

Free and Bound Volatile Phenols in Smoke-Exposed Wines- Biomarkers, Machine-learning, and Model Prediction

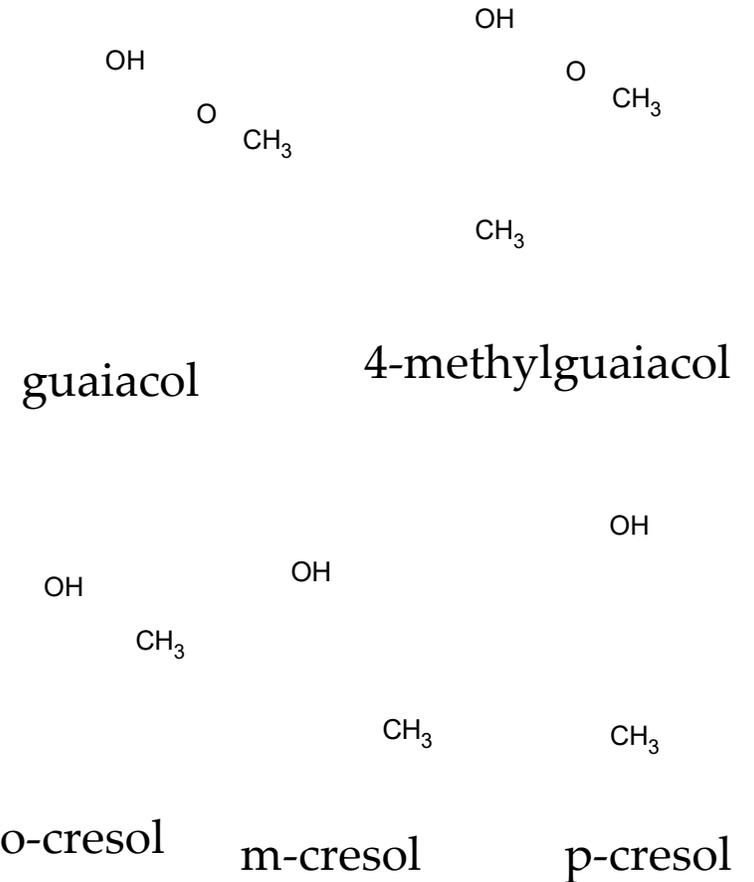
- Ruiwen Yang¹, Yanping L. Qian¹, Armando Alcazar^{1,2}, Ye Feng³, Ling Huang¹, Danye Zhu¹, and Michael C. Qian^{1*}
- ¹Department of Food Science and Technology, Oregon State University, 100 Wiegand Hall, Corvallis OR 97331
- ²OSU Mass Spectrometry Center, Oregon State University, Corvallis, OR 97330
- ³President and CEO, Spectra Scientific, Portland, OR 97229

Outlines

- Introduction to smoke related volatile phenols
- Free and total volatile phenols by GC-MS
- Volatile phenol glycoside by LC-MS
- Machine-learning and modeling
- Conclusion

What we know for smoke compounds

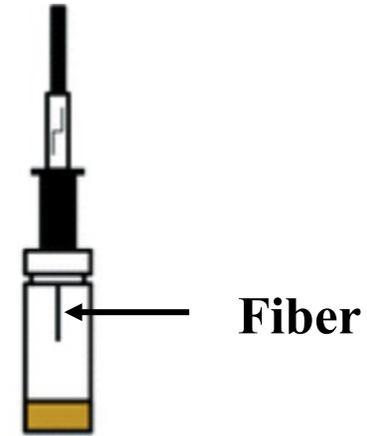
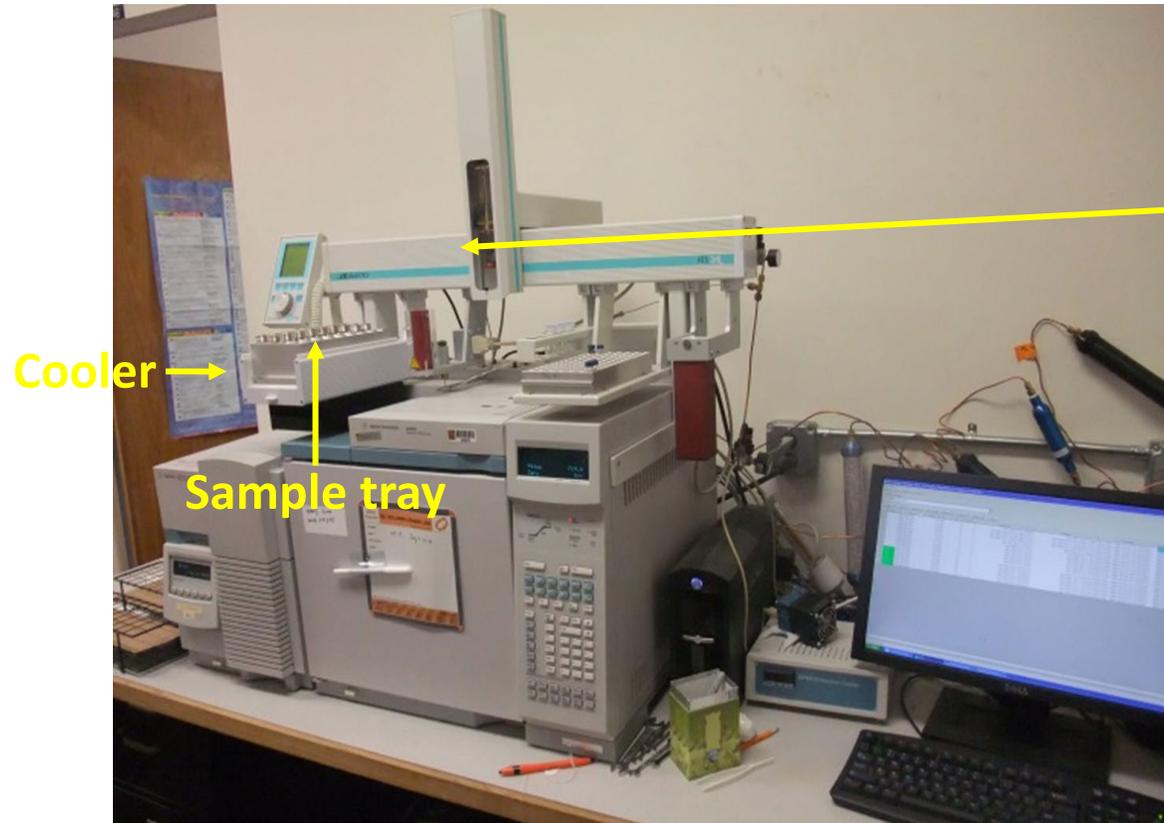
- Volatile phenols have been associated with smoke exposure
- The most frequently studied volatile phenols are guaiacol, 4-methylguaiacol and o-cresol, m-cresol and p-cresol, although a few other volatiles phenols are sometimes studied
- In the event of smoke, the grapevine can absorb these volatile phenols and convert them to phenol-glycosides
- Bound phenol glycosides do not have smoke taint, however, they can convert to free form during fermentation, aging or in mouth, impart off-flavor



Challenges with smoke analysis

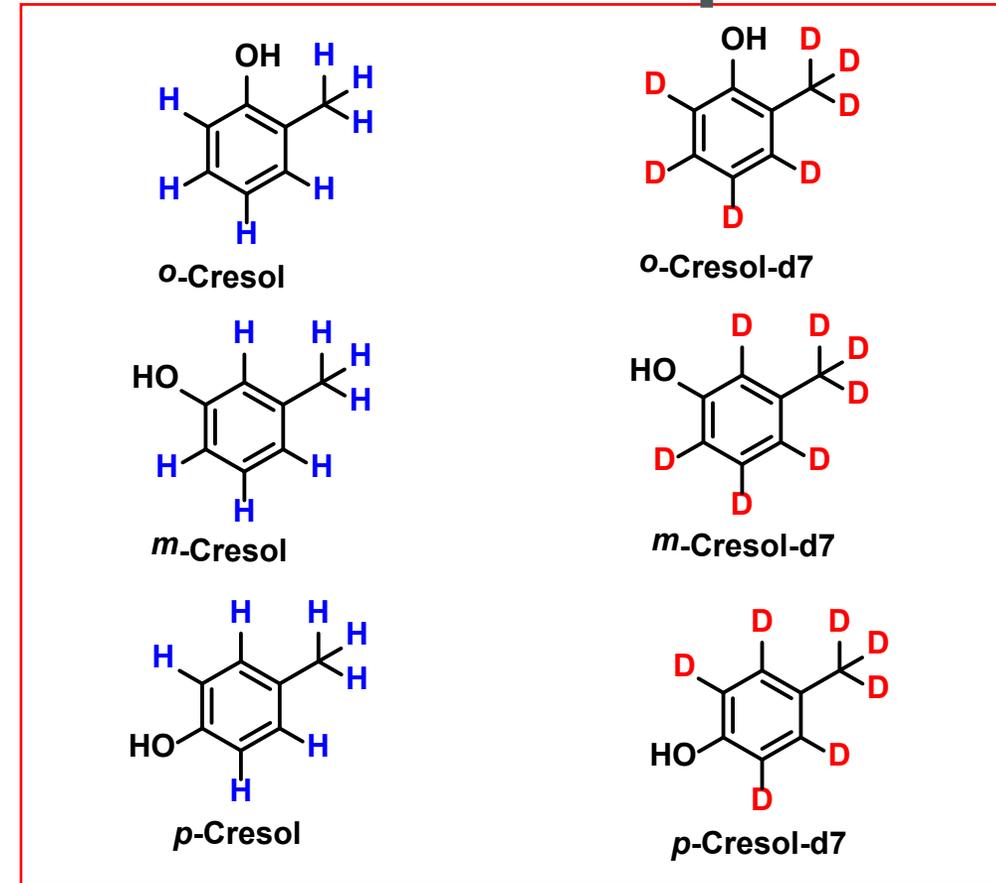
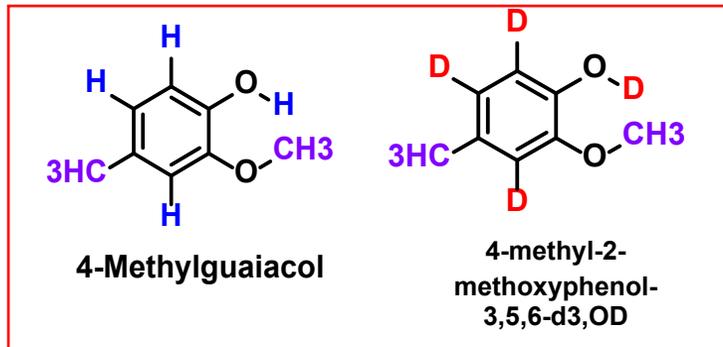
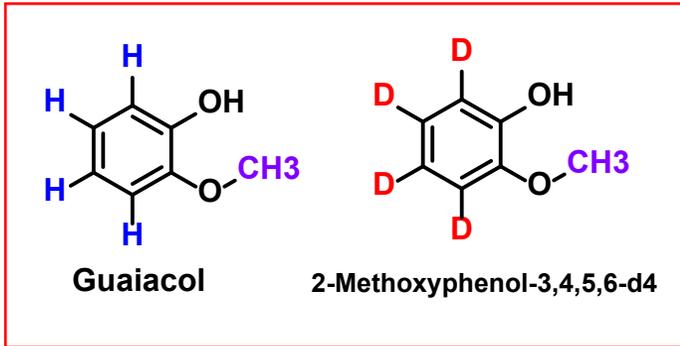
- Volatile phenols exist in normal wine, are part of wine aroma
 - Difficult to distinguish the “good” from the “bad”
- Volatile phenols present at very low concentration
 - Need reliable and sensitive instrumentation and robust analytical method

Sensitive volatile phenol analysis by Solid-phase microextraction (SPME)-GC-MS



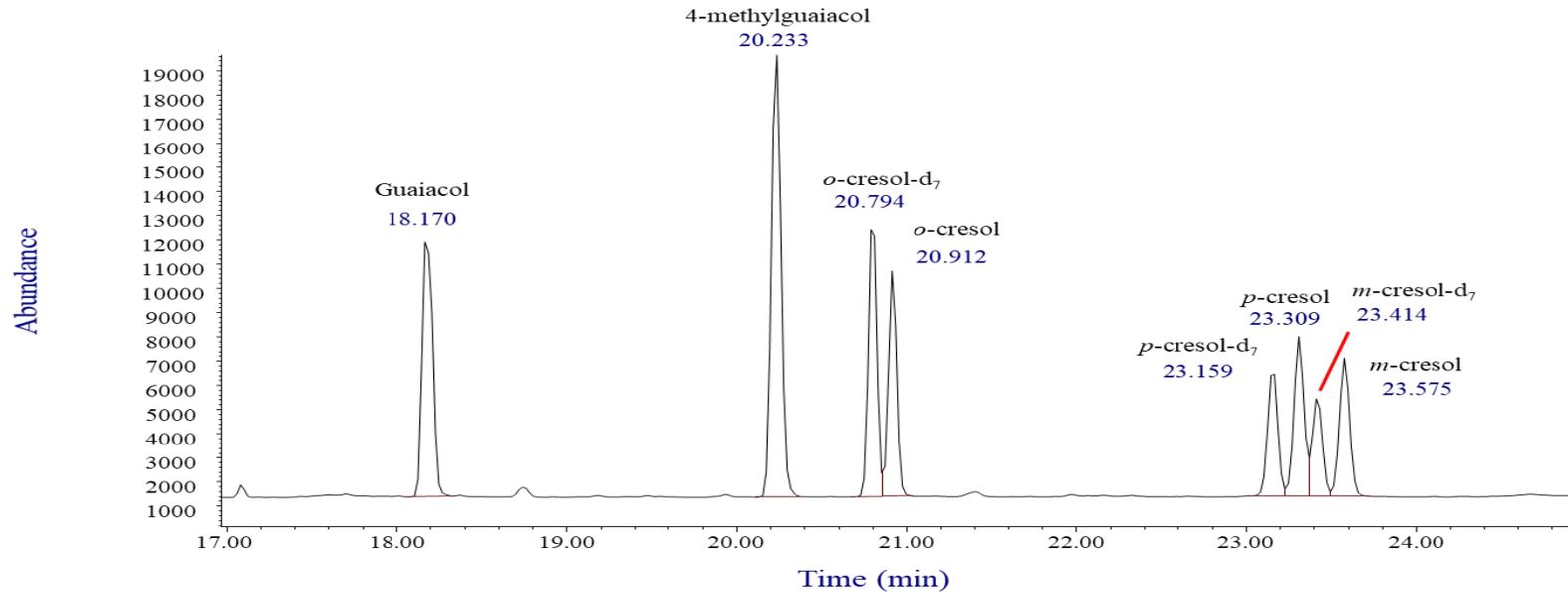
Fiber:
CAR/DVB/PDMS, 2cm

Isotope-labeled Compounds as Internal Standards to Eliminate Wine Matrix Impact



D- Exact mass 2.0141; **H**- Exact mass 1.0078

Free Volatile Analysis

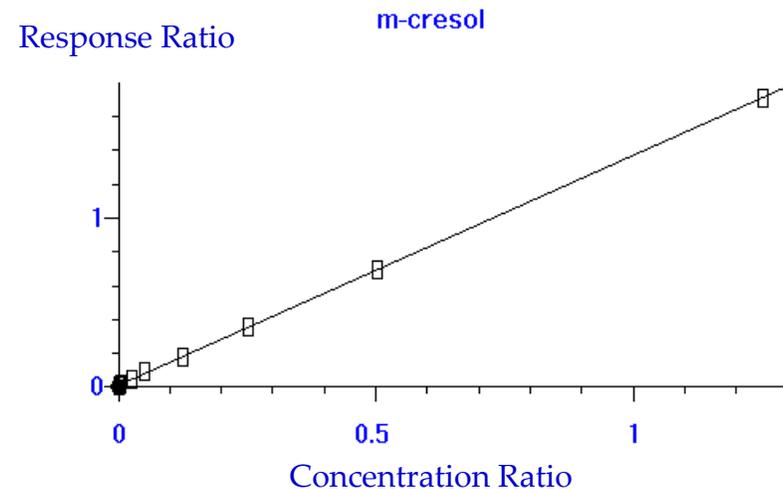
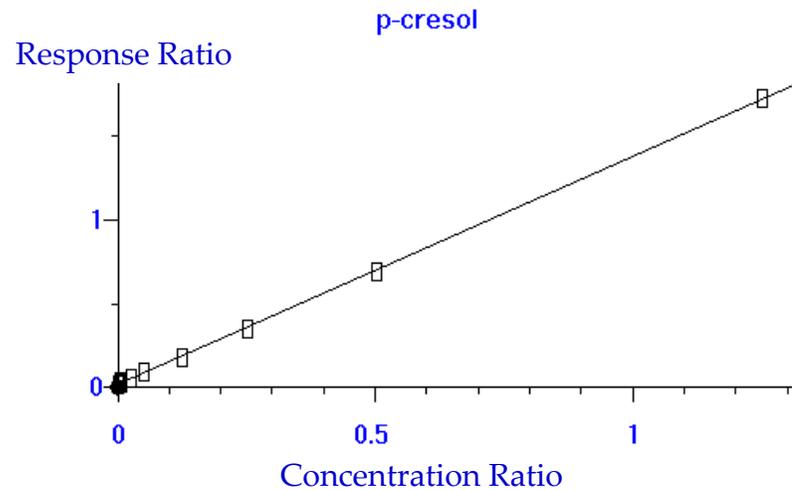
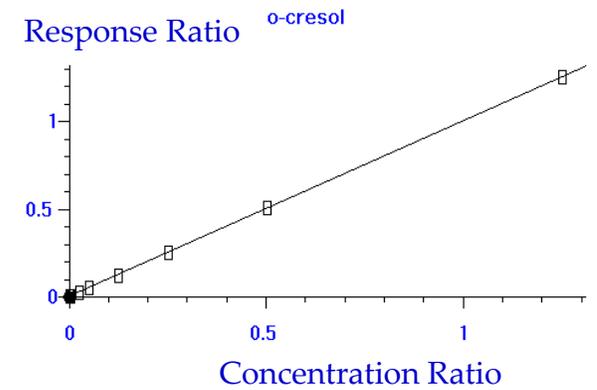
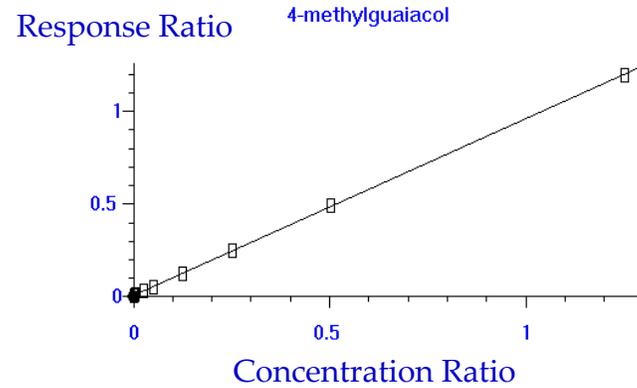
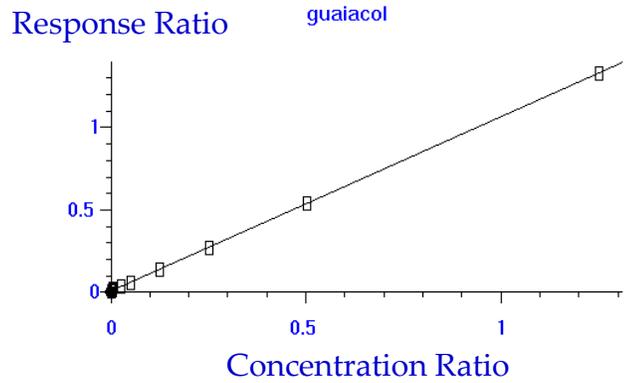


- Free phenol analysis
 - 10 ml wine + 10 ul internal standard
 - Solid-phase Micro-extraction
50°C/25min
 - GC-MS analysis (30 min!)

Calibration Curve, Quantitation Limits

SPME-GC-MS for total VPs

Standard calibration curves: limitation 0.1 $\mu\text{g/L}$ free, 0.01 $\mu\text{g/L}$ total



Analytical Data Comparison with Commercial Certified Lab (free)

| | guaiacol (ug/L) | | 4-methylguaiacol | | o-cresol | p-cresol | m-cresol |
|------------------------|-----------------|----------|------------------|----------|---------------------------------|----------|----------|
| | Qian's lab | Cert lab | Qian's lab | Cert lab | (Cert lab did not test cresols) | | |
| sample 1 | 4.0 | | 1.0 | | 3.3 | 1.1 | 1.3 |
| Sample2 | 3.7 | 3.9 | 0.8 | 0.8 | 3.1 | 0.9 | 1.6 |
| sample 3 | 4.0 | 4.2 | 0.9 | 1.1 | 2.8 | 1.5 | 1.4 |
| sample 4 | 3.9 | 4.0 | 0.9 | 1.1 | 3.0 | 1.9 | 1.5 |
| sample 5-1 | 2.9 | 3.0 | 0.7 | 0.7 | 2.4 | 1.2 | 1.0 |
| sample 5-2 (duplicate) | 3.0 | | 0.7 | | 2.5 | 0.9 | 1.1 |
| sample 6 | 3.3 | 3.4 | 0.8 | 0.8 | 2.3 | 1.1 | 1.1 |
| sample 7 | 4.2 | 4.4 | 1.0 | 1.1 | 3.0 | 0.8 | 1.4 |
| sample 8 | 6.5 | | 1.7 | | 3.8 | 1.5 | 2.2 |
| sample 9 | 21.3 | | 5.5 | | 8.7 | 2.7 | 6.2 |



Analytical Quality Control

- Verify check sample for every 20 analysis
- Duplicate analysis every 10 samples

| | Guaiacol | 4-methylguaiacol | <i>m</i> -cresol | <i>o</i> -cresol | <i>p</i> -cresol |
|----------------|----------|------------------|------------------|------------------|------------------|
| Sample 1-1 | 12.06 | 2.36 | 6.78 | 5.88 | 3.67 |
| Sample 1-2 | 11.85 | 2.36 | 6.68 | 5.95 | 3.76 |
| CV (%) | 1% | 0% | 1% | 1% | 2% |
| Check-1.0 µg/L | 1.04 | 1.06 | 0.98 | 1.05 | 1.07 |



Total volatile phenol analysis

- Much more phenol glycosides than free volatile phenols in grapes and wine
- Direct analysis of phenol glycosides by LC-MS is costly and time consuming
- Less expensive method is “total volatile phenol analysis”
 - SPME-GC-MS analysis after strong acid hydrolysis
 - Not reliable due to artifact formation under strong acid condition and low pH
 - No good universal method
 - Results are lab-dependent (varied hydrolysis conditions)
- Total phenol analysis
 - 2 ml wine, pH 1-pH1.2/100°C/4hr
 - 8 mL citrate buffer, pH 3.5
 - SPME 50°C/25min
 - GC-MS analysis

Volatile phenol analysis in smoke exposed wines



- In 2020, we analyzed 377 Smoke exposed red wine, 91 Smoke exposed white (rose) wine
- Different degree of smoke exposure
- Analyzed both free and total volatile phenols

With collaboration with Dr. Elizabeth Tomasino

Objective 1

- Build database information for smoke exposed wine in Oregon and understand volatile phenol correlation in smoke exposed wine

Control Wine Analysis-86 Pinot noir for Baseline

- 21 wines from 2013
 - 21 wines from 2014
 - 24 wines from 2015
 - 20 wines from 2016
-
- All wines were made commercially in industrial scale
 - No barrel aging

Free Phenols in 86 Control Pinot noir Wine (2013-2016)

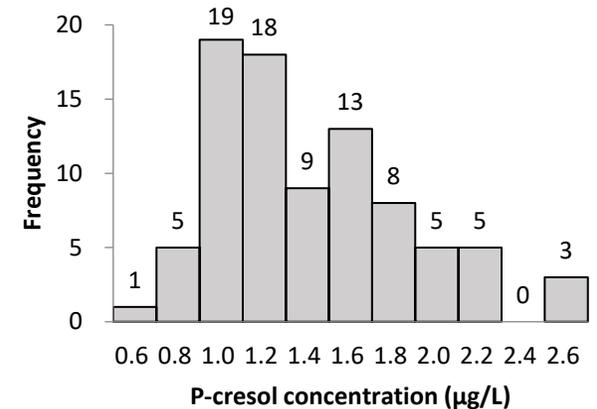
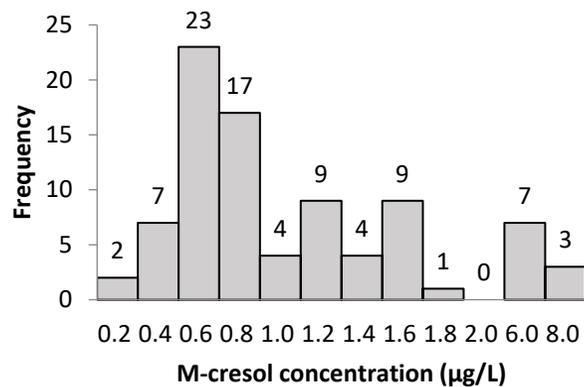
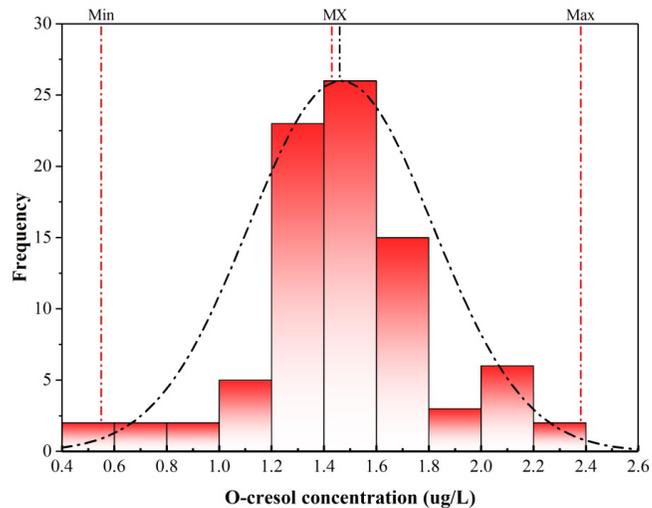
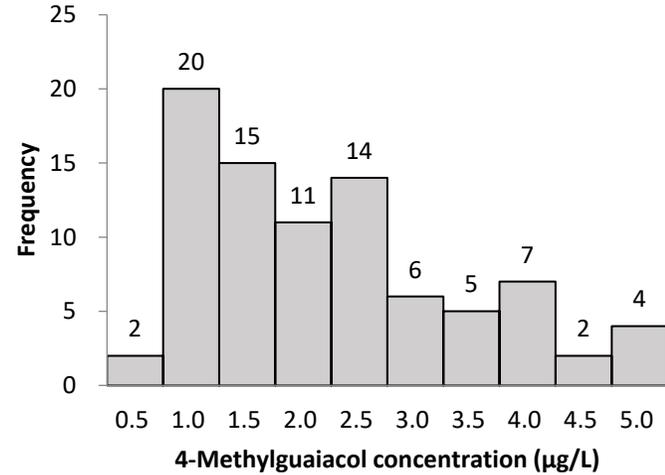
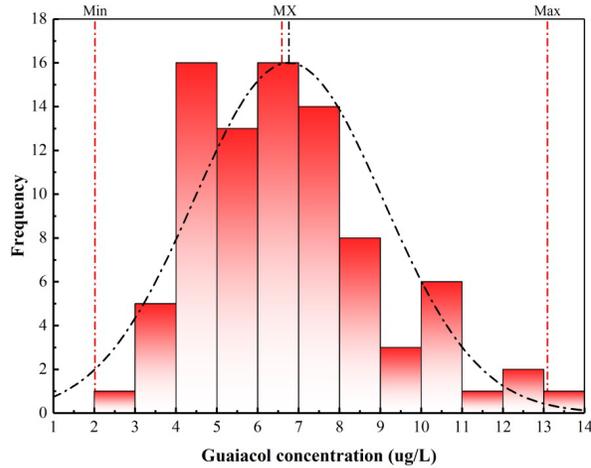
Concentration ($\mu\text{g/L}$)

| | | guaiacol | 4-methylguaiacol | o-cresol | p-cresol | m-cresol |
|-----|---------|----------|------------------|----------|----------|----------|
| X | Average | 6.76 | 1.98 | 1.46 | 1.31 | 1.16 |
| M | Median | 6.59 | 1.76 | 1.43 | 1.20 | 0.74 |
| Min | Minimum | 2.02 | 0.41 | 0.55 | 0.51 | 0.17 |
| Max | Maximum | 13.09 | 4.86 | 2.38 | 2.55 | 6.87 |



Free phenol concentration distribution in 86 control Pinot noir wine

(Number of samples in a concentration range)



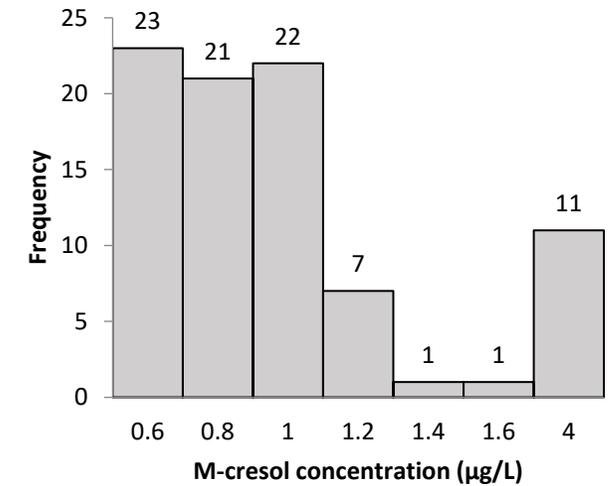
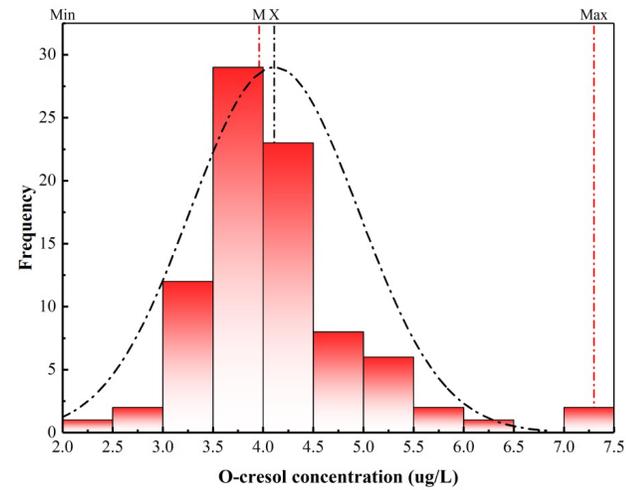
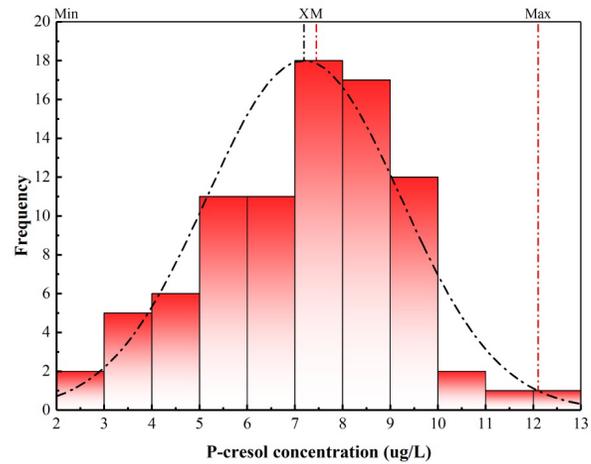
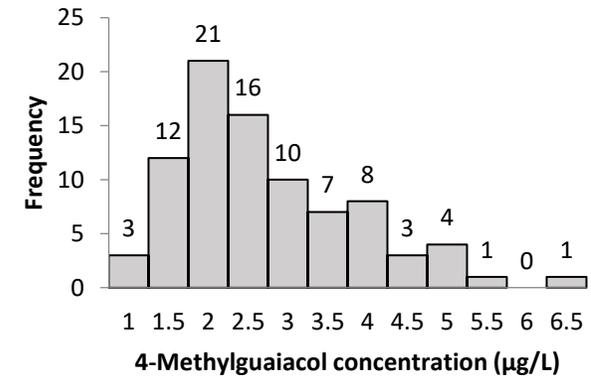
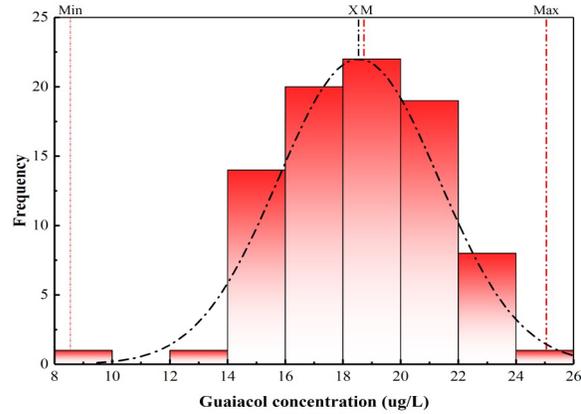
Total phenols in 86 Control Pinot noir Wines- 2013-2016

Concentration ($\mu\text{g/L}$)

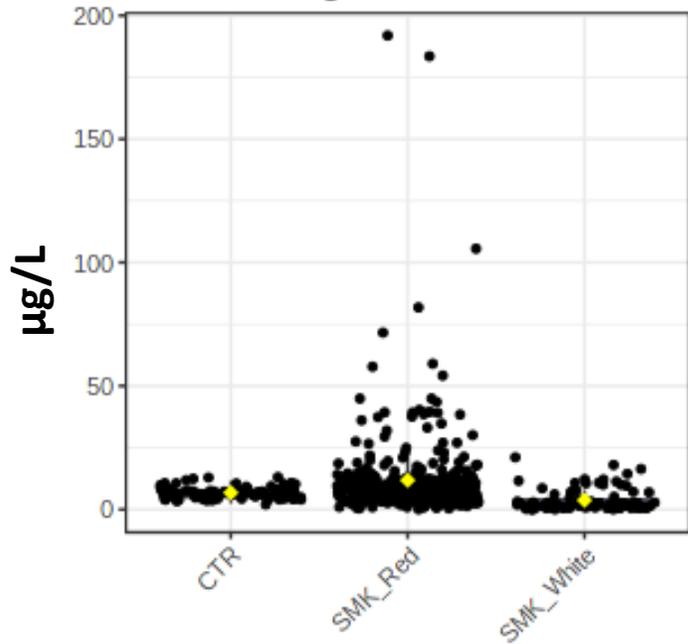
| | | guaiacol | 4-methylguaiacol | o-cresol | p-cresol | m-cresol |
|-----|---------|----------|------------------|----------|----------|----------|
| X | Average | 18.5 | 2.5 | 4.1 | 7.2 | 1.0 |
| M | Median | 18.7 | 2.3 | 4.1 | 7.4 | 0.8 |
| Min | Minimum | 8.6 | 0.7 | 2.0 | 2.0 | 0.4 |
| Max | Maximum | 25.1 | 6.2 | 7.3 | 12.1 | 3.8 |



Total Phenol Distribution in 86 Control Pinot noir Wines-2013-2016



guaiacol_F



Volatile Phenol Analysis- Free (µg/L)

377 Smoke exposed red wine
 91 Smoke exposed white (rose) wine
 86 Control red wine

Nomenclature

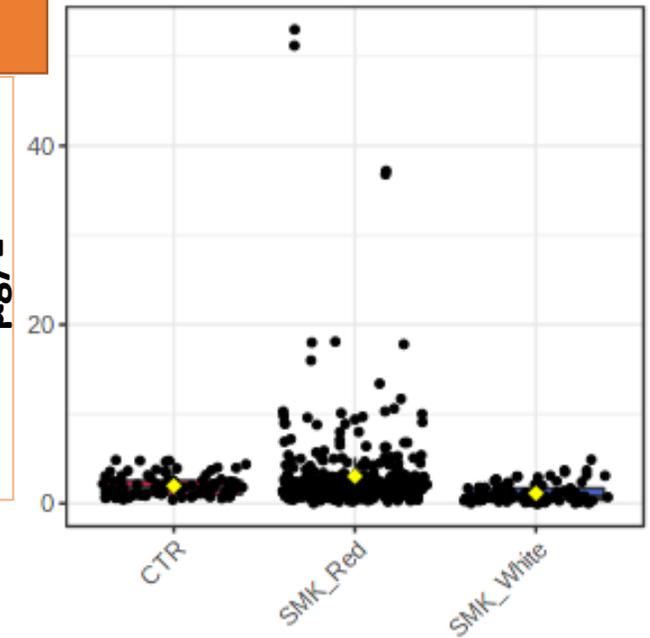
_F- Free form

-T- total after hydrolysis (HCl, 100C/4hr) (µg/L)



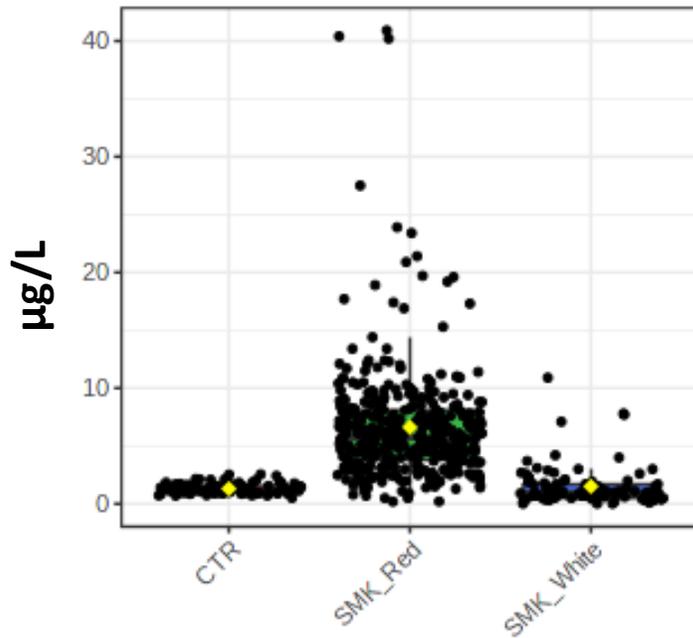
µg/L

4-methylguaiacol_F

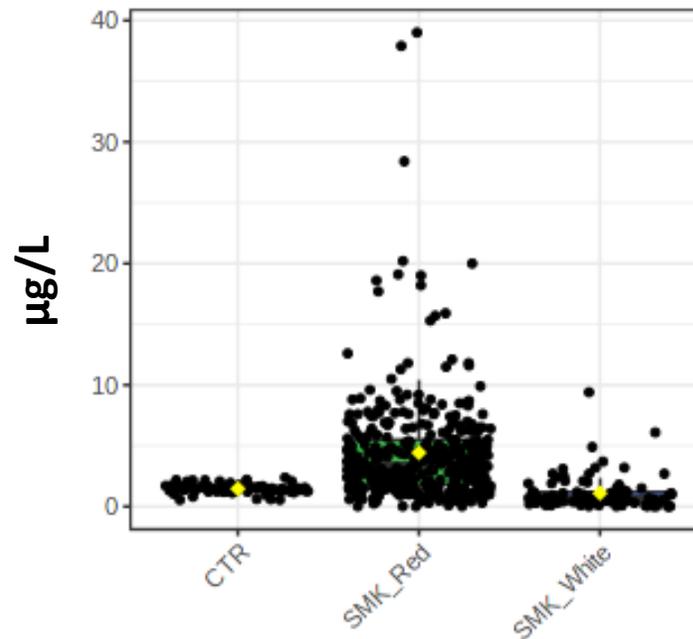


Smoke exposed red wine presents the biggest dispersion of the data. o-,m- and p-cresol exhibit the biggest difference across sample.

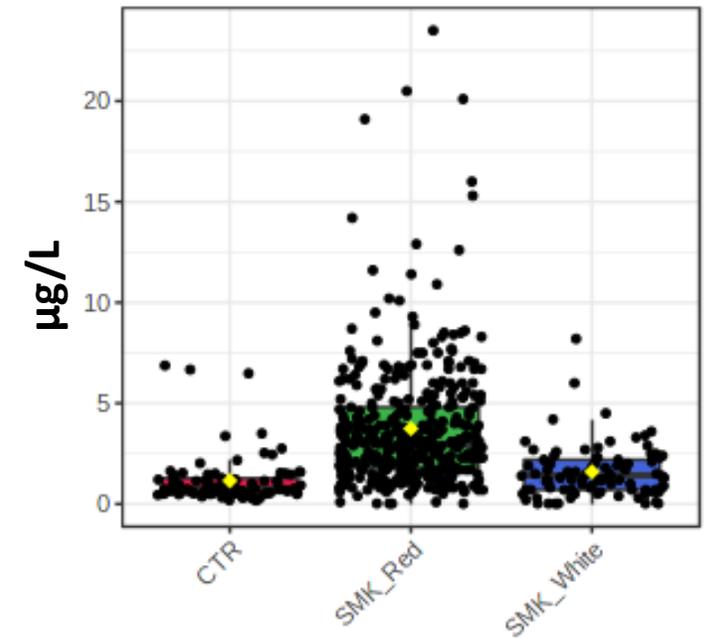
o-cresol_F



m-cresol_F



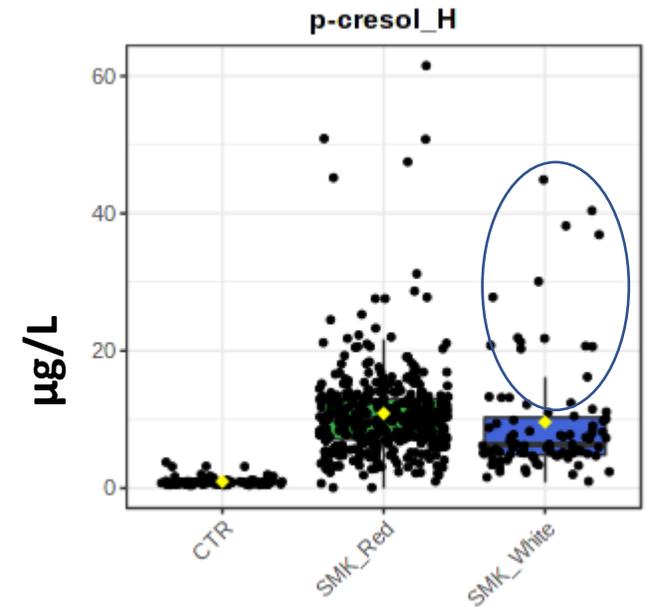
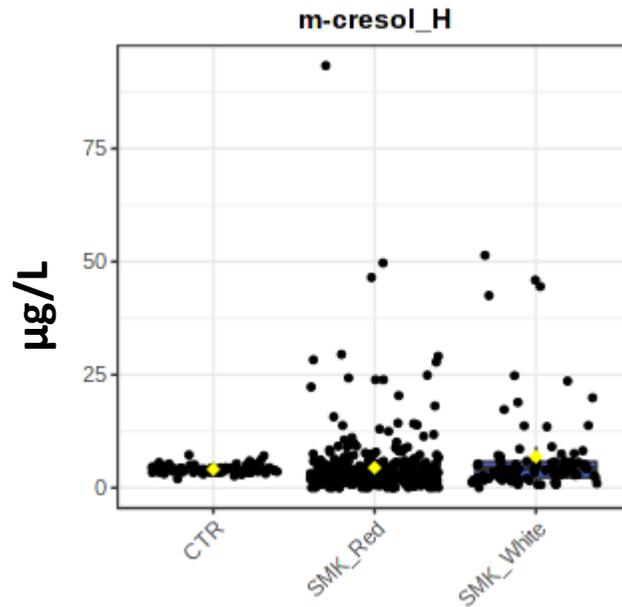
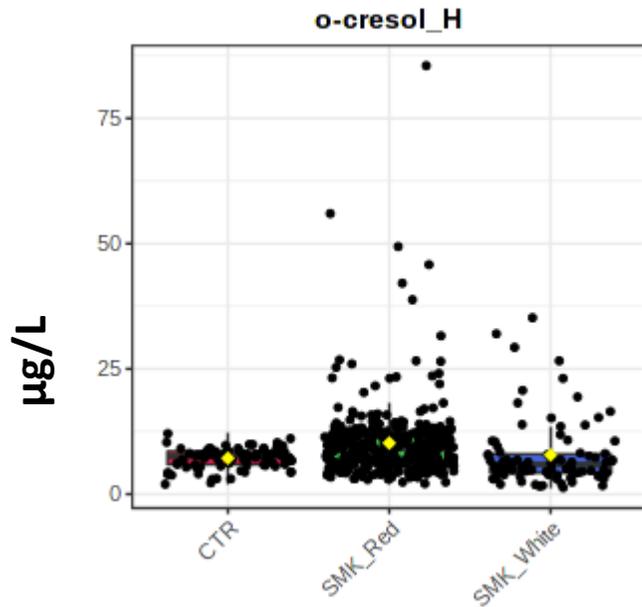
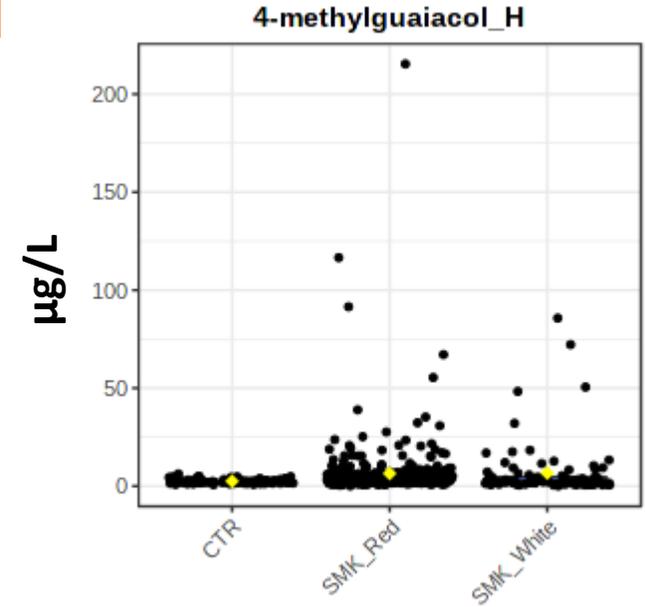
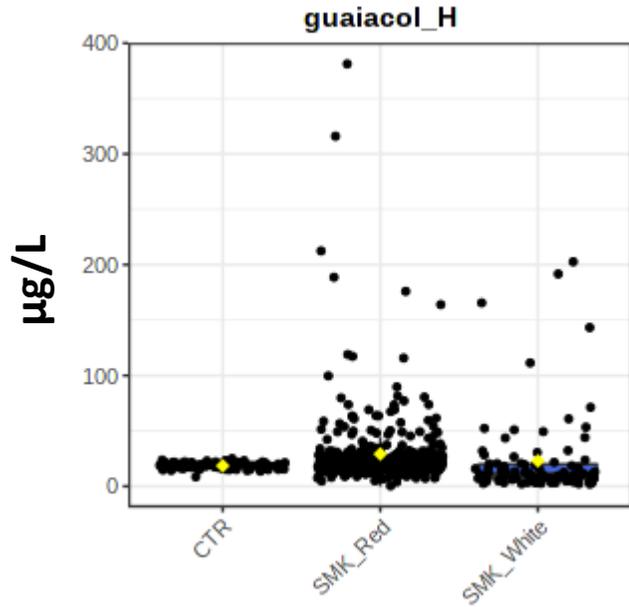
p-cresol_F



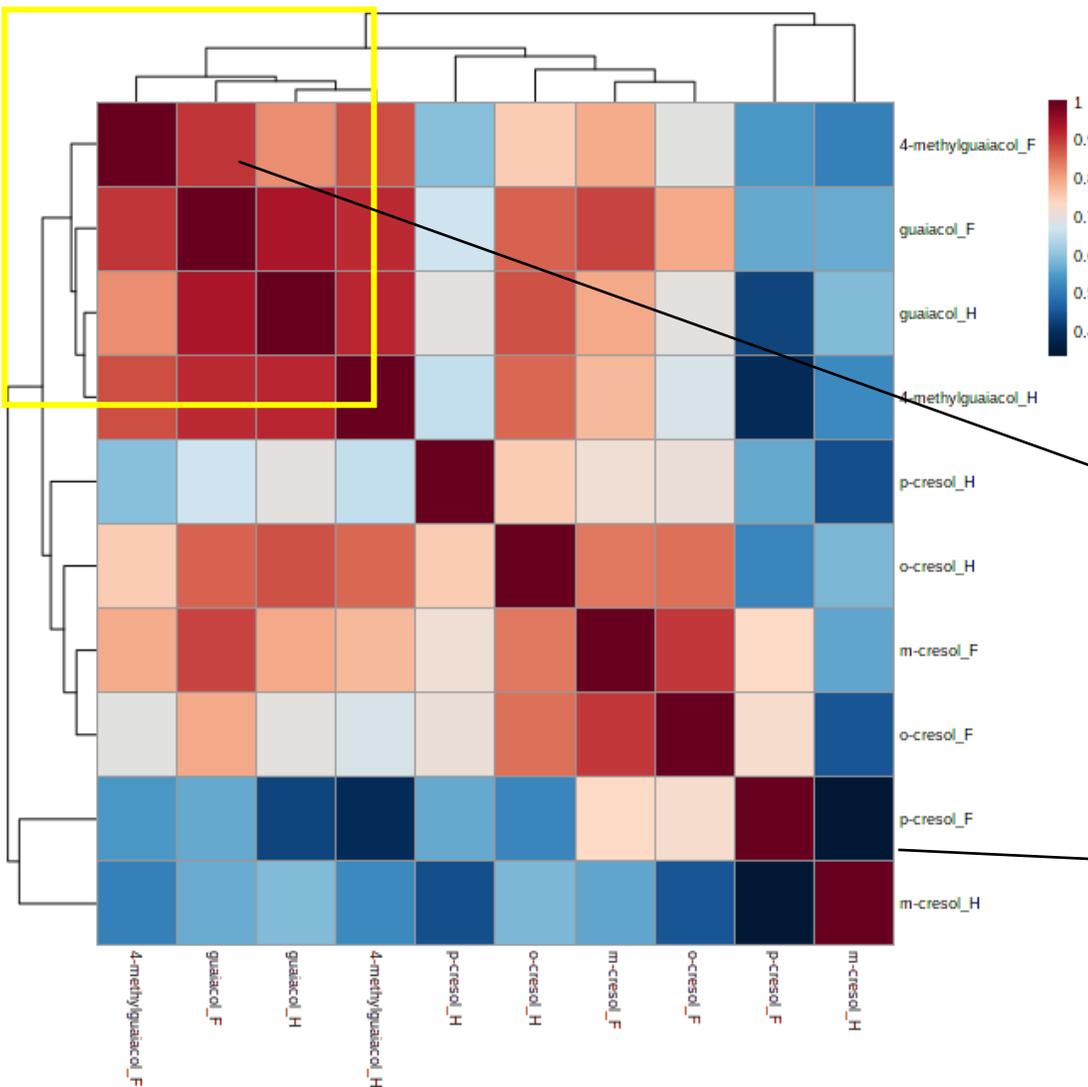
Volatile Phenol Analysis-Total ($\mu\text{g/L}$)

377 Smoke exposed red wine
91 Smoke exposed white wine
86 Control red wine

Smoke exposed red wine presents the biggest dispersion of the data. p-cresol exhibited the highest difference.



Correlation Heatmap and Matrix



Heatmap was used to identify correlation of compounds. The value ranges from -1 to 1. As shown in the heatmap, all correlations are positive, ranging from 0.34 (blue, bound m-cresol vs free p-cresol) to 0.94 (deep red, guaiacol free form vs bound guaiacol).

| | 4-methylguaiacol_F | guaiacol_F | guaiacol_H | 4-methylguaiacol_H | p-cresol_H | o-cresol_H | m-cresol_F | o-cresol_F | p-cresol_F | m-cresol_H |
|--------------------|--------------------|------------|------------|--------------------|------------|------------|------------|------------|------------|------------|
| 4-methylguaiacol_F | 1.00 | 0.91 | 0.82 | 0.88 | 0.60 | 0.75 | 0.79 | 0.69 | 0.55 | 0.51 |
| guaiacol_F | 0.91 | 1.00 | 0.94 | 0.92 | 0.67 | 0.86 | 0.89 | 0.80 | 0.57 | 0.57 |
| guaiacol_H | 0.82 | 0.94 | 1.00 | 0.92 | 0.70 | 0.88 | 0.80 | 0.70 | 0.43 | 0.59 |
| 4-methylguaiacol_H | 0.88 | 0.92 | 0.92 | 1.00 | 0.66 | 0.86 | 0.78 | 0.68 | 0.40 | 0.52 |
| p-cresol_H | 0.60 | 0.67 | 0.70 | 0.66 | 1.00 | 0.75 | 0.72 | 0.71 | 0.57 | 0.45 |
| o-cresol_H | 0.75 | 0.86 | 0.88 | 0.86 | 0.75 | 1.00 | 0.84 | 0.85 | 0.52 | 0.59 |
| m-cresol_F | 0.79 | 0.89 | 0.80 | 0.78 | 0.72 | 0.84 | 1.00 | 0.90 | 0.74 | 0.56 |
| o-cresol_F | 0.69 | 0.80 | 0.70 | 0.68 | 0.71 | 0.85 | 0.90 | 1.00 | 0.73 | 0.45 |
| p-cresol_F | 0.55 | 0.57 | 0.43 | 0.40 | 0.57 | 0.52 | 0.74 | 0.73 | 1.00 | 0.34 |
| m-cresol_H | 0.51 | 0.57 | 0.59 | 0.52 | 0.45 | 0.59 | 0.56 | 0.45 | 0.34 | 1.00 |

Pearson correlation

Control Red Wine vs Smoke Exposed Red Wine (t-test)

377 Smoke exposed red wine
86 Control red wine

Nomenclature

F- Free form

T- After hydrolysis (HCl, 100C/4hr) (µg/L)

Smoke exposed red/ CTR red wine

| | FC | log2(FC) | P-pval |
|--------------------|-------|----------|---------|
| p-cresol_T | 11.12 | 3.47 | 3.9E-35 |
| o-cresol_F | 5.05 | 2.34 | 2.5E-21 |
| p-cresol_F | 3.23 | 1.69 | 1.7E-13 |
| m-cresol_F | 3.03 | 1.60 | 9.2E-10 |
| o-cresol_T | 1.42 | 0.51 | 1.7E-04 |
| guaiacol_T | 1.57 | 0.66 | 3.1E-03 |
| guaiacol_F | 1.75 | 0.81 | 6.7E-03 |
| 4-methylguaiacol_T | 2.57 | 1.36 | 1.3E-02 |
| 4-methylguaiacol_F | 1.54 | 0.62 | 5.5E-02 |
| m-cresol_T | 1.09 | 0.12 | 6.5E-01 |

Univariate data analysis was performed to establish differences between control and smoke exposed red wine.

Total p-cresol is most discriminating compound for smoke exposure, followed by free o-,p-,m-cresol.

With exception of free 4-methylguaiacol and total m-cresol, all compounds were significantly higher in smoke exposed wine ($p > .05$). After hydrolysis, total p-cresol presented the most significant difference (FC = 11.12, $p = 3.9 \times 10^{-35}$).

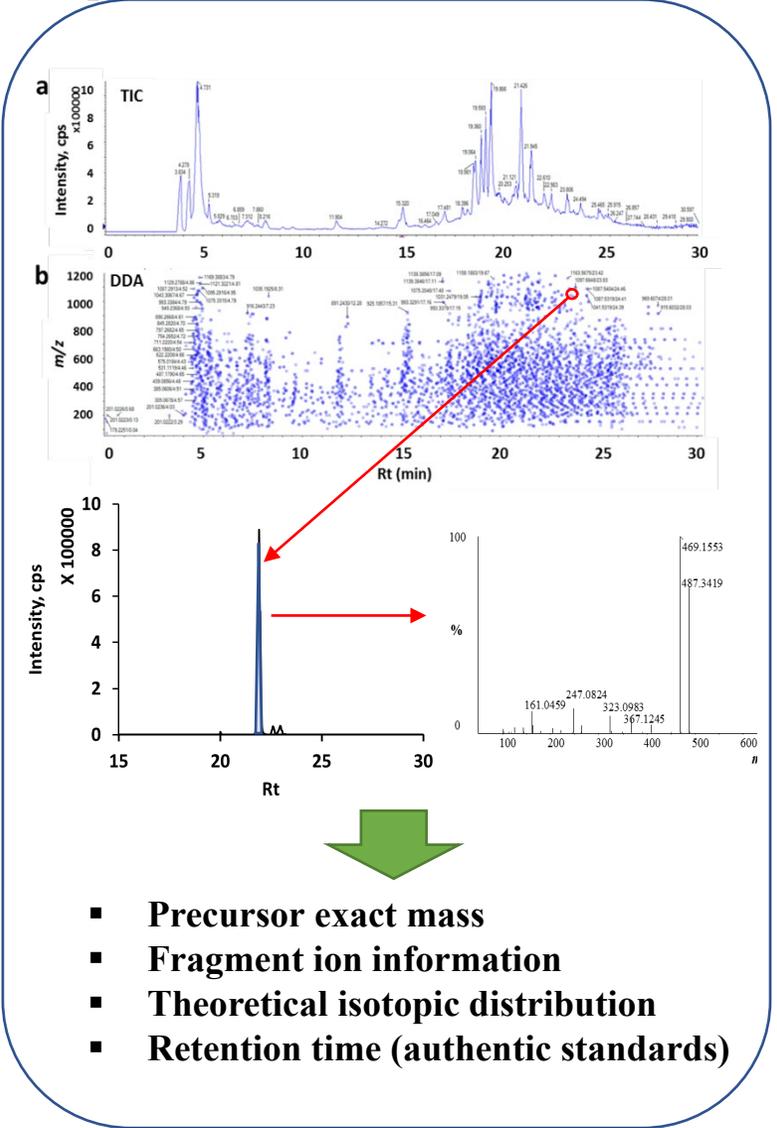
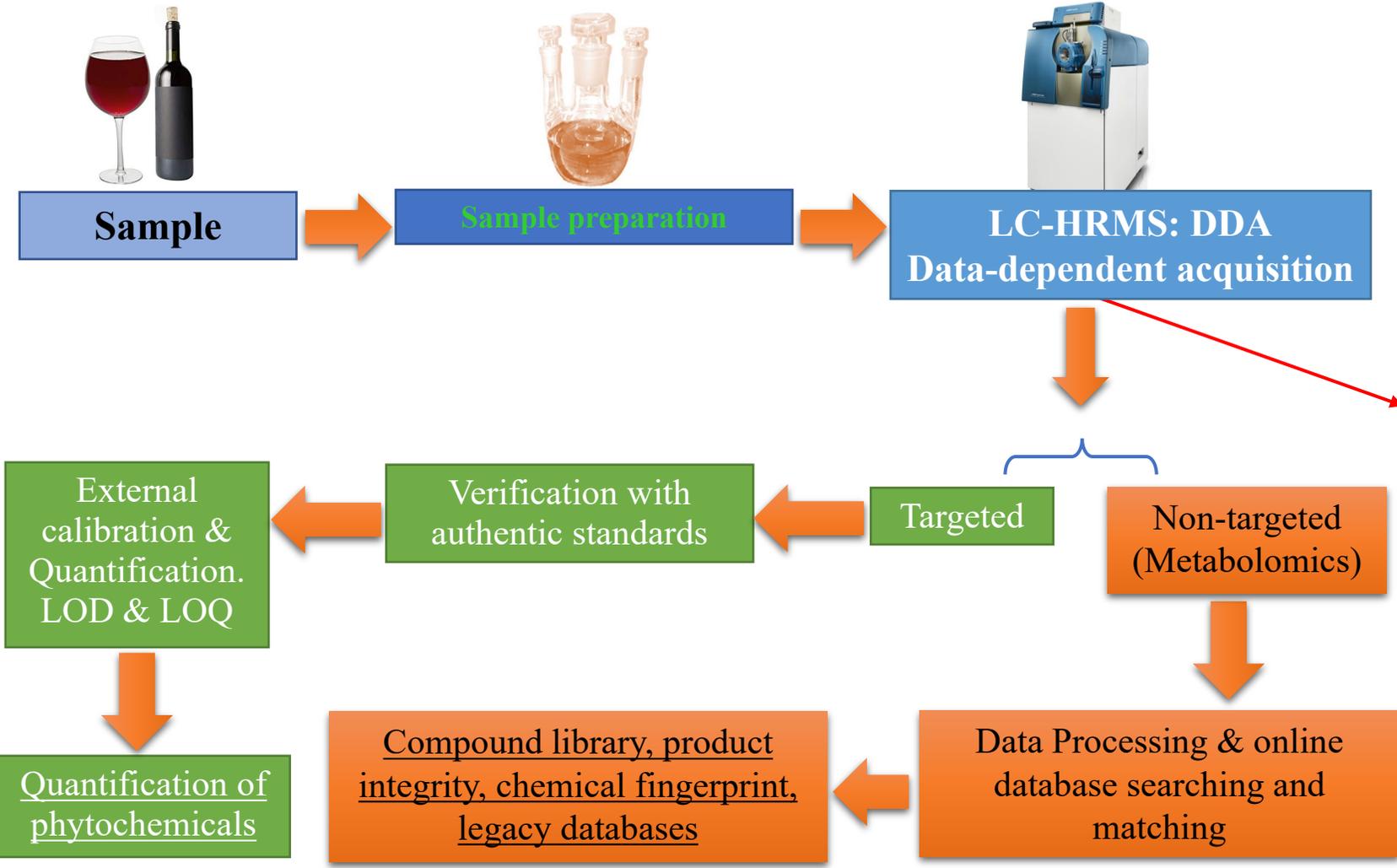
Major findings

- Smoke exposed red wine presents the biggest dispersion of the data. o-,m- and p-cresol exhibit the biggest difference across sample.
- Free guaiacol is highly correlated with free 4-methylguaiacol in smoke exposed wine
- Total guaiacol is highly correlated with free guaiacol
- m-cresol and o-cresol is highly correlated

Objective 2

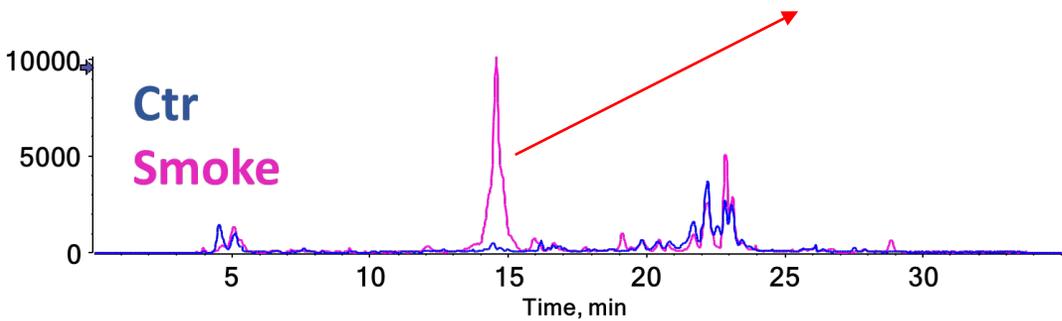
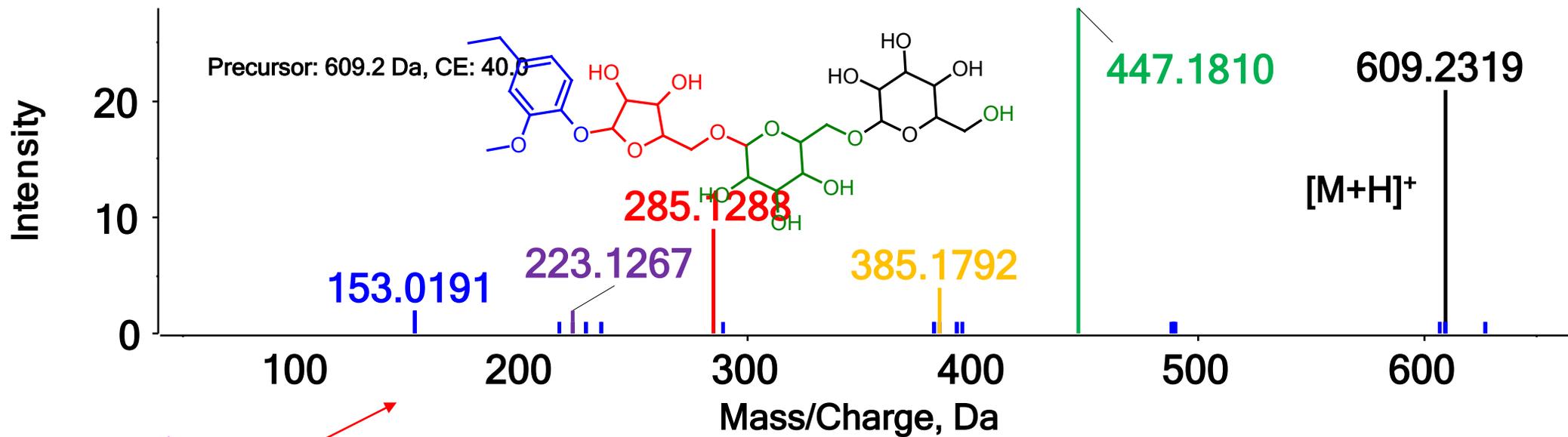
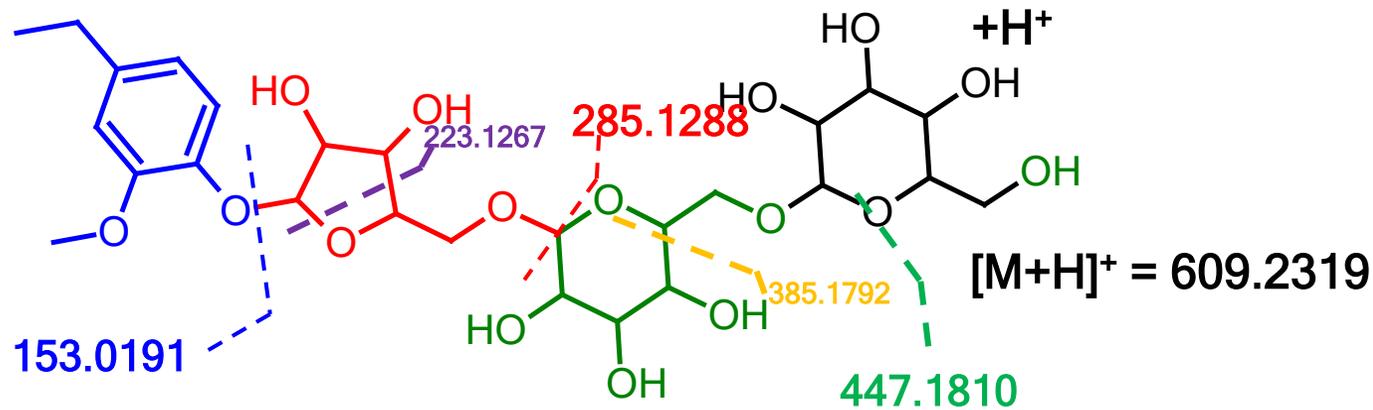
- Understand how the phenol glycosides are related to the volatile phenols
 - Selected 26 different degree of smoke-exposed wine and 14 control wines
 - LC-MS glycoside analysis
 - GC-MS analysis for free and total volatile phenols

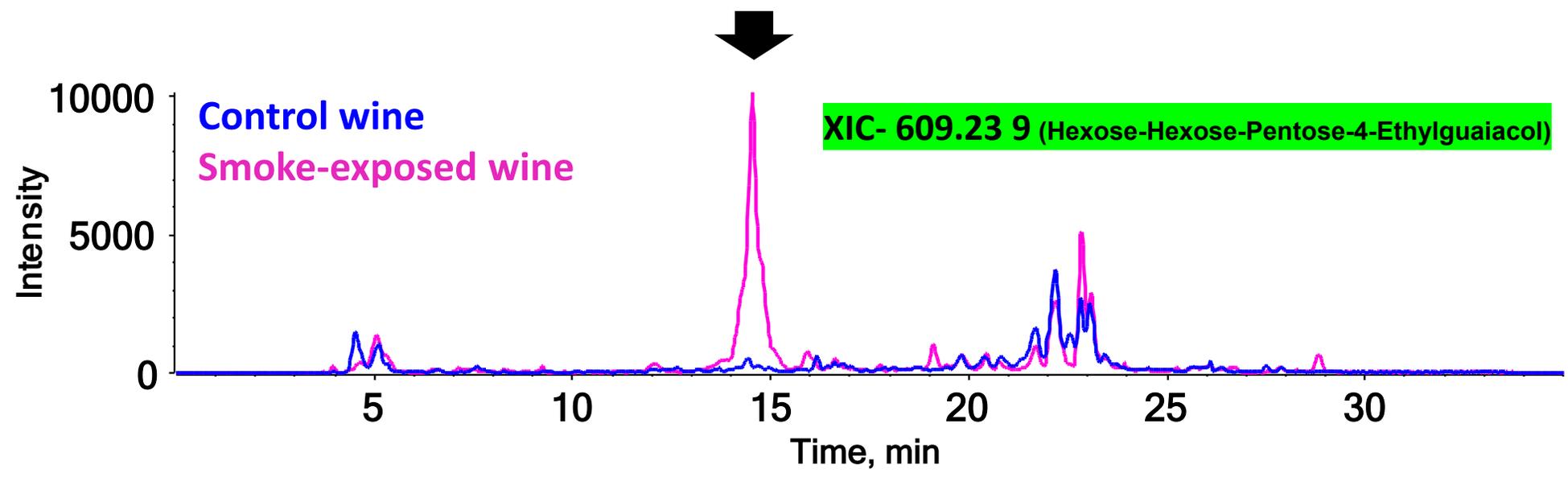
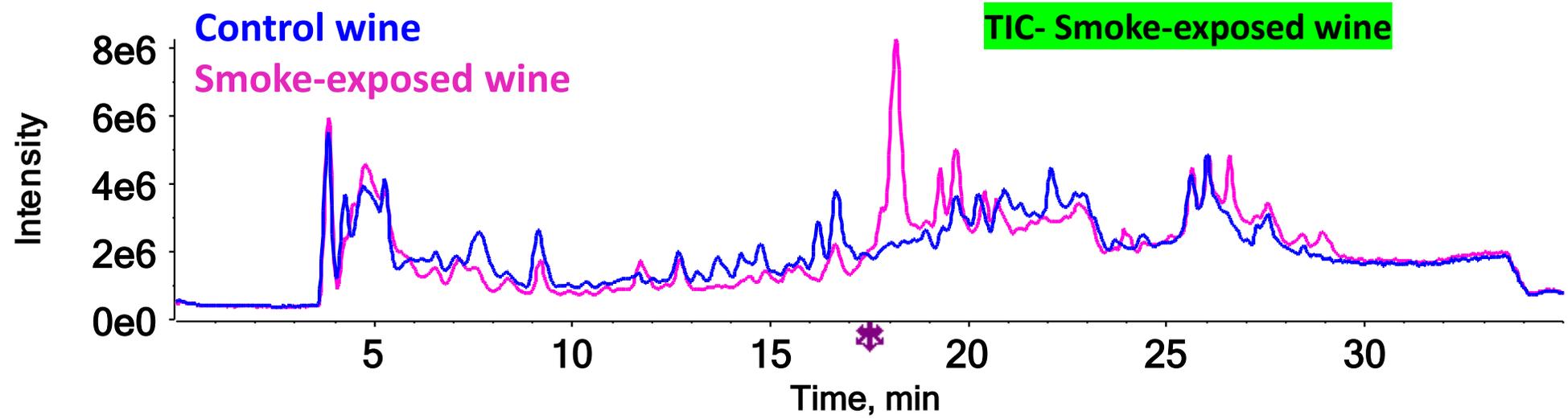
LC MS/MS Mass Spectral Fingerprinting and Quantification of Marker Compounds (14 control, 26 smoke exposed wine)



Hexose-Hexose-Pentose-4-Ethylguaiaicol

ESI+





Latest Metabolomics-Data Analysis in Progenesis Q1

File | Review Alignment | Experiment Design Setup | Peak Picking | Review Deconvolution | Identify Compounds | Review Compounds | Compound Statistics

Import Data | Review Alignment | Experiment Design Setup | Peak Picking | Review Deconvolution | Identify Compounds | Review Compounds | Compound Statistics

Identify Compounds
 Select your identification method:
 Progenesis MetaScope

Filter the compounds
 Using the list below, filter the compounds to show only those you want to identify.

Choose search parameters
 Select your MetaScope search parameters or create a new parameter set:
 Kegg_HMDB

Search for identifications
 Identifications will be assigned to the relevant compounds automatically.

No filter applied

| Compound | Accepted ID | Tag | Identifications | MS/MS | Neutral mass | m/z | Adducts |
|------------------------|-------------|-----|-----------------|-------|--------------|-----------|---------|
| 21.03_227.0702m/z | 44878 | | 55 | 10 | <unknown> | 227.0702 | 1 |
| 21.00_387.1633m/z | | | 0 | 0 | <unknown> | 387.1633 | 1 |
| 21.00_803.0623m/z | | | 0 | 0 | <unknown> | ----- | 1 |
| 21.00_579.1715m/z | | | 20 | 3 | <unknown> | | |
| 21.00_493.1350m/z | | | 0 | 0 | <unknown> | | |
| 21.00_411.3486m/z | | | 0 | 0 | <unknown> | | |
| 21.00_445.2074m/z | | | 0 | 0 | <unknown> | | |
| 21.03_647.1607m/z | | | 6 | 2 | <unknown> | | |
| 21.03_575.2722m/z | HMD80034679 | | 1 | 1 | <unknown> | | |
| 21.03_833.1616m/z | | | 14 | 2 | <unknown> | 833.1616 | 1 |
| 21.03_1140.7983m/z | | | 0 | 0 | <unknown> | 1140.7983 | 1 |
| 21.03_453.1768m/z | | | 0 | 0 | <unknown> | 453.1768 | 1 |
| 21.03_202.0854m/z | | | 0 | 0 | <unknown> | 20. | |
| 21.00_100.9328m/z | | | 0 | 0 | <unknown> | 100 | |
| 21.10_435.1090m/z | | | 15 | 4 | <unknown> | 43! | |
| 21.10_695.3670m/z | | | 2 | 5 | <unknown> | 69! | |
| 21.10_651.1716m/z | | | 2 | 2 | <unknown> | 65. | |
| 21.17_353.0852 | | | 5 | 1 | <unknown> | 353.0852 | 1 |
| 21.18_803.1866 | | | 3 | 1 | <unknown> | 803.1866 | 1 |
| 21.18_196.0373 | | | 3 | 1 | <unknown> | 196.0373 | 1 |
| 21.18_605.8872 | | | 3 | 1 | <unknown> | 605.8872 | 1 |
| 21.18_218.0524 | | | 4 | 1 | <unknown> | 218.0524 | 1 |
| 21.18_800.3073 | | | 3 | 1 | <unknown> | 800.3073 | 1 |
| 327 of 14999 compounds | | | | | | | |

Databases

09052019 Wine Neg - Progenesis... | 09052019 Wine POS - Progenesis...

PubChem

hmdb

Metlin

KEGG

Knapsack Family

Compound 21.03_227.0702m/z (cis-Resveratrol)

Legend: Matched fragment | Unmatched fragment

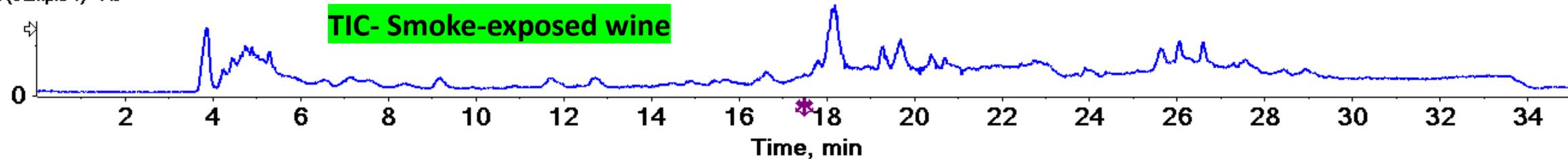
Possible identifications: 55

| Compound ID | Description | Adducts | Formula | Retention time | Score | Fragment |
|-------------|-------------------------------|--------------------|--|----------------|--------|----------|
| HMDB0039493 | Dihydroxyerone acid | M-H ₂ O | C ₁₄ H ₁₄ O ₄ | | C 53.1 | C 78.6 |
| 44878 | cis-Resveratrol | M-H | C ₁₄ H ₁₂ O ₃ | | B 53 | B 78.2 |
| HMDB0132692 | 8-[(3,3-dimethylloxiran-2-yl) | M-H ₂ O | C ₁₄ H ₁₄ O ₄ | | C 52.8 | C 77.3 |
| HMDB0033914 | Aegelinol | M-H ₂ O | C ₁₄ H ₁₄ O ₄ | | C 52.8 | C 76.5 |
| HMDB0136681 | 7-hydroxy-5-[(3-methylbut | M-H ₂ O | C ₁₄ H ₁₄ O ₄ | | C 52.6 | C 76.1 |
| HMDB0132701 | 7-hydroxy-8-(2-hydroxy-3- | M-H ₂ O | C ₁₄ H ₁₄ O ₄ | | C 52.5 | C 75.8 |
| HMDB0130516 | 7-hydroxy-6-(3-hydroxy-3- | M-H ₂ O | C ₁₄ H ₁₄ O ₄ | | C 52.3 | C 74.5 |
| 6979 | trans-Resveratrol | M-H | C ₁₄ H ₁₂ O ₃ | | B 52.1 | B 73.6 |
| HMDB0136947 | 6-[(E)-2-(4-hydroxyphenyl) | M-H ₂ O | C ₁₄ H ₁₄ O ₄ | | C 51.7 | C 71.7 |
| HMDB0034270 | 5,6-Dihydro-5-hydroxy-6-r | M-H | C ₁₄ H ₁₂ O ₃ | | C 50.2 | C 64 |
| HMDB0015575 | Trioxalen | M-H | C ₁₄ H ₁₂ O ₃ | | C 49.5 | C 60.8 |
| HMDB0127763 | 5-[(3-methoxyphenyl)meth | M+Na... | C ₁₂ H ₁₄ O ₃ | | C 49.5 | C 59.5 |
| HMDB0030786 | Marmesin | M-H ₂ O | C ₁₄ H ₁₄ O ₄ | | C 48 | C 53.2 |
| HMDB0034384 | Corticocin | M-H ₂ O | C ₁₄ H ₁₄ O ₄ | | C 47.5 | C 50.6 |
| HMDB0034257 | Seselin | M-H | C ₁₄ H ₁₂ O ₃ | | C 47.3 | C 49.6 |
| HMDB0130827 | 6-(2-hydroxybut-1-en-1-yl) | M-H ₂ O | C ₁₄ H ₁₄ O ₄ | | C 47.1 | C 49.6 |

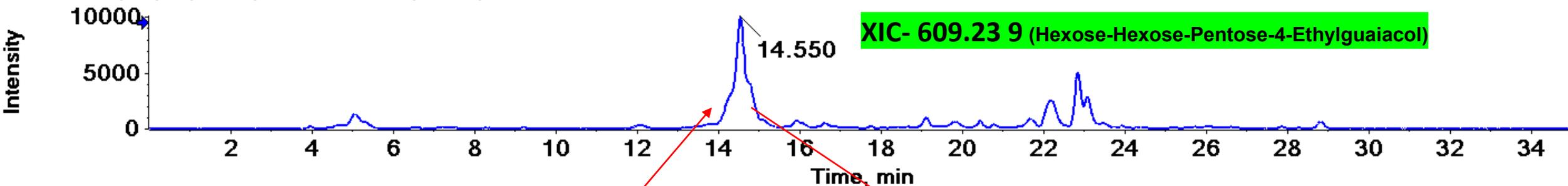
Section Complete

Phenol-Explorer
 Database on polyphenol content in foods

TIC from R5.wiff (sample 1) - R5



XIC from R5.wiff (sample 1) - R5, Experiment 1, +TOF MS (60 - 1500): 609.240 +/- 0.040 Da, Gaussian smoothed

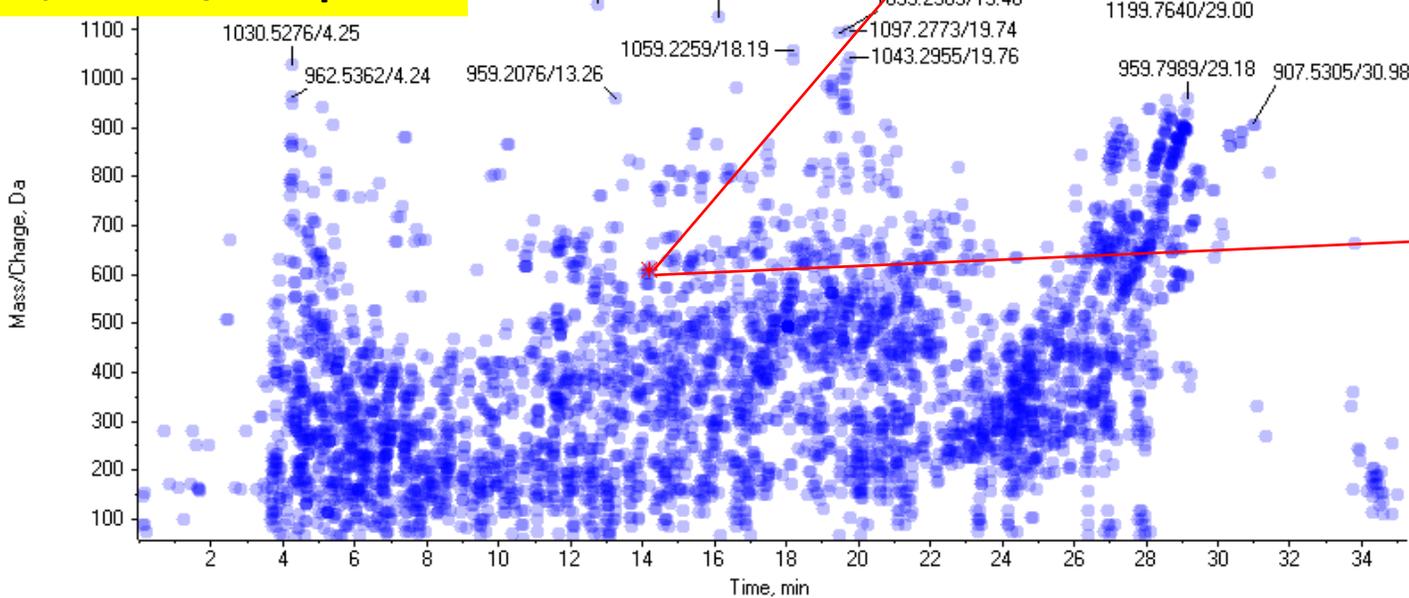


Filtering Controls

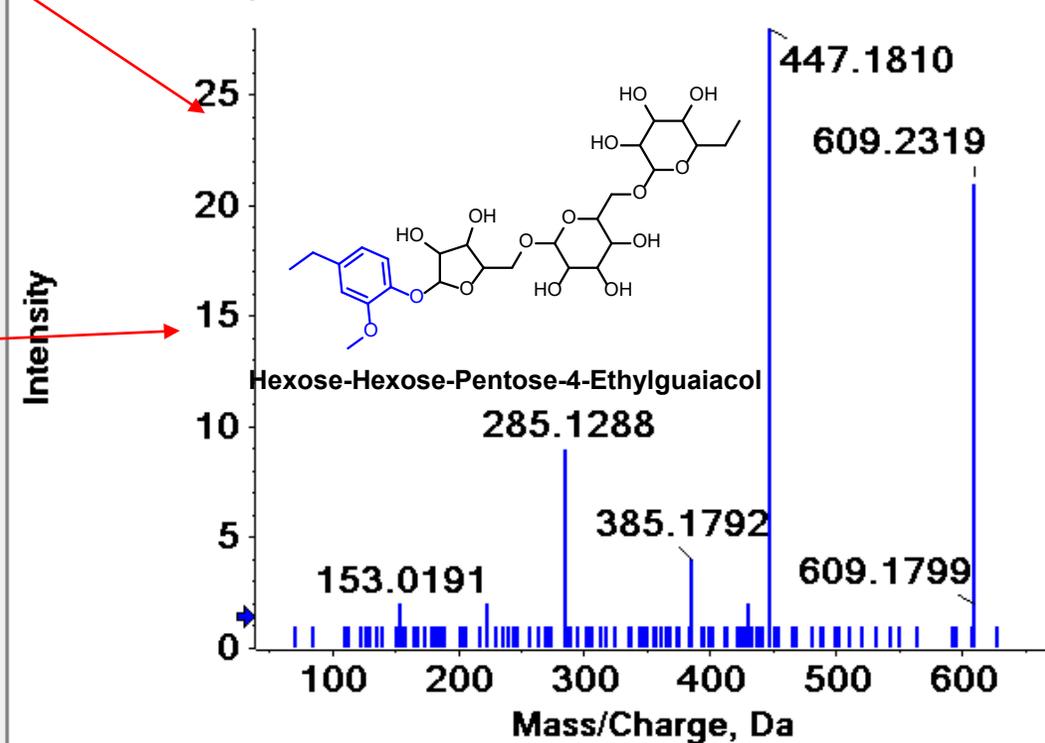
Graph Table

Time versus Precursor Mass/Charge for IDA Dependents

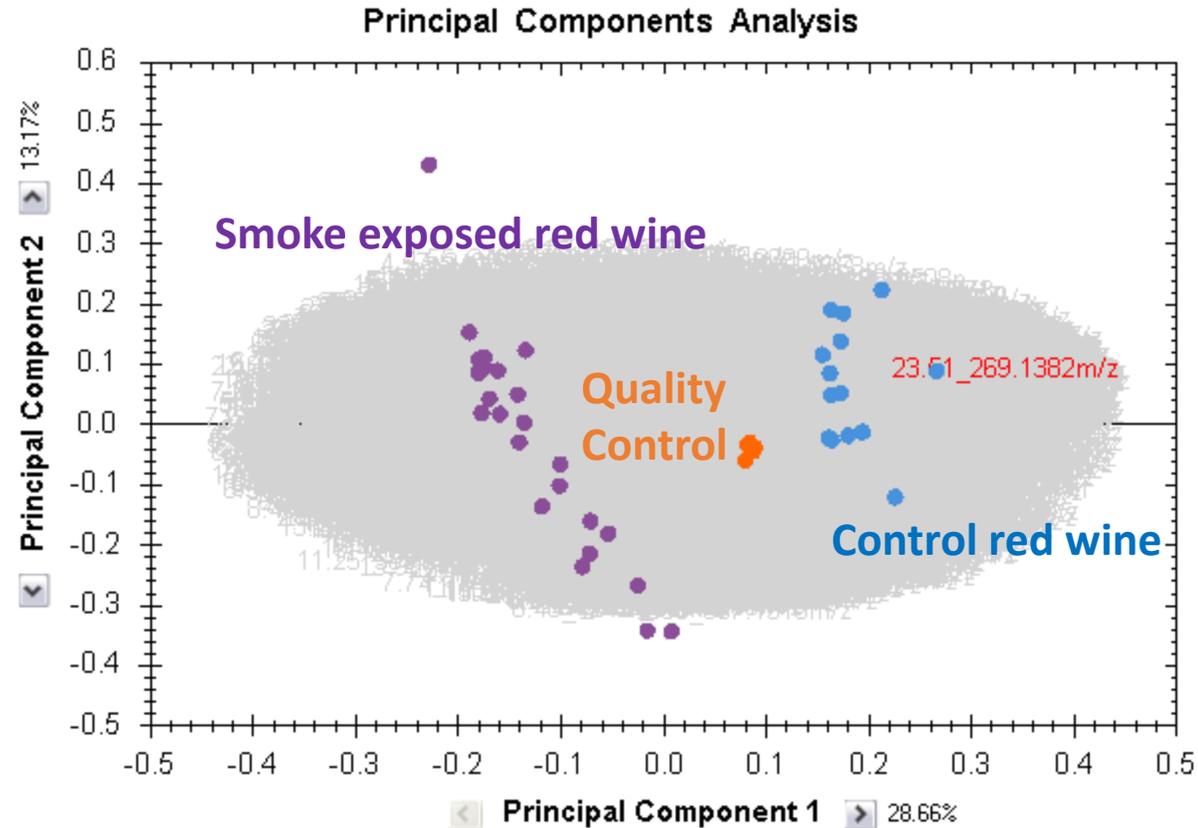
4,135 MS/MS spectra



Spectrum from R5.wiff (sample 1) - R5, Exper...ent 2, +TOF MS² (40 - 1000) from 14.188 min
Precursor: 609.2 Da, CE: 40.0



LC-HRMS/MS Untargeted Metabolomics Analysis of Wine (14 control, 26 smoke exposed wine)



We obtained more than 4000 ions by MS/MS

Control wines and smoke exposed wines are well clustered by PCA analysis

We are continuing with data mining to figure out underlying relationship

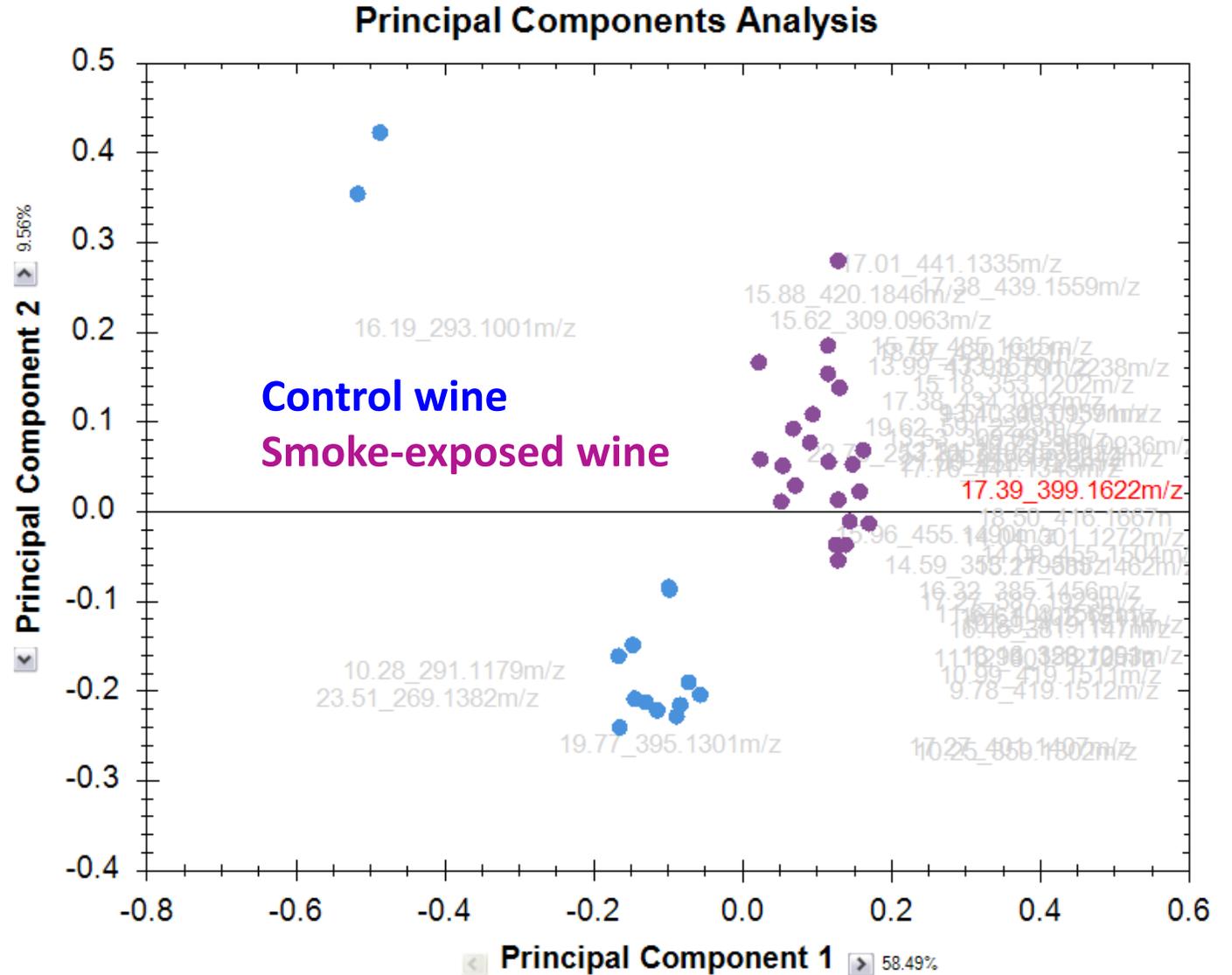
LC-HRMS/MS targeted phenol glycosides Analysis

35 tentative annotated volatile phenol glycosides

| Accepted Compound ID | Adducts detected | Formula | Mass Error (ppm) |
|----------------------------------|------------------|-----------|------------------|
| syringyl-β-D-glucopyranoside_Iso | M+H, M+NH4, M+Na | C14H20O8 | -3.14 |
| syringyl-β-D-glucopyranoside | M+NH4, M+Na | C14H20O8 | -2.24 |
| DeoxyH-H-Cresol | M+NH4, M+Na | C19H28O10 | -3.50 |
| DeoxyH-H-Cresol_iso | M+Na | C19H28O10 | -3.69 |
| DeoxyH-H-P-Phenol | M+Na | C23H34O14 | -4.10 |
| DeoxyH-P-4-Ethylguaiacol | M+Na | C20H30O10 | -1.93 |
| guaiacyl-β-D-gentiobioside -Der | M+NH4 | C19H28O12 | -4.50 |
| guaiacyl-β-D-gentiobioside | M+Na | C19H28O12 | -4.12 |
| guaiacyl-β-D-gentiobioside_2 | M+Na | C19H28O12 | -4.11 |
| guaiacyl-β-D-gentiobioside_3 | M+Na | C19H28O12 | 7.61 |
| H-4-Methylguaiacol | M+NH4 | C14H20O7 | -3.68 |
| H-4-Methylguaiacol_2 | M+H | C14H20O7 | -3.18 |
| H-4-Methylsyringol | M+Na | C15H22O8 | -1.53 |
| H-Guaiacol | M+Na | C13H18O7 | -2.00 |
| H-Guaiacol_2 | M+Na | C13H18O7 | -3.35 |
| H-Guaiacol_3 | M+Na | C13H18O7 | -3.56 |
| H-H-P-4-Ethylguaiacol_Iso1 | M+H-H2O | C24H36O14 | -7.46 |
| H-H-P-4-Ethylguaiacol | M+H | C24H36O14 | -7.68 |
| H-H-P-4-Ethylguaiacol_iso2 | M+H-H2O | C24H36O14 | -9.13 |
| HMDB0041514 | M+NH4 | C18H26O10 | -4.49 |
| H-P-4-Methylguaiacol_iso | M+Na | C19H28O11 | -4.63 |
| H-P-4-Methylguaiacol | M+H | C19H28O11 | -7.97 |
| H-P-Guaiacol | M+Na | C18H26O11 | -7.79 |
| H-P-Guaiacol_Iso1 | M+Na | C18H26O11 | -5.29 |
| H-P-Guaiacol_so2 | M+Na | C18H26O11 | -3.63 |
| H-P-Guaiacol_iso3 | M+Na | C18H26O11 | -4.93 |
| H-P-P-4-Methylguaiacol | M+Na | C24H36O15 | -4.72 |
| P-H-Cresol | M+H | C18H26O10 | -6.84 |
| P-H-Cresol_iso2 | M+H-H2O | C18H26O10 | -7.78 |
| P-H-Cresol_iso3 | M+H-H2O | C18H26O10 | -9.25 |
| P-H-Cresol_iso4 | M+H, M+NH4, M+Na | C18H26O10 | -3.61 |
| P-P-H-Cresol | M+Na | C23H34O14 | -2.80 |
| syringyl-β-D-gentiobioside | M+NH4 | C20H30O13 | -4.53 |
| syringyl-β-D-gentiobioside_iso1 | M+H, M+Na | C20H30O13 | -8.89 |
| syringyl-β-D-gentiobiosideliso2 | M+H-H2O | C20H30O13 | -8.01 |

H-Hexose, P- Pentose, Iso- Isomer

PCA- Using 35 tentative annotated volatile phenol glycosides



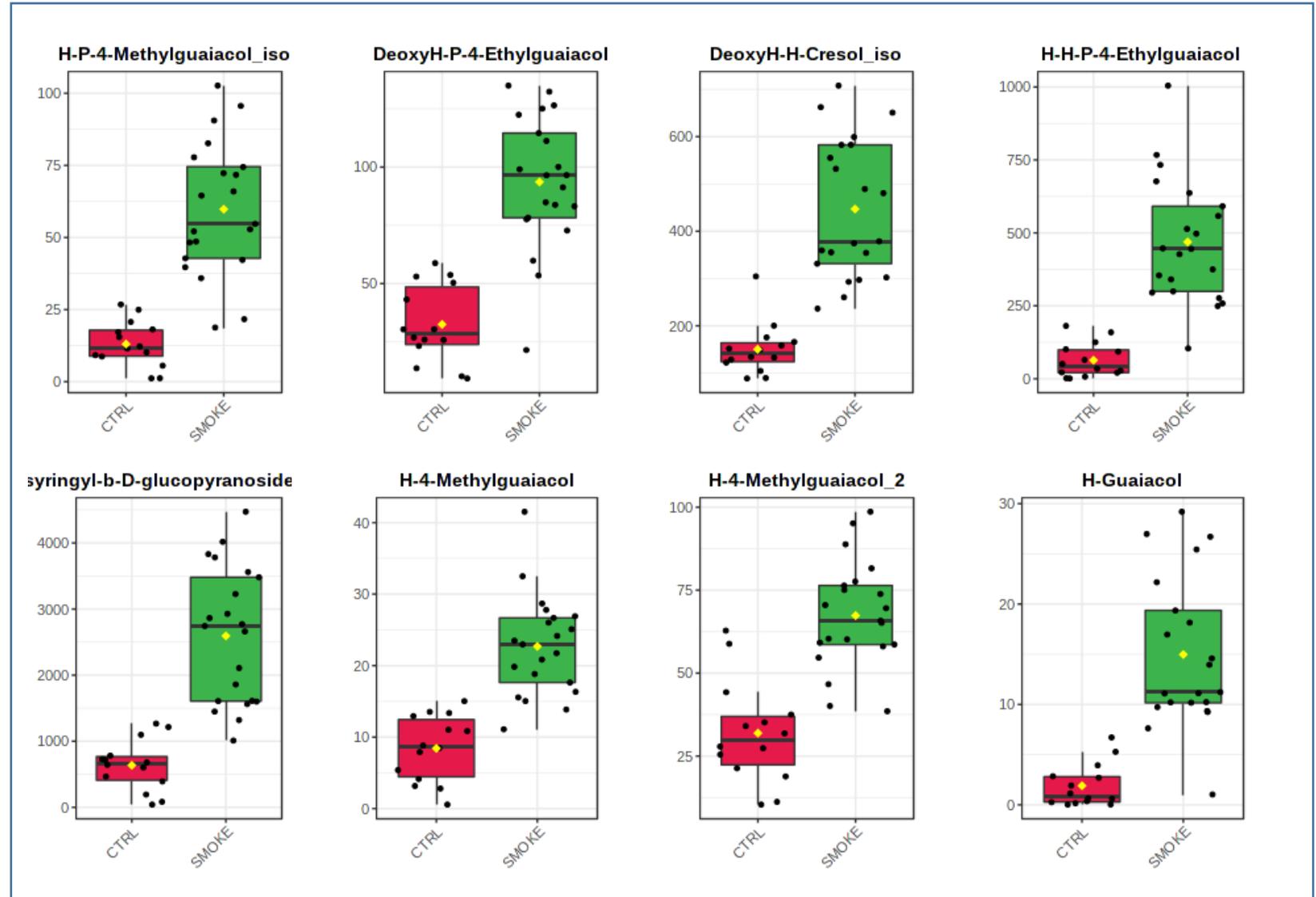
T-test

Top 8, volatile phenol glycosides ($p < 4 \cdot 10^{-7}$)

Top 15, volatile phenol glycosides

| Tentative glycoside | FC | SMK/CTR raw.pval |
|----------------------------------|-------|------------------|
| H-P-4-Methylguaiacol_iso | 4.58 | 2.37E-08 |
| DeoxyH-P-4-Ethylguaiacol | 2.88 | 2.59E-08 |
| DeoxyH-H-Cresol_iso | 2.97 | 3.02E-08 |
| H-H-P-4-Ethylguaiacol | 7.39 | 6.61E-08 |
| syringyl-b-D-glucopyranoside_Iso | 4.08 | 9.50E-08 |
| H-4-Methylguaiacol | 2.69 | 1.26E-07 |
| H-4-Methylguaiacol_iso1 | 2.11 | 2.52E-07 |
| H-Guaiacol | 7.88 | 3.95E-07 |
| P-H-Cresol_iso2 | 3.90 | 5.13E-07 |
| syringyl-b-D-gentiobioside_iso1 | 2.60 | 8.96E-07 |
| H-H-P-4-Ethylguaiacol_Iso1 | 31.88 | 2.21E-05 |
| P-H-Cresol_iso3 | 3.52 | 3.23E-05 |
| H-P-4-Methylguaiacol | 2.66 | 4.37E-05 |
| DeoxyH-H-Cresol | 3.13 | 5.73E-05 |
| syringyl-b-D-gentiobioside | 5.65 | 8.64E-05 |

H-Hexose, P- Pentose, Iso- Isomer



Take home message from LC-MS glycoside analysis

- With very few exceptions, volatiles measured by GC-MS and phenol glycosides by LC-MS were positively correlated
- This means higher concentration of volatile phenols predict higher concentrations of phenol glycosides and vice-versa.

Objective 3

- How bad is my wine if exposed to smoke?

Smoke exposed \neq Smoke taint

Machine learning and modeling for smoke evaluation and prediction

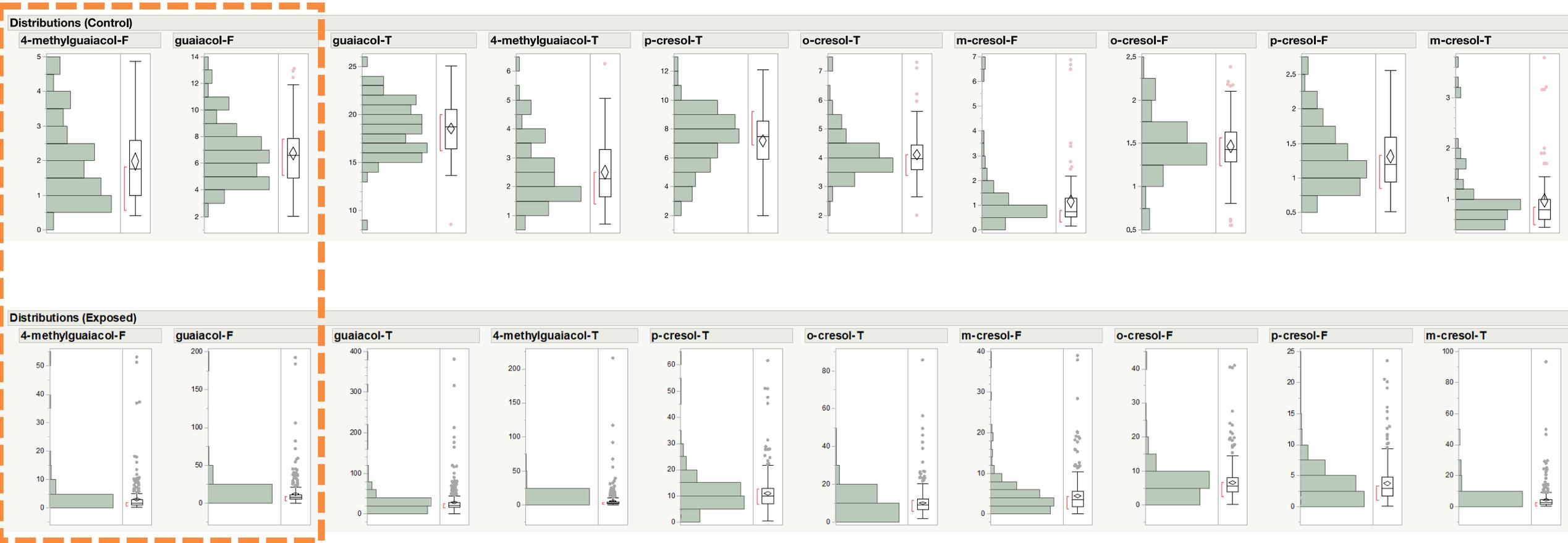
- 86 control wines from 2013-2016
- 377 wines from 2020 (exposed)

- Free and total volatile phenols were used as markers

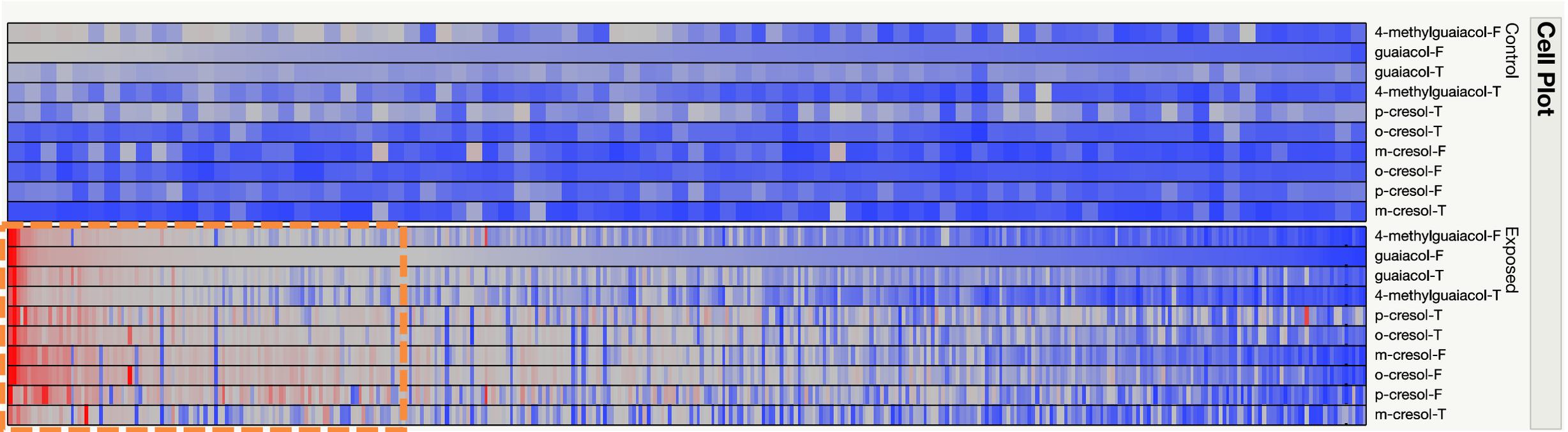


Dr. Ye Feng, Former director of Machine-learning at Lam Research

Sample Distribution

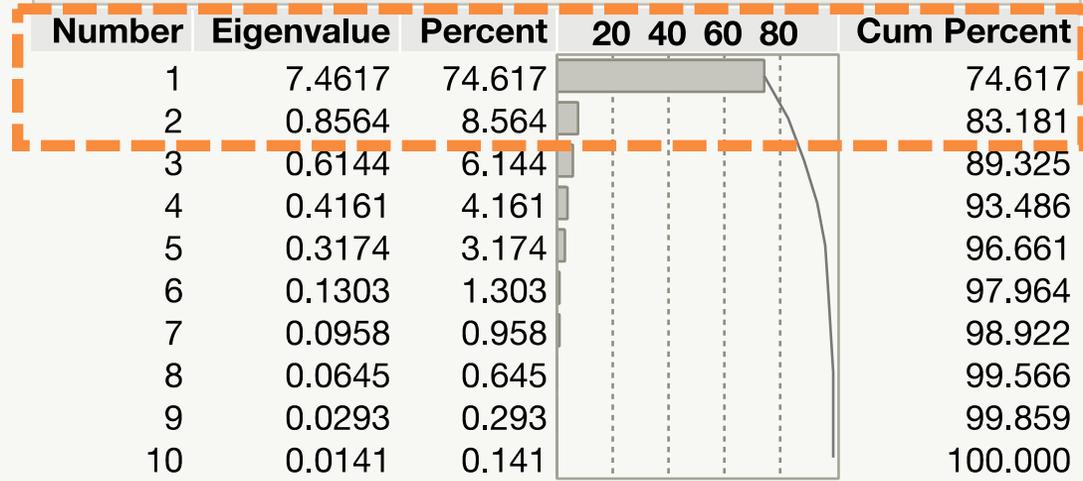


Multivariate Sample Distribution

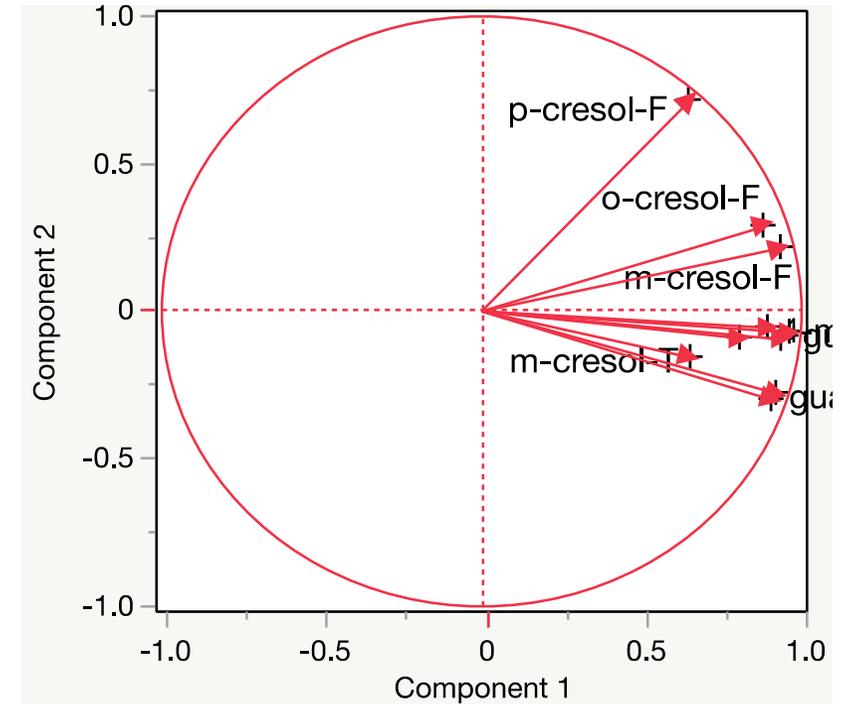


Principal Component Analysis

Principal Components: on Correlations



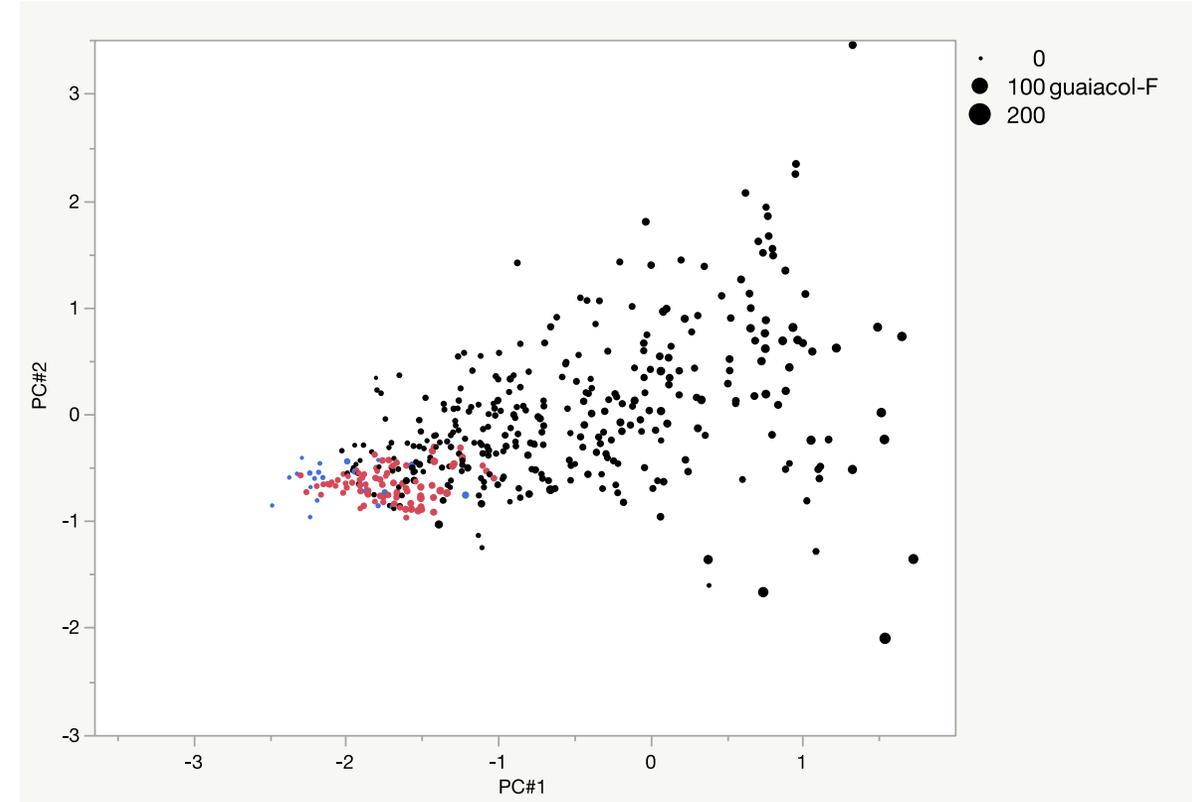
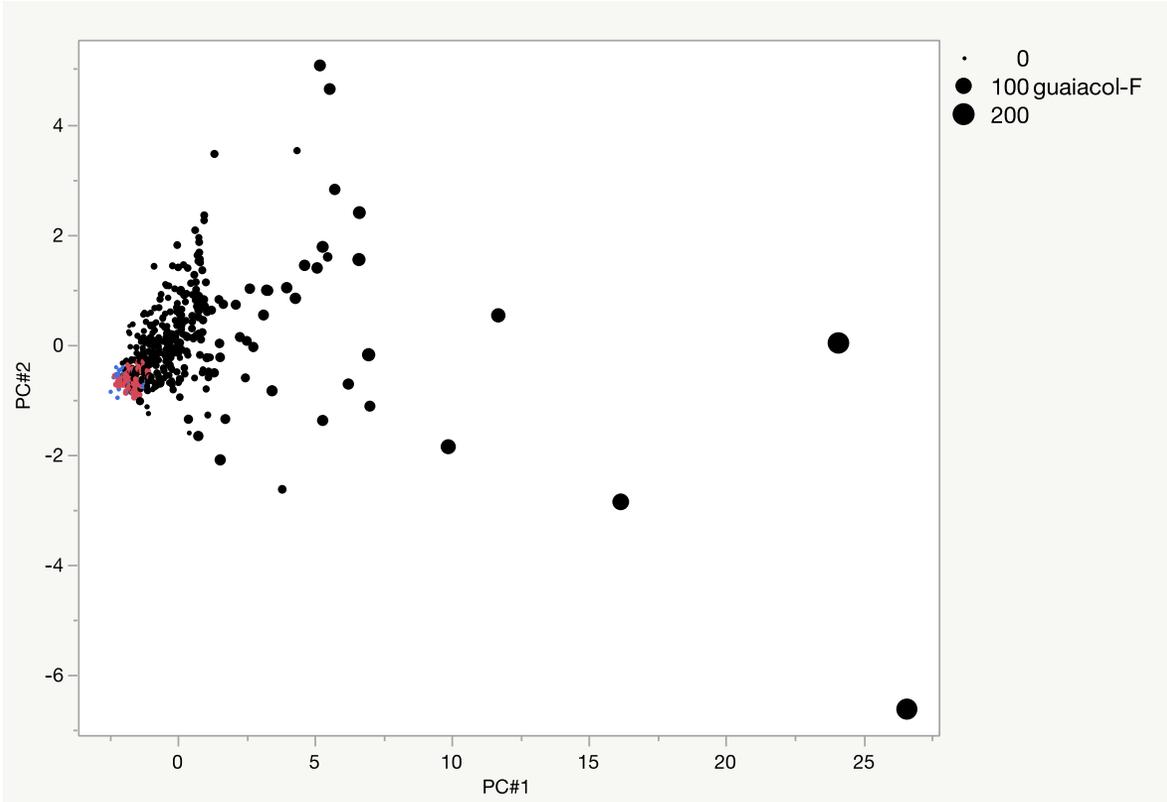
| Eigenvectors | PC#1 | PC#2 |
|--------------------|---------|----------|
| 4-methylguaiacol-F | 0.32844 | -0.06079 |
| guaiacol-F | 0.35343 | -0.07758 |
| guaiacol-T | 0.33792 | -0.30178 |
| 4-methylguaiacol-T | 0.33257 | -0.32487 |
| p-cresol-T | 0.29652 | -0.09926 |
| o-cresol-T | 0.34353 | -0.10365 |
| m-cresol-F | 0.34308 | 0.23174 |
| o-cresol-F | 0.32346 | 0.31042 |
| p-cresol-F | 0.23796 | 0.77070 |
| m-cresol-T | 0.23985 | -0.17038 |



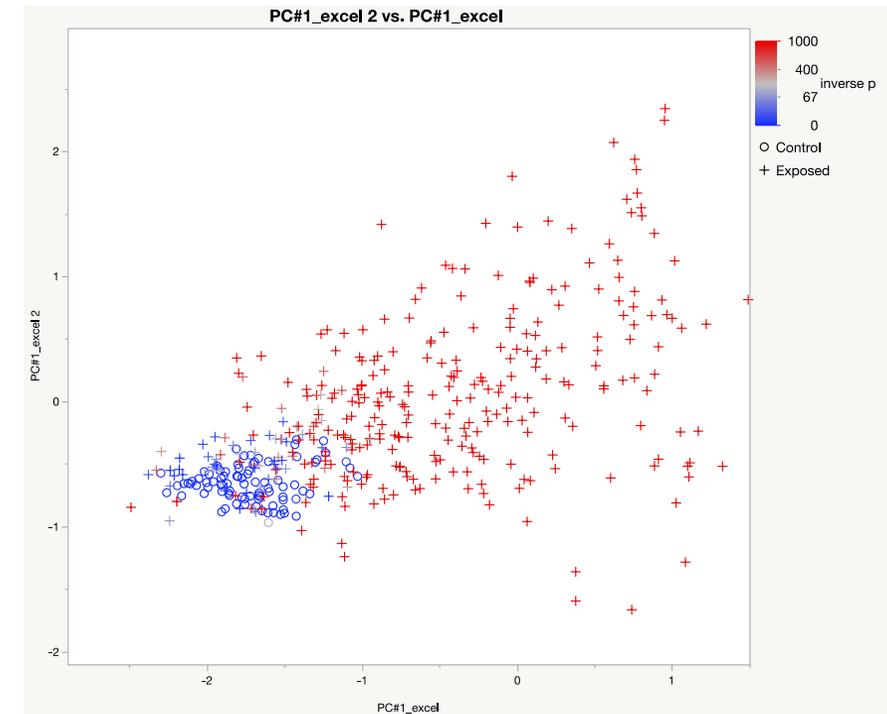
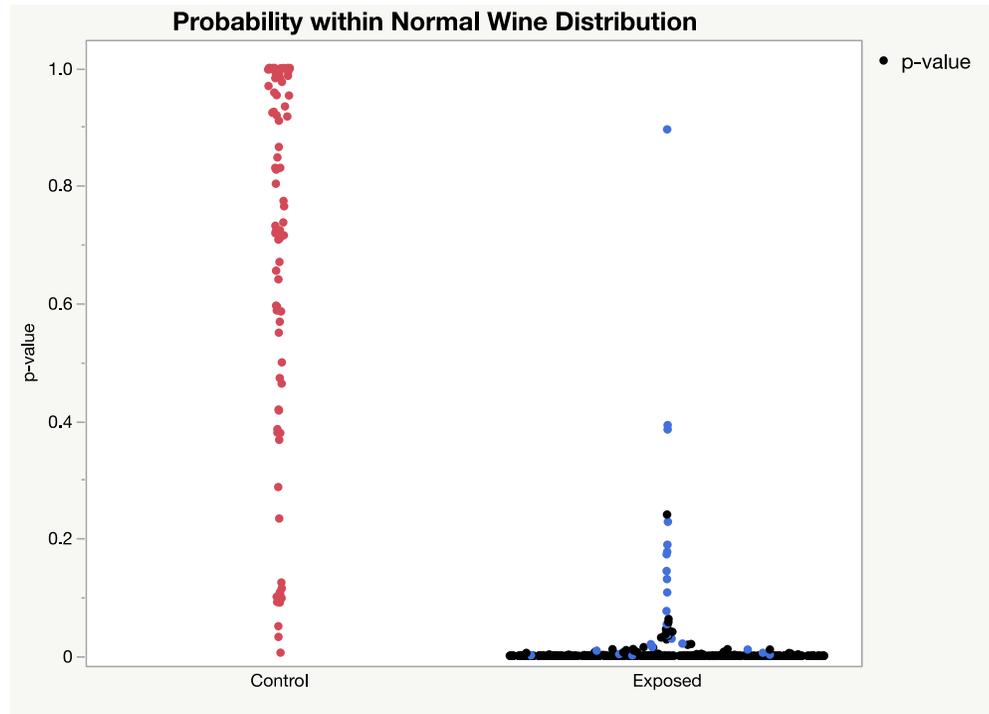
The first principal component (PC) represents overall phenolic changes

The second PC points to unique covariant fingerprints in free cresols

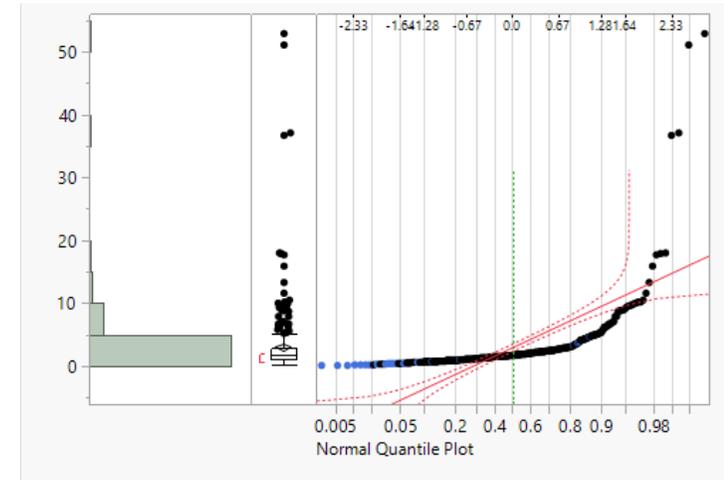
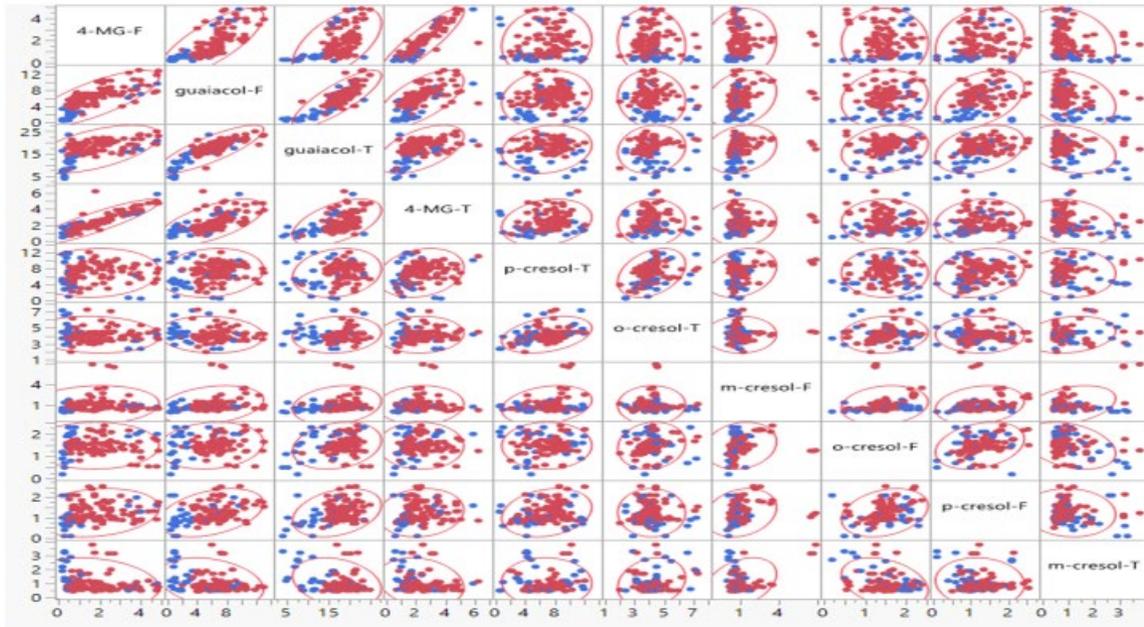
Evaluation of 2020 Wines (sized by Guaiacol-F)



Evaluation of Wines (sized by p-value)



Correlation and modeling



Model prediction: p-value

| | | | | | | | | | | | | | | | | | |
|-------------|-------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|-------------|
| 3.91666667 | 6.592592593 | -1.33112 | -0.66458 | -0.77733 | -2.05628 | -0.74186 | 0.194477 | 4.911919 | 1.888081 | 0.544012 | 2.135581 | -0.31035 | 0.621221 | 3.411382 | 77.95991 | 6.960706 | 0.001055611 |
| 2.972222222 | 5.682926829 | -0.8037 | -0.3853 | 1.622674 | 3.943721 | 4.75814 | 1.594477 | 2.511919 | 2.988081 | 1.444012 | 2.935581 | 0.889651 | 0.621221 | 1.743567 | 94.58311 | 8.444921 | 0.000450617 |
| 5.2 | 6.78125 | -0.71179 | -0.16982 | -0.47733 | 1.043721 | 3.15814 | 0.694477 | 1.211919 | 4.888081 | 2.744012 | 4.635581 | 0.989651 | 2.721221 | 0.646487 | 265.8569 | 23.73722 | 3.65367E-06 |
| 4.875 | 6.12195122 | -0.31389 | -0.69874 | -0.37733 | 1.043721 | 6.55814 | 1.594477 | 9.711919 | 5.588081 | 2.544012 | 4.035581 | -0.01035 | 4.621221 | 5.480444 | 385.4543 | 34.41557 | 6.06225E-07 |
| 3.857142857 | 4.612244898 | -1.12923 | -1.13684 | -1.27733 | -4.05628 | 4.05814 | 2.394477 | 7.311919 | 4.788081 | 0.144012 | 0.035581 | -0.61035 | 4.421221 | 4.537723 | 154.4551 | 13.79064 | 4.79318E-05 |
| 4.631578947 | 5.42 | 0.016748 | -0.69662 | -0.07733 | 2.043721 | 8.55814 | 2.494477 | 13.31192 | 5.988081 | 2.744012 | 4.735581 | 0.189651 | 4.521221 | 11.03934 | 494.8929 | 44.18686 | 1.79067E-07 |
| 4.722222222 | 4.885245902 | -0.21436 | -0.73772 | -0.17733 | 1.743721 | 11.25814 | 3.594477 | 8.811919 | 5.788081 | 2.544012 | 3.935581 | 0.189651 | 4.621221 | 4.108217 | 367.3541 | 32.79947 | 7.65886E-07 |
| 4.882352941 | 5 | -0.52074 | -0.4816 | -0.27733 | 1.543721 | 3.95814 | 1.994477 | 6.311919 | 4.588081 | 2.844012 | 4.135581 | 0.289651 | 2.721221 | 3.089621 | 263.8451 | 23.5576 | 3.78942E-06 |
| 4.75 | 5.28 | -0.24115 | -0.43869 | 0.022674 | 2.743721 | 7.85814 | 2.494477 | 8.011919 | 5.088081 | 3.044012 | 5.135581 | 0.489651 | 2.421221 | 5.507082 | 363.5363 | 32.4586 | 8.05747E-07 |
| 4.642857143 | 5.272727273 | -0.86358 | -0.69621 | -0.57733 | -0.25628 | 4.65814 | 1.894477 | 3.811919 | 4.188081 | 1.744012 | 3.135581 | -0.21035 | 3.521221 | 1.490751 | 218.5799 | 19.51606 | 9.32128E-06 |
| 4.736842105 | 5.255319149 | -0.53183 | -0.43205 | -0.07733 | 2.243721 | 6.15814 | 2.194477 | 1.511919 | 6.288081 | 2.844012 | 5.135581 | 0.189651 | 2.921221 | 1.43106 | 362.7333 | 32.3869 | 8.14445E-07 |
| 5.0625 | 6.177777778 | -0.62515 | -0.69753 | -0.37733 | 1.343721 | 9.25814 | 1.994477 | 0.911919 | 5.488081 | 2.444012 | 4.335581 | -0.31035 | 6.021221 | 2.103738 | 466.2665 | 41.63094 | 2.39647E-07 |
| 4.666666667 | 5.048387097 | -0.03815 | -0.5024 | 0.122674 | 3.043721 | 12.75814 | 3.694477 | 5.411919 | 7.188081 | 3.544012 | 4.935581 | 0.689651 | 5.021221 | 1.620323 | 436.5137 | 38.97444 | 3.3069E-07 |
| 5 | 5.826086957 | -0.32814 | -0.27774 | -0.17733 | 2.243721 | 8.25814 | 2.094477 | 3.211919 | 5.988081 | 3.244012 | 4.835581 | 1.089651 | 4.421221 | 0.78918 | 361.4582 | 32.27305 | 8.2849E-07 |
| 4.5 | 8.05 | -1.98962 | -0.55526 | -1.17733 | -3.15628 | -2.44186 | -0.50552 | -4.78808 | 0.688081 | 0.144012 | 1.635581 | -0.51035 | 0.121221 | 2.593103 | 43.90425 | 3.920022 | 0.011311237 |
| 2.909090909 | 8.625 | -1.95435 | -0.52808 | -0.87733 | -3.55628 | -4.74186 | -0.90552 | -3.58808 | 0.588081 | 0.344012 | 1.635581 | -0.51035 | -0.07878 | 1.991242 | 43.70191 | 3.901956 | 0.01151535 |
| 3.702970297 | 3.724637681 | 3.431154 | -0.84644 | 8.122674 | 30.64372 | 58.55814 | 18.19448 | 11.91192 | 13.08808 | 7.644012 | 9.535581 | 2.589651 | 12.82122 | 7.701365 | 2638.31 | 235.5634 | 4.53383E-11 |
| 2.363636364 | 6.291666667 | -1.78698 | -0.85924 | -0.87733 | -4.15628 | -3.44186 | -0.10552 | 3.511919 | 0.388081 | -0.45599 | 0.135581 | -0.71035 | 0.621221 | 2.938962 | 18.73934 | 1.673155 | 0.188994681 |
| 3.333333333 | 6.074074074 | -1.69191 | -0.53215 | -0.77733 | -2.75628 | -2.14186 | 0.194477 | 0.211919 | 0.888081 | 0.044012 | 1.035581 | 0.189651 | 0.421221 | 1.235536 | 17.03465 | 1.520951 | 0.240265045 |
| 5 | 5.057692308 | -0.40944 | -0.56437 | 0.122674 | 3.743721 | 7.75814 | 2.694477 | 6.011919 | 6.088081 | 1.044012 | 5.435581 | 0.289651 | 2.221221 | 4.485602 | 430.0916 | 38.40103 | 3.55489E-07 |
| 4.35 | 5.510638298 | -0.66816 | -0.62343 | 0.022674 | 1.943721 | 7.35814 | 2.194477 | 5.411919 | 5.188081 | 1.344012 | 3.335581 | 0.289651 | 1.721221 | 1.386922 | 162.2397 | 14.48569 | 3.80933E-05 |
| 4.791666667 | 5.642857143 | 0.23035 | -0.42914 | 0.422674 | 4.743721 | 13.05814 | 3.094477 | 8.811919 | 7.188081 | 3.744012 | 6.635581 | 0.589651 | 3.721221 | 6.057231 | 624.7588 | 55.78204 | 5.70944E-08 |
| 3.916666667 | 6.567567568 | -0.76448 | -0.28099 | 0.422674 | 2.643721 | 5.75814 | 1.194477 | -1.98808 | 5.088081 | 2.944012 | 4.835581 | 0.389651 | 1.721221 | 2.647964 | 267.3275 | 23.86853 | 3.55812E-06 |



- The model needs to be verified with sensory evaluation
- So the p-value should only be used as a reference!

Acknowledgement

- American Vineyard Foundation
- Oregon Wine Board
- Oregon Wine Research Institute

