Supplemental Information for

Combination of Two Analytical Techniques Improves Wine Classification by Vineyard, Region, and Vintage

Alexandra A. Crook,[a, †] Diana Zamora-Olivares,[c,d,†] Fatema Bhinderwala,[a,b,f†] Jade Woods,[a] Michelle Winkler,[d] Sebastian Rivera,[d] Cassandra E. Shannon,[d] Holden R. Wagner,[d] Deborah L. Zhuang, [d] Jessica E. Lynch,[d] Nathan R. Berryhill, [d] Ron C. Runnebaum,[e],* Eric Anslyn,[c,]* and Robert Powers,[a,b,]*

[a] Department of Chemistry, University of Nebraska-Lincoln, Lincoln NE 65888, United States

[b] Nebraska Center for Integrated Biomolecular Communication, University of Nebraska-Lincoln, Lincoln NE 68588, United States

[c] Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States

[d] Texas Institute for Discovery Education in Science and Freshman Research Initiative, The University of Texas at Austin, Austin, Texas 78712, United States.

[e] Department of Viticulture and Enology, and Department of Chemical Engineering, University of California−Davis, Davis, California 95616, United States

[f] Current Affiliation: Department of Structural Biology, University of Pittsburgh School of Medicine, 3501 Fifth Avenue, Pittsburgh, PA 15261, United States of America

[†]equal contribution

*To whom correspondence should be addressed:

Robert Powers
University of Nebraska-Lincoln
Department of Chemistry
722 Hamilton Hall
Lincoln, NE 68588-0304
Email: rpowers3@unl.edu
Phone: (402) 472-3039
Fax: (402) 472-9402
ORCID: 0000-0001-9948-6837

Eric Anslyn
The University of Texas at Austin
Department of Chemistry
Austin, Texas 76712
Email: anslyn@austin.utexas.edu
Phone: 512-471-0068
Fax: (512) 471-7791
ORCID: 0000-0002-5137-8797

Ron Runnebaum
University of California–Davis
Department of Viticulture & Enology
Department of Chemical Engineering
Davis, CA 95616
Email: rcrunnebaum@ucdavis.edu
Phone: (530) 752-9078
ORCID: 0000-0001-5872-8596

Contents

Tables:

Figure:

Table S1. Binding Ratios of the DS Array Ensembles.

Sensor Code	Sensing ensemble	Indicator (mM)	Binding ratio [Peptide:Metal:Indicator]
MM1	RN8-Cu^{2+}-PCV	0.075	0.5:1:1
MM2	TT2-Cu^{2+}- PCV	0.075	0.5:1:1
MM3	SEL1-Cu^{2+}- PCV	0.075	1:1:1
MM4	RN8-Cu^{2+}-CAS	0.06	0.4:1:1
MM5	TT2-Cu^{2+}- CAS	0.06	0.4:1:1
MM6	SEL1-Cu^{2+}- CAS	0.06	0.5:1:1
MM7	RN8-Ni^{2+}- BPR	0.018	1:1:1
MM8	TT2-Ni^{2+}- BPR	0.018	0.4:1:1
MM9	SEL1-Ni^{2+}- BPR	0.018	0.75:1:1

Table S2. 2015 Vineyard MB-PCA p-value Matrix

	SMV2	SRH1	AS1	AS2	SNC1	SNC2	CRN1	RRV1	RRV2	RRV3	AV1	AV2	OR1	OR2
SMV1	2.84E-10	1.68E-08	7.15E-03	9.27E-10	1.00E-07	1.74E-09	5.04E-07	4.20E-07	5.78E-10	1.13E-08	5.30E-03	4.23E-09	1.78E-08	5.68E-07
SMV2		4.21E-04	1.28E-07	9.53E-06	6.47E-11	8.34E-13	2.55E-11	5.49E-07	6.46E-13	1.67E-12	4.35E-04	4.68E-13	2.49E-13	3.97E-11
SRH1			1.68E-07	2.83E-06	1.33E-09	8.21E-13	5.71E-10	2.79E-05	8.14E-13	2.44E-12	9.31E-03	1.79E-11	2.78E-10	4.92E-07
AS1				4.16E-09	8.00E-08	1.59E-10	9.62E-07	4.93E-05	6.81E-10	1.67E-10	4.47E-01	3.39E-07	1.26E-06	2.90E-07
AS2					4.18E-10	1.00E-12	7.81E-11	4.46E-08	6.34E-13	4.85E-13	4.74E-06	4.45E-11	3.66E-11	6.27E-09
SNC1						1.26E-06	1.27E-05	5.41E-10	7.43E-06	9.24E-07	5.37E-06	5.43E-07	1.75E-05	4.80E-09
SNC2							1.95E-06	1.55E-10	1.42E-06	2.34E-07	2.58E-06	4.84E-10	1.64E-08	3.92E-12
CRN1								9.49E-10	1.86E-04	1.29E-08	1.54E-03	8.87E-05	8.08E-02	1.53E-06
RRV1									3.49E-11	3.37E-10	1.08E-04	4.44E-12	4.06E-11	3.84E-08
RRV2										6.92E-09	1.02E-03	6.61E-05	5.95E-07	4.45E-11
RRV3											3.06E-07	3.70E-11	1.77E-09	5.47E-12
AV1												7.61E-09	5.52E-07	1.88E-08
AV2													4.37E-05	7.71E-10
OR1														1.32E-07

Table S3. 2016 Vineyard MB-PCA p-value Matrix

	SMV2	SRH1	AS1	AS2	SNC1	SNC2	CRN1	RRV1	RRV2	RRV3	AV1	AV2	OR1	OR2
SMV1	6.98E-06	2.36E-05	8.12E-06	1.00E-10	2.88E-07	1.80E-08	1.42E-11	4.27E-06	6.14E-09	4.25E-07	1.26E-11	4.41E-08	5.23E-10	2.02E-09
SMV2		8.77E-08	7.38E-09	1.49E-11	5.36E-08	2.60E-06	3.69E-07	2.12E-08	5.18E-05	1.87E-08	1.74E-10	1.88E-05	2.35E-06	2.70E-05
SRH1			2.10E-05	2.40E-09	2.82E-08	8.44E-11	2.89E-13	3.44E-04	2.10E-11	4.09E-04	2.66E-11	1.38E-09	5.61E-11	1.70E-09
AS1				6.98E-11	1.99E-09	2.85E-09	3.85E-11	1.99E-02	9.24E-11	1.22E-08	1.37E-11	1.23E-09	5.20E-11	2.39E-11
AS2					3.12E-13	3.11E-13	1.11E-15	2.51E-08	6.46E-13	4.06E-10	3.52E-12	8.21E-12	1.38E-13	1.12E-13
SNC1						1.01E-06	1.80E-11	5.82E-09	9.44E-09	6.88E-08	8.88E-09	2.14E-09	3.82E-10	1.53E-07
SNC2							2.09E-07	2.33E-09	5.48E-06	5.46E-08	1.35E-07	4.55E-10	1.31E-06	3.30E-08
CRN1								5.03E-11	1.65E-04	1.64E-10	6.92E-08	5.93E-10	1.01E-04	7.11E-12
RRV1									1.08E-09	2.07E-08	5.57E-11	1.34E-08	4.89E-11	5.01E-10
RRV2										9.49E-10	2.17E-13	3.74E-13	1.60E-03	3.70E-13
RRV3											4.63E-09	1.70E-09	3.71E-10	5.56E-09
AV1												1.82E-09	3.90E-08	3.52E-07
AV2													1.15E-07	4.87E-05
OR1														4.66E-07

Table S4. 2015 AVA Region NMR p-value Matrix

	SRH	AS	SNC	CRN	RRV	AV	OR
SMV	5.33E-12	7.80E-07	0.00E+00	2.66E-15	0.00E+00	0.00E+00	0.00E+00
SRH		7.62E-05	5.70E-14	1.77E-11	3.75E-11	1.36E-12	0.00E+00
AS			1.97E-09	1.03E-09	0.00E+00	0.00E+00	5.55E-16
SNC				0.00E+00	0.00E+00	9.58E-07	0.00E+00
CRN					0.00E+00	0.00E+00	1.11E-16
RRV						0.00E+00	0.00E+00
AV							1.30E-09

Table S5. 2016 Region NMR p-value Matrix

	SRH	AS	SNC	CRN	RRV	AV	OR
SMV	7.50E-07	0.00E+00	0.00E+00	0.00E+00	0.00E+00	3.86E-14	0.00E+00
SRH		0.00E+00	0.00E+00	1.70E-12	0.00E+00	0.00E+00	0.00E+00
AS			0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
SNC				1.23E-08	0.00E+00	0.00E+00	0.00E+00
CRN					0.00E+00	0.00E+00	1.55E-15
RRV						0.00E+00	0.00E+00
AV							0.00E+00

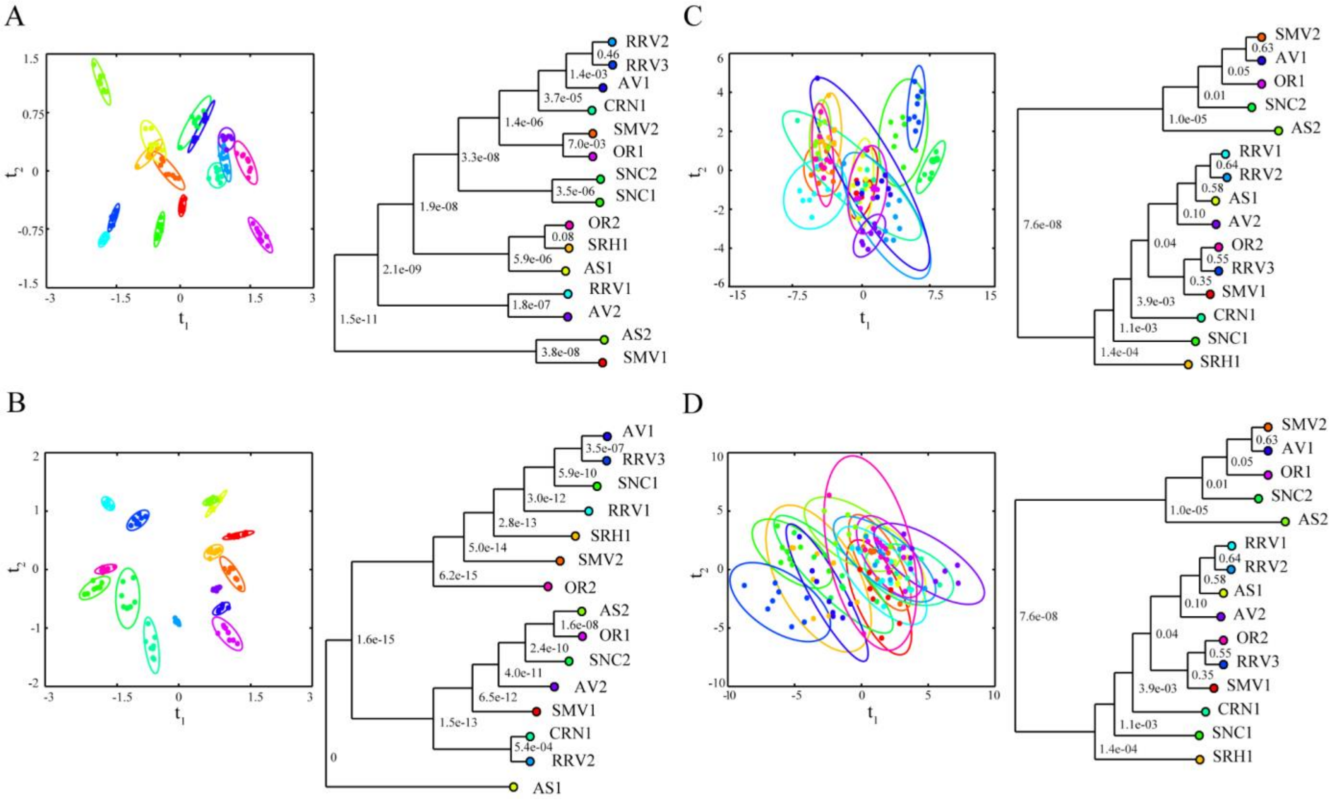

Figure S1. Multivariate Analysis of NMR and DS Array Datasets. LDA-PCA scores plots and associated dendrograms generated from the NMR **(A)** 2015 (R^2 0.79, Q^2 0.75) or **(B)** 2016 (R^2 0.75, Q^2 0.70) data sets. PCA scores plots and associated dendrograms generated from the DS array (**C**) 2015 (R^2 0.78, Q^2 0.72) or (**D**) 2016 (R^2 0.62, Q^2 0.52) data sets. LDA-PCA scores plot, PCA scores plots and dendrograms are displayed with color-coded groups labels, symbols and ellipses as defined in Figure 2. Ellipses represent the 95% confidence interval from a normal distribution. Each node of the dendrogram is labeled with a p-value based on Mahalanobis distances between the groups. LDA models are generated from the first four components of the PCA model.

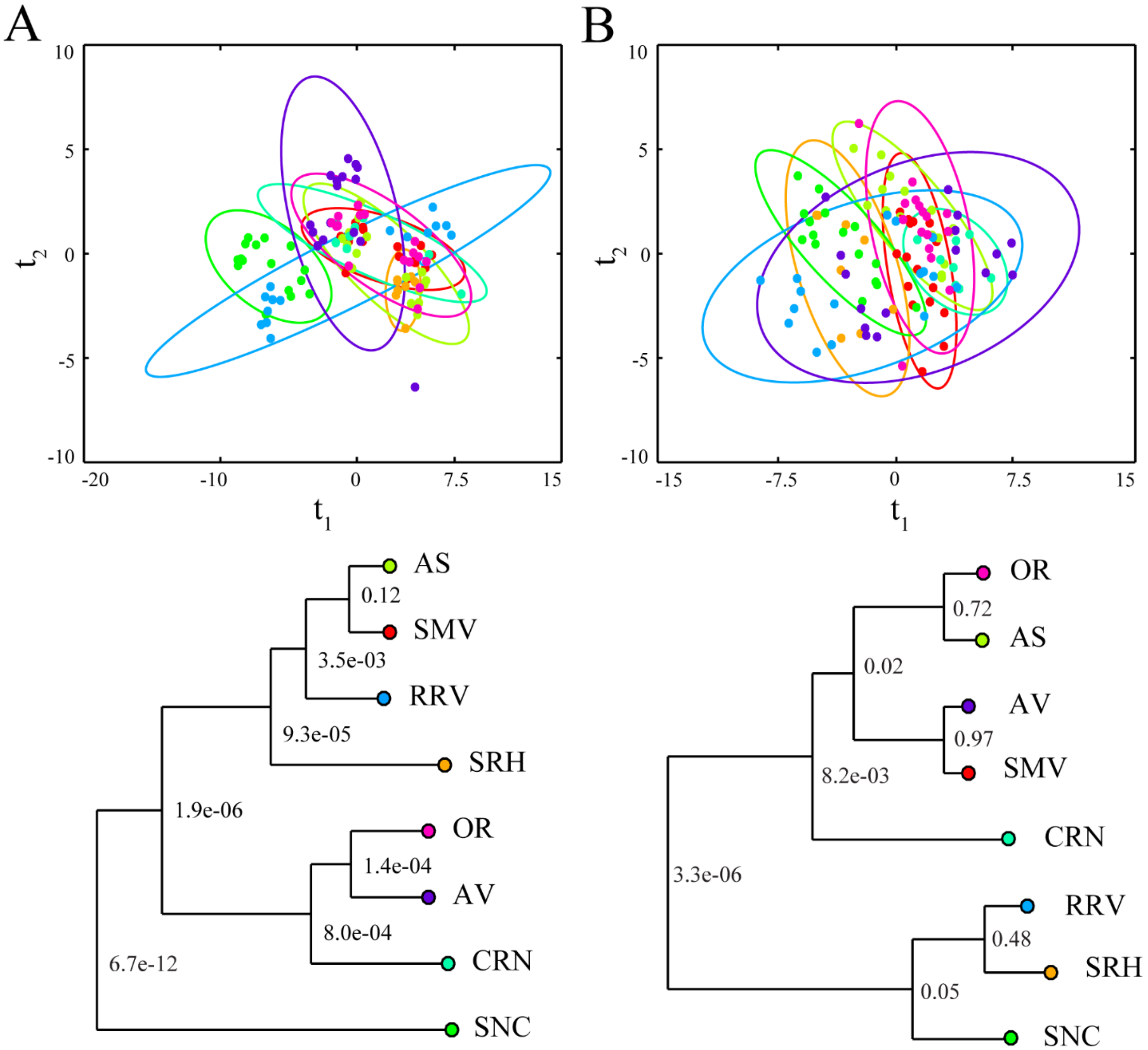

Figure S2. AVA Region Multivariate Analysis of DS Array Datasets. PCA scores plots and associated dendrograms generated from the DS array **(A)** 2015 (R^2 0.88, Q^2 0.754) or **(B)** 2016 (R^2 0.63, Q^2 0.52) data sets. PCA scores plots and dendrograms are displayed with color-coded groups labels, symbols and ellipses as defined in Figure 2. Ellipses represent the 95% confidence interval from a normal distribution. Each node of the dendrogram is labeled with a p-value based on Mahalanobis distances between the groups.

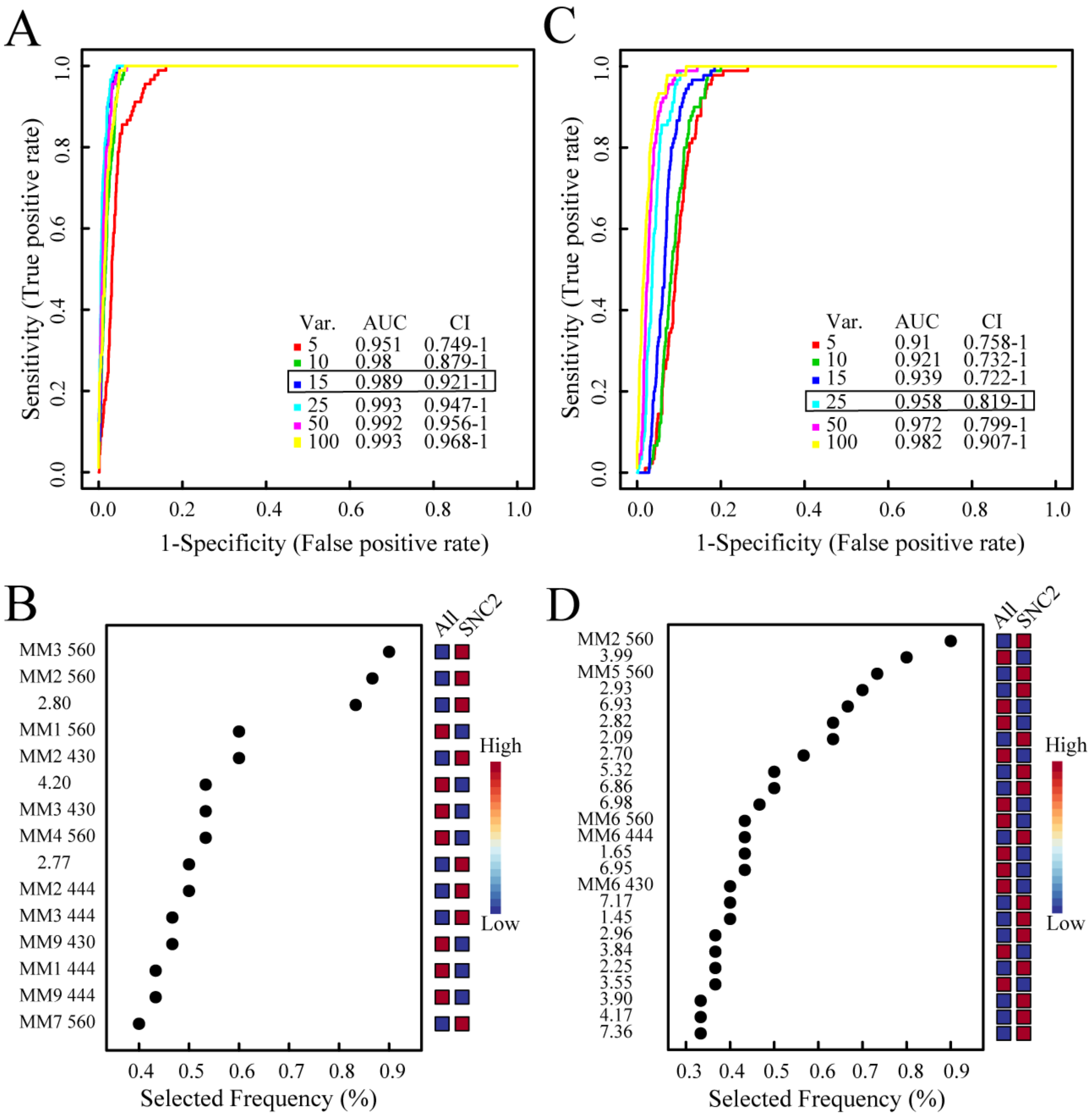

Figure S3. Vineyard and AVA Region Univariate Analysis. Representative ROC curves created from the combined NMR and DS array dataset for the SNC2 vineyard for the **(A)** 2015 and **(C)** 2016 vintages. The top 6 ROC curves are displayed with Area Under the Curve (AUC) and the confidence interval (CI) for all ROC curves indicated. The selected ROC curve is indicated by a box. The top NMR and DS array features that contribute to the ROC model for the **(B)** 2015 and **(D)** 2016 vintages are listed. The ROC curves were generated with MetaboAnalyst 4.0 (https://www.metaboanalyst.ca/) (Xia et al., 2015).

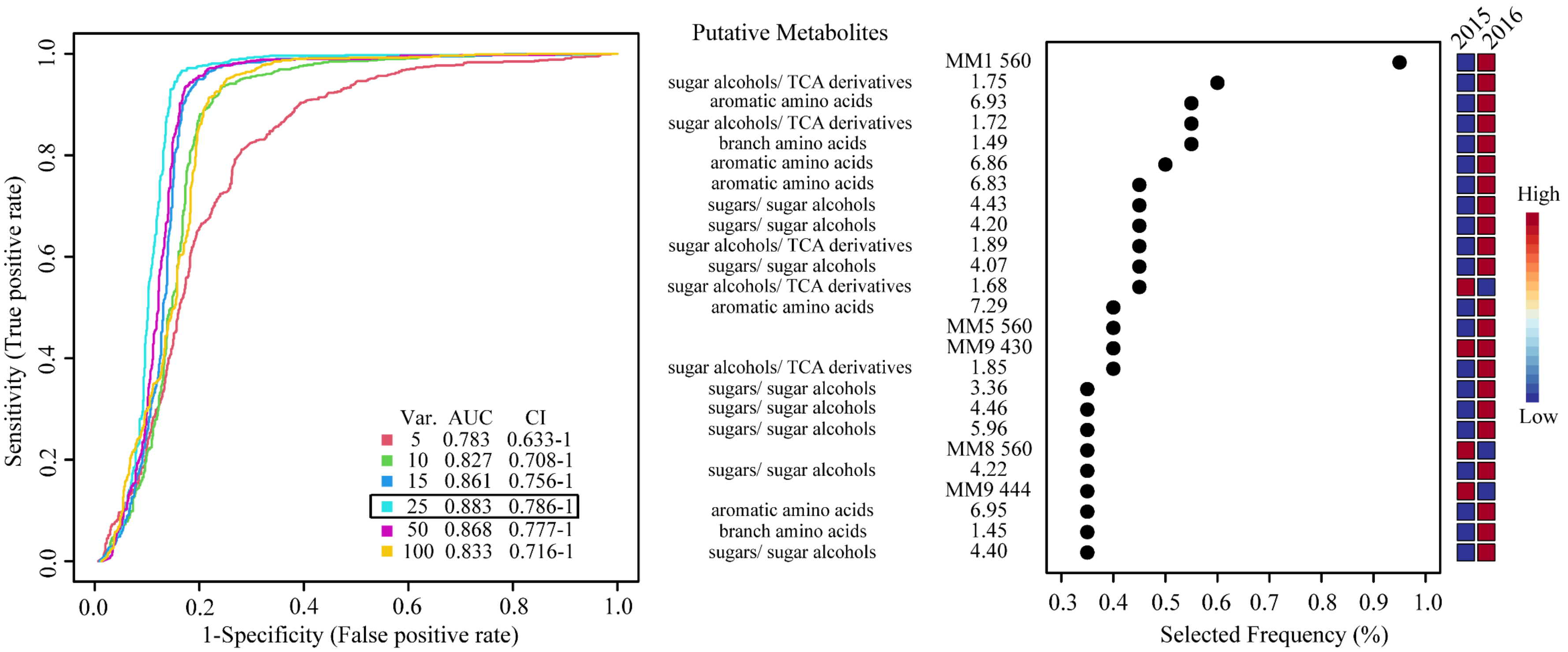

Figure S4. Vintage Univariate Analysis. (*left*) ROC curves created from the combined NMR and DS array dataset for the 2015 and 2016 vintage years. The top 6 ROC curves are displayed with Area Under the Curve (AUC) and the confidence interval (CI) for all ROC curves indicated. The selected ROC curve is indicated by a box. (*right*) The top NMR and DS array features that contribute to the ROC model are listed with (*middle*) putative metabolite class assignments. The ROC curves were generated with MetaboAnalyst 4.0 (https://www.metaboanalyst.ca/) (Xia et al., 2015).