APP NOTE-13/2019GCMS



Метод профилирования различных образцов вина в режиме полного сканирования с использованием хромато-масс-спектрометрической системы **«Маэстро-αMC»**.

Колунтаев Д. А. *к. б. н., научный сотрудник отдела исследований и разработок,* ***ООО Интерлаб****, Россия, Москва*

Ключевые слова

Газовая хромато-масс-спектрометрия, полулетучие органические соединения, профилирование, сканирование в полном массовом диапазоне, метод главных компонент, алгоритм автоинтегрирования.

Резюме

Продемонстрированы возможности хромато-масс-спектрометрической системы Маэстро-αMC на примере профилирования нескольких типов вина для их интерпретирования и статистического разделения на множество подгрупп.

Введение

Виноградно-винодельческая отрасль занимает особое место среди других отраслей экономики. Продукция данной отрасли имеет большой потребительский спрос и поэтому потребитель начинает все больше обращать внимание на органолептические характеристики вин, задумываться о престиже и марке вина, которое он пьет. [1]

Сложный и приятный вкус вина – это результат идеального равновесия и баланса между его компонентами, в частности такими, как сахара, спирты, полиспирты, кислоты, танины, минеральные вещества и другие. Аромат вина определяется сотнями летучих соединений, присутствующих в широком диапазоне концентраций. Характеристика этих ароматических соединений имеет важное значение при классификации вин, а, следовательно, и в контроле их качества.

Хромато-масс-спектрометрическая система «Маэстро-αMC» (ООО «Интерлаб», Россия) позволяет получать масс-спектрометрические данные полного сканирования множества компонентов, составляющих аромат вин, с достаточной чувствительностью, динамическим диапазоном и оптимальной скоростью сбора данных. Последнее, в частности, позволяет надёжно интерпретировать присутствующие компоненты.

В работе рассмотрена реализация метода главных компонент в вычислительной среде *R*, где в качестве исходных данных использовались результаты интегрирования хроматографических профилей вин, полученных в ходе анализа с помощью масс-детектора «Маэстро-αMC

Экспериментальная часть

Вина марок «Krstac», «Rose», «Vranac» и «Rubin» (все производства Черногория) были куплены в специализированном винном магазине. По 20 мл каждого вина смешивали с 10 мл дихлорметана и далее смесь перемешивали при комнатной температуре в течение 30 минут. После перемешивания образцы центрифугировали со скоростью 3500 об/мин в течение 10 мин. Нижний слой дихлорметана осторожно отбирали пипеткой и переносили в дополнительную центрифужную пробирку (рис. 1), содержащую безводный сульфат натрия (2 г). Далее собранный элюат встряхивали в течение 30 секунд. По завершению встряхивания экстракт снова центрифугировали со скоростью 3500 об/мин в течение 10 минут, после чего отбирали органический слой и упаривали его в токе азота до 1 мл. 1 мкл финального экстракта вводили в газохроматографическую систему «Маэстро-αMC» (табл. 1).

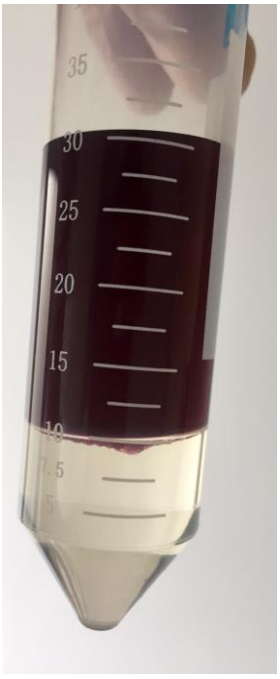


Рис. 1. Экстракция образцов вина.

Табл. 1. *Параметры инструментального метода.*

Оборудование	
ГХ-МС система	Interlab «Маэстро-αMC»
Испаритель	Split/Splitless (liner: Restek 4 mm x 6.3 x 78.5)
Хроматографическая колонка	Rtx-5MS (30m x 0.25mm; 250um)

Параметры хроматографического разделения	
Объём пробы	1,0 мкл
Режим испарителя	Split; 1:10 деление
Температура испарителя	250°C
Режим работы испарителя	Режим Constant Flow
Градиент t° печи	45°C в течение 2,0 мин; 13,0°C/мин до 180°C, удержание 0,0 мин; 10,0°C/мин до 240°C, удержание 0,0 мин.
Поток газа-носителя	(He) 1 мл/мин
T° интерфейса МС	250°C
Параметры детектирования	
Источник ионизации	*ЭИ (EI)*
T° источника	230°C
Задержка растворителя	3,0 мин
Режимы сканирования	SCAN (40 – 400 Да)
Скорость сканирования	2000 Да/сек

Результаты и обсуждение

Обработка полученных скан-данных профилей вина была выполнена алгоритмом «автоинтегрирования», который содержится в пакете программного обеспечения «Маэстро-αМС». Автоинтегратор производит обработку данных, полученных в режиме полного сканирования. В частности, производится детектирование хроматографических пиков на хроматограммах выделенных ионов каждой из масс. Полученные пики группируются по времени хроматографического выхода, в результате чего формируются выделенные, очищенные от химического фона спектры компонентов пробы. Финальной стадией процедуры автоинтегрирования является опрос множества библиотек масс-спектров на наличие подобных спектров. В результате, в ходе работы были детектированы свыше 70 полулетучих компонентов в каждой пробе вина. Наиболее характерные из них представлены в таблице 2. Стоит также отметить, что подтверждение результатов библиотечного поиска было выполнено также с помощью деконволюционного алгоритма AMDIS (ver. 2.73, NIST).

Рис. 2. Процедура автоинтегрирования пробы белого вина для нецелевого скрининга.

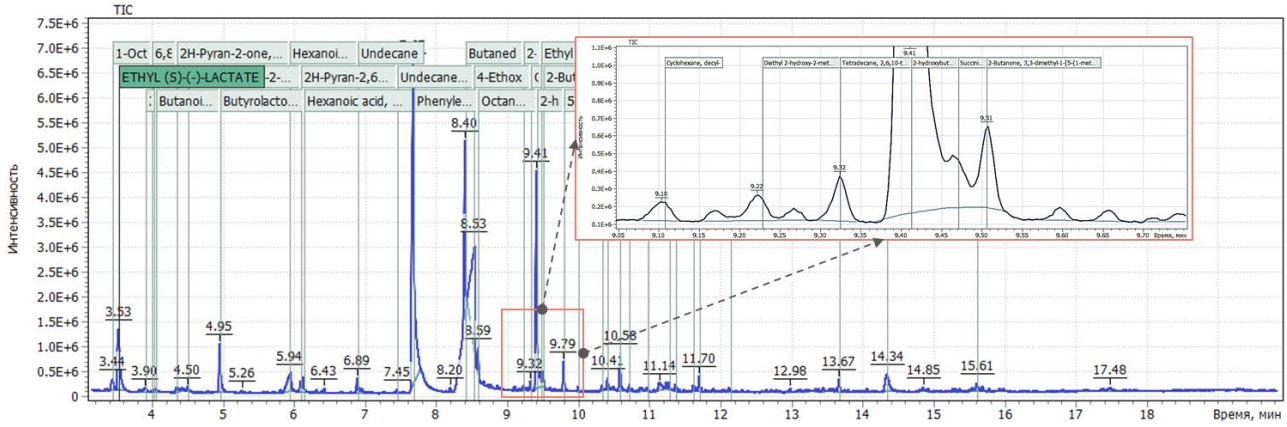


Табл. 2. Общие компоненты, детектированные с помощью библиотеки NIST в профилях вин.

RT, min	Компонент	Score NIST	Rev. Score NIST	RT, min	Компонент	Score NIST	Rev. Score NIST
3.1	Butanoic Acid	730	750	7.67	*Unknown*		
3.39	Butanoic Acid, ethyl ester	744	801	7.82	Phenylethyl Alcohol	828	864
3.44	Acetic acid, anhydride with formic acid	740	841	7.91	N-(3-Methylbutyl)acetamide	894	894
3.53	Ethanol, 2-methoxy-	858	924	8.27	Phenol, 4-ethyl-	914	924
3.83	Furfural	855	855	8.40	*Unknown*		
3.9	3-methylbutanoic acid	751	763	8.59	Octanoic acid, ethyl ester	809	813
4.05	Butanoic acid, 2-methyl	780	780	9.46	Succinic acid, ethyl 2-hexyl ester	750	750
4.4	2H-Pyran-2-one, tetrahydro-3,6-dimethyl-	800	824	10.72	Decanoic acid, ethyl ester	750	750
4.42	1-Butanol, 3-methyl-, acetate	850	850	11.95	Acetamide, N-(2-phenylethyl)-	900	900
4.5	2,3-Epoxybutane	702	905	12.4	N-Acetyltyramine	871	871
4.95	Butyrolactone	933	937	12.64	Benzoic acid, 4-hydroxy-3-methoxy-, ethyl ester	890	890
5.26	Ethyl, 3-hydroxybutanoate	740	740	13.67	Pentadecane	800	800
5.81	*Unknown*			14.35	Tryptofol	763	789
5.93	Hexanoic acid	865	865	14.71	Benzoic acid, 4-hydroxy-3,5-dimethoxy-	808	834
6.09	2H-Pyran-2,6(3H)-dione	851	887	14.76	*Unknown*		
6.13	Hexanoic acid, ethyl ester	877	877	14.80	p-Hydroxycinnamic acid, ethyl ester	867	867
6.31	Hexane, 2-methyl-	829	829	14.85	Ethanone, 1-(1,3-dimethyl-1H-indol-2-yl)-	737	737
6.42	Nonane, 2,6-dimethyl-	781	781	14.92	Tridecane, 2-methyl-	743	743
6.89	Decane, 2,4,6-trimethyl-	840	853	15.53	1H-Indole-3-ethanol, acetate (ester)	940	940
6.95	Undecane, 4,7-dimethyl-	715	715	15.54	Undecane, 2,10-dimethyl-	770	770
7.05	Isopentyl methoxyacetate	752	752	15.79	Pyrrolo[1,2-a]pyrazine-1,4-dione, hexahydro-3-(2-methylpropyl)-	799	803
7.09	*Unknown*			15.93	Hexadecanoic acid	700	700
7.45	Octane, 2-methyl-	808	808				

Анализ данных

Задачей данной работы являлось не только апробация скрининга полулетучих компонентов вин, но также и их аналитическое разделение на основные компоненты, которые в дальнейшем могли бы характеризовать тот или иной тип вина. Для выполнения подобной задачи был сформирован общий лист всех детектированных полулетучих компонентов и проведена статистическая обработка массива с использованием метода главных компонент (PCA). Реализация метода была выполнена в вычислительной среде R (RStudio ver. 1.2.5019).

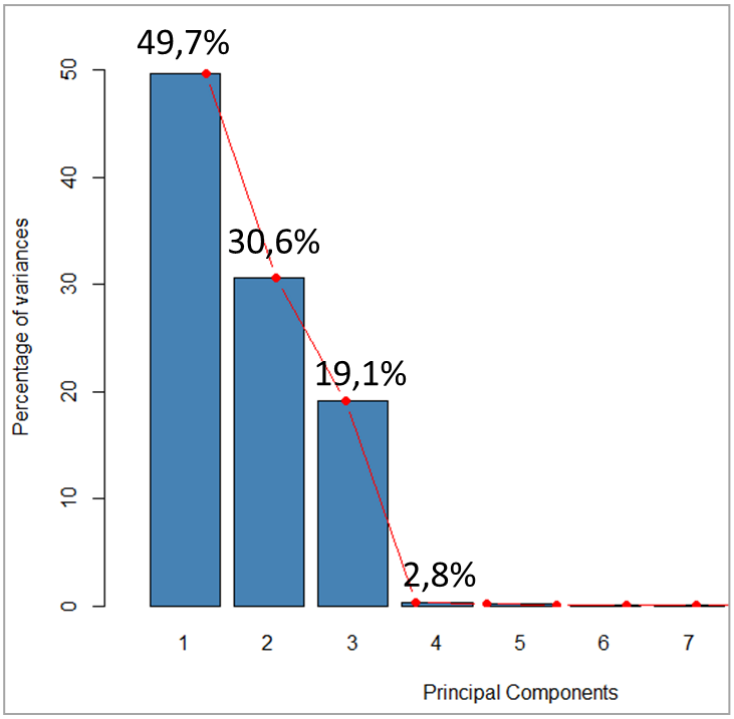
Метод главных компонент позволяет заменить несколько исходных переменных на меньшее количество новых переменных (выполнить снижение размерности). Новые искусственные переменные называются *главными компонентами*.

На рисунке 3 представлены основные статистические характеристики компонентов и снижение размерности массива данных до 13 главных компонент (PC1-PC13). Каждая компонента характеризуется значением среднеквадратичного отклонения главных компонент («Standard deviation»), объясняет долю от общего отклонения компонент («Proportion of Variance») и из последней строки «Cumulative Proportion» следует, что первая компонента PC1 описывает 49,7% отклонений исходного набора данных профилей вин, а первая и вторая главные компоненты объясняют уже порядка 80% отклонений. Остальные компоненты не вносят сильного влияния на различие.

Рис. 3. Основные статистические характеристики компонент.

```
Importance of components:
                          PC1    PC2    PC3     PC4    PC5    PC6     PC7    PC8     PC9    PC10    PC11    PC12
Standard deviation     3.7948 2.9771 2.3534 0.28530 0.2220 0.1521 0.12285 0.1081 0.09587 0.06703 0.05175 0.02545
Proportion of Variance 0.4966 0.3056 0.1910 0.00281 0.0017 0.0008 0.00052 0.0004 0.00032 0.00015 0.00009 0.00002
Cumulative Proportion  0.4966 0.8022 0.9932 0.99599 0.9977 0.9985 0.99901 0.9994 0.99973 0.99989 0.99998 1.00000
                          PC13
Standard deviation     2.759e-16
Proportion of Variance 0.000e+00
Cumulative Proportion  1.000e+00
```

Более наглядно данные можно продемонстрировать на диаграмме процентов отклонений, объясняемых главными компонентами. Как уже было сказано выше, из диаграммы видно, что компонента PC1 объясняет 49,7% всех различий в данных, что означает, что почти половина информации массива (30 исходных признаков) могут быть инкапсулированы только этой компонентой. PC2 объясняет 30% различий.



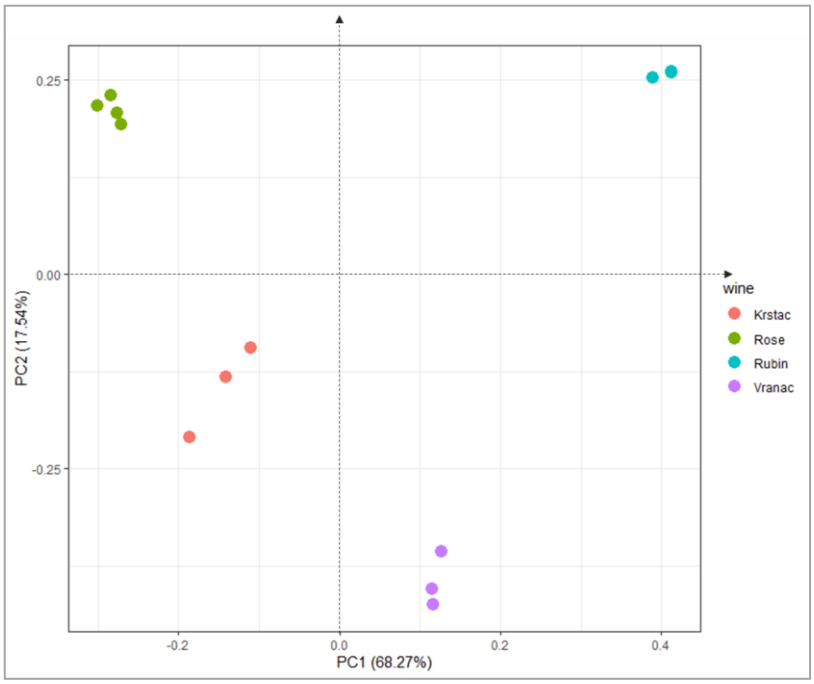
Таким образом, зная положение данных (целевых пиков) для какого-либо типа образца вина по отношению только к главным компонентам PC1 и PC2, мы можем получить очень точное представление о том, где образец будет находиться по отношению к другим образцам, поскольку только PC1 и PC2 могут объяснить 80% различий

Рис. 4. Проценты отклонений, объясняемые главными компонентами.

Таким образом, дальнейшие расчёты предполагается проводить только согласно первой и второй главной компоненты (PC1/PC2).

На рисунке 5 приведён график величин вклада каждого из 30 исходных признаков в первые две компоненты. Видно, как различные типы вин распределились согласно анализу двух главных компонент. При этом легко заметить, что каждый тип вина и каждая марка вина имеет своё точное расположение.

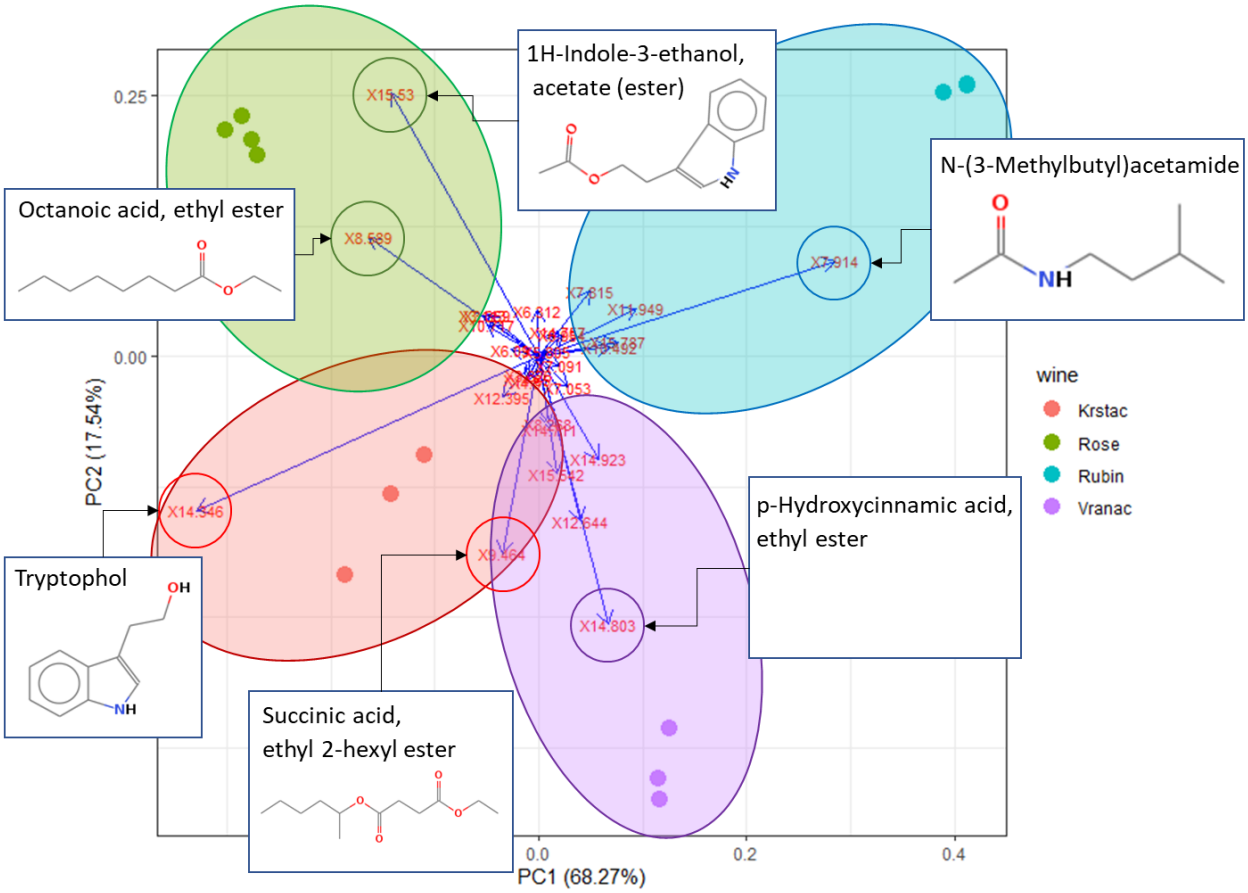
Рис. 5. Визуализация первой и второй главных компонент.



Ниже на рисунке 6 добавлены векторы, которые соответствуют временам элюирования наиболее характеристичных компонент (в зависимости от направления вектора) к тому или иному распределению типов вина.

Для завершения работы необходимо только определить компоненты, которые демонстрируют наибольшее влияние на отличие данных и которые могут быть использованы для их классификации.

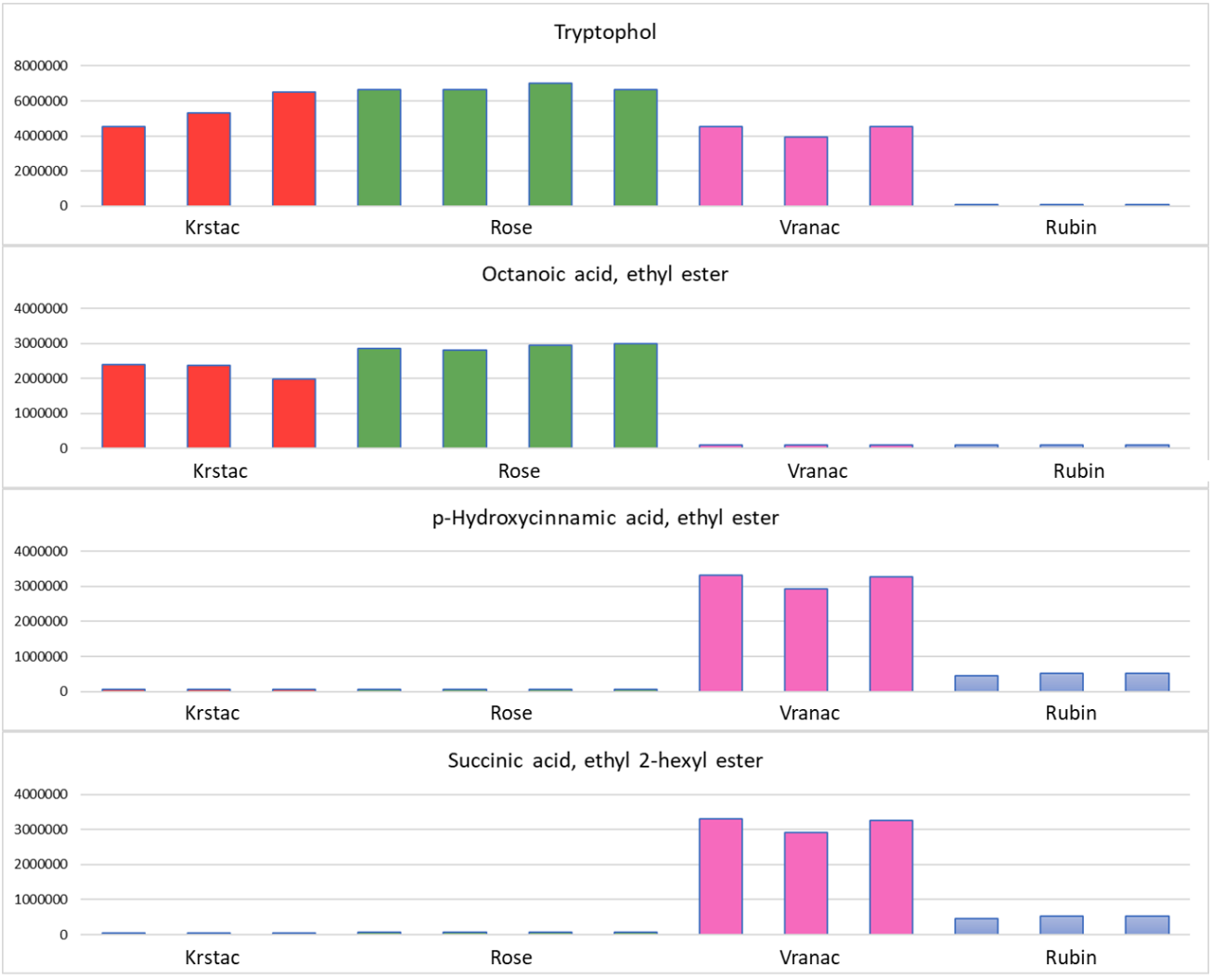
Рис. 6. Визуализация первой и второй главных компонент с векторами соответствующих компонент.



В результате анализа данных можно сделать предположение, что этиловый эфир октановой кислоты и 1H-индол-3-этанол ацетат в большей концентрации содержатся в розовом вине «Rose» в сравнении с красными винами «Rubin» и «Vranac», а также практически полностью отсутствует в белом вине «Krstac». Компонент триптофол в большей концентрации содержится в белом вине марки «Krstac» относительно других марок вин, а также этиловый эфир янтарной кислоты соответствует в равной степени винам «Krstac» и «Vranac». Согласно графику, можно также сделать предположение, что компонент N-(3-метилбутил) ацетамид является уникальным компонентом вина марки «Rubin». Данное соединение относится к семейству амидов вторичных карбоновых кислот и отсутствует в профиле других вин [2]. В то же время компонент этиловый эфир п-гидроксициннамоновой кислоты в большей степени был детектирован в винах марки «Vranac».

Пакет гистограмм на рисунке 7 дополнительно подчеркивает химическую разницу между различными сортами вин и подтверждает присутствие в них маркерных компонентов.

Рис. 7. Гистограмма химической разницы между различными сортами вин.



Выводы.

1. Представленная работа показывает, что инструменты быстрой обработки данных, используемые в программном обеспечении «Маэстро-αMC», позволяют проводить надёжное детектирование множества компонентов смеси с автоматическим опросом коммерческих библиотек масс-спектров для их идентификации в биологических матрицах, в частности таких, как вина различных сортов.
2. При этом результаты интегрирования могут быть легко импортированы в любое внешнее программное обеспечение для их дальнейшей обработки. Такое преимущество может быть использовано для простого интегрирования данных получаемых с помощью масс-спектрометрической системы и обрабатываемых в программном обеспечение «Маэстро-αMC» и особенно удобно при написание различных научных статей, публикаций, а также расчёта «сырых» данных в любой другой внешней программе их обработки.

Литература

1. *Барабаш И. П., Романенко Е. С., Сосюра Е. А., Нуднова А. Ф., Чернов А. И., Селиванова М. В., Юхнова А. А.* Отечественное виноделие: перспективы развития // Современные проблемы науки и образования. – 2013. – № 1.
2. *http://www.thegoodscentscompany.com/data/rw1634481.html*