likelihood to form based on the bond strengths.

Once the initial database was created, a representative grape sample was run and immediately analyzed for database hits. The hits could be used for a tandem MS inclusion list where the same sample could be rerun within minutes of the first analysis to obtain MS/MS spectra. The resulting data could be analyzed in MSC for potential compound identification. In this manner, the samples could be quickly analyzed to ensure that the experiment ran as intended, to determine if more tandem MS spectra were needed, or to ascertain if instrument settings, such as collision energies, were suitable for the analysis. Using MassHunter's cross-program compatibility and automation increased overall productivity and freed up the instrument for further experimentation.

In this study, it was necessary to determine if ESI or APCI should be used for analysis. Recent published studies have used both APCI and ESI. Hayasaka *et al.* (2013)⁷ found that APCI was preferred over ESI for increased signal; however, in the work of Noestheden *et al.* (2018),⁵ ESI was used for their analysis without mention of APCI. The two sources were compared for each of the compounds tentatively identified. Results of this analysis are found in Figures 4, 5, and 6. Even after source optimization, the use of APCI resulted in a loss of signal for trisaccharides with an *m/z* greater than 550. This difference and the increased abundance of many of the other glycosides led to the choice of ESI for analysis. The loss of trisaccharide signal with APCI could be due to a lack of volatility at a certain mass threshold. ESI showed more versatility with higher mass compounds and was less sensitive to mass bias.

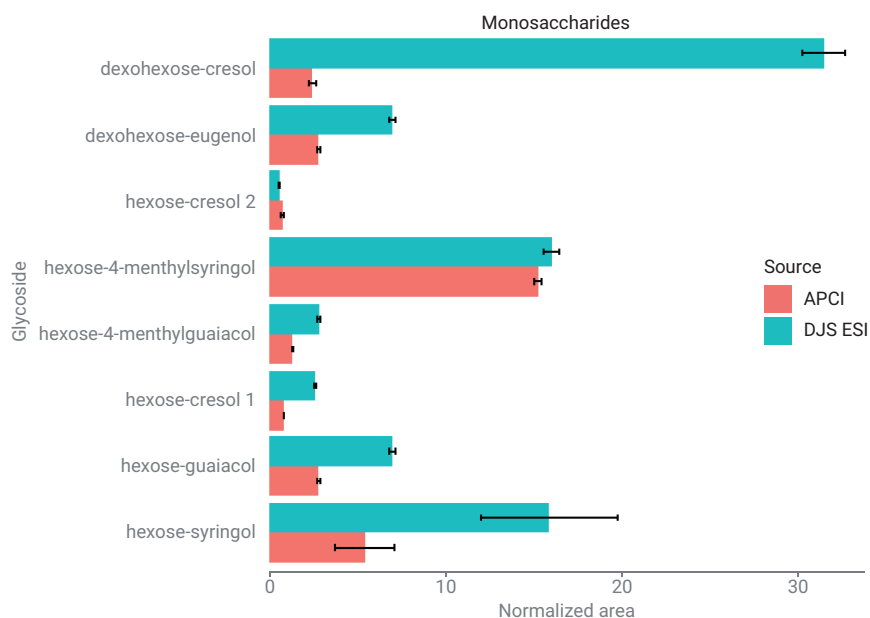


Figure 4. A comparison of the Dual Jet Spray ESI and APCI LC/MS sources. Source settings were optimized for the highest abundance for monosaccharides. In every case, ESI had the same or higher abundance for each molecule.



Figure 5. A comparison of the Dual Jet Spray ESI and APCI LC/MS sources. Source settings were optimized for the highest abundance for disaccharides. In every case, ESI had the same or higher abundance for each molecule.

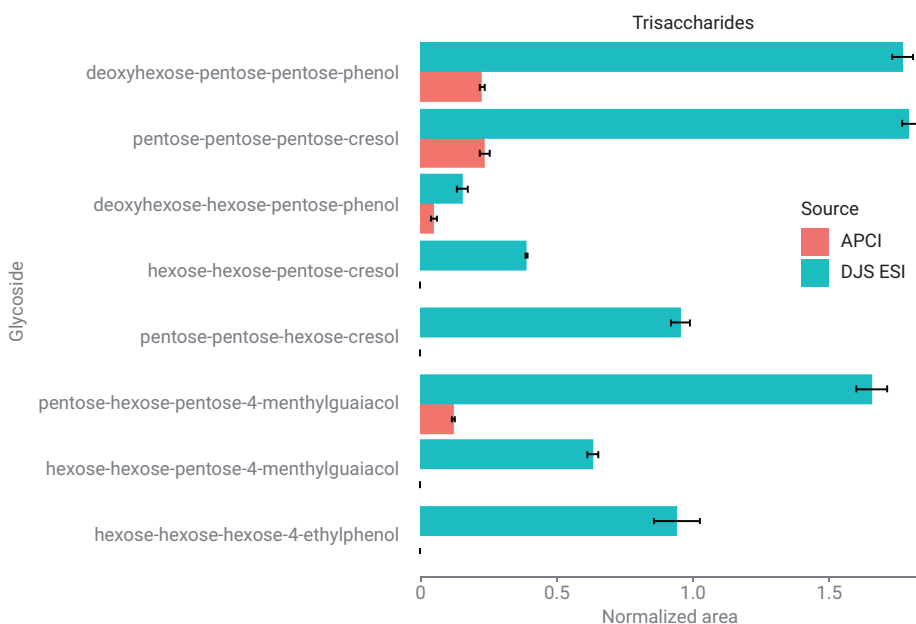


Figure 6. A comparison of the Dual Jet Spray ESI and APCI LC/MS sources. Source settings were optimized for the highest abundance for trisaccharides. In every case, ESI had the same or higher abundance for each molecule. Four trisaccharide glycosides were not detectable with APCI.

Semiquantitation of winemaking samples revealed that the relative abundances of volatile-phenol glycosides change throughout the winemaking process. Overall, the glycosides are highest in abundance at the start of winemaking and lose roughly one-fourth of their abundance throughout winemaking. Most of the hydrolysis occurred during the primary fermentation with *Saccharomyces cerevisiae* yeast. The “all ions” workflow allowed more definitive identification of compound peaks within MassHunter Quantitative Analysis even when retention times changed due to matrix effects. In addition, using the “all ions” approach helped distinguish compounds from noise, preventing misidentification and incorrect integrations of low-abundance or nondetected compounds such as hexose-guaiacol or hexose-hexose-pentose-4-methylguaiacol. Without the ability to see qualifier ions from the “all ions” data, during integration, it is possible that a false positive would have been selected because there are no standards for comparison.

Conclusion

An Agilent 6545 Accurate Mass Q-TOF LC/MS coupled with Agilent MassHunter was used to identify and semiquantitate volatile-phenol glycosides in smoke-affected grapes. Roughly three-quarters of the glycosides remained in the finished wines, and can continue to act as a reservoir for off-flavors as the wine ages. The Agilent MassHunter PCDL Manager, Find-by-Formula algorithm, and MSC permitted a broader search for volatile-phenol glycosides resulting in the identification of newly reported glycosidic compounds, including the first report of smoke-related trisaccharide glycosides in grapes. An "all ions" workflow for semiquantitation allowed a robust method to be used across samples of varying analyte abundances and sample matrices. Smoke taint encompasses hundreds of potential compounds; the use of Agilent software in an untargeted LC/MS analysis facilitated the tentative identification and semiquantitation of smoke taint compounds in the absence of commercial standards. Future studies may build on this method to effectively evaluate winemaking techniques to reduce the amount of volatile-phenol glycosides in the finished wines.

References

1. Caffrey, A. *et al.* Changes in Smoke-Taint Volatile Phenol Glycosides in Wildfire Smoke-Exposed Cabernet Sauvignon Grapes Throughout Winemaking. *American Journal of Enology and Viticulture* **2019**, *70*(4), **Pages To be determined**
2. Parker, M. *et al.* Contribution of Several Volatile Phenols And Their Glycoconjugates to Smoke-Related Sensory Properties of Red Wine. *J. Agric. Food Chem.* **2012**, *60*, 2629–2637.
3. Kennison, K. R. *et al.* Effect of Timing and Duration of Grapevine Exposure to Smoke on the Composition and Sensory Properties of Wine. *Aust. J. Grape Wine Res.* **2009**, *15*, 228–237.
4. Hayasaka, Y. *et al.* Investigation into the Formation of Guaiacol Conjugates in Berries and Leaves of Grapevine *Vitis vinifera* L. cv. Cabernet Sauvignon Using Stable Isotope Tracers Combined with HPLC-MS And MS/MS Analysis. *J. Agric. Food Chem.* **2010a**, *58*, 2076–208.
5. Noestheden, M. *et al.* Detailed Characterization of Glycosylated Sensory-Active Volatile Phenols in Smoke-Exposed Grapes and Wine. *Food Chem.* **2018**, *259*, 147–156.
6. Caffrey, A.; Hjelmeland, A.; Zweigenbaum, J.; Ebeler, S. E.; Identification and Semiquantitation of Monoterpene Glycosides in Ripening Muscat of Alexandria Grapes. *Agilent Technologies Application Note*, publication number 5991-9448EN, **2018**. Retrieved from Agilent Technologies: https://www.agilent.com/cs/library/applications/application_monoterpene_glycosides_grapes_lc_qtof_ms_5991-9448EN_agilent.pdf
7. Hayasaka, Y. *et al.* Assessing the Impact of Smoke Exposure in Grapes: Development and Validation of a HPLC-MS/MS Method for the Quantitative Analysis of Smoke-Derived Phenolic Glycosides in Grapes and Wine. *J. Agric. Food Chem.* **2013**, *61*(1), 25–33.

www.agilent.com/chem

This information is subject to change without notice.

© Agilent Technologies, Inc. 2019
Printed in the USA, November 5, 2019
5994-1563EN