International PhD Program in Biomolecular Sciences

Department of Cellular, Computational

and Integrative Biology – CIBIO

34th Cycle

"Application of GC×GC-MS in VOC analysis of fermented beverages"

Tutor/Advisor

Urska VRHOVSEK *urska.vrhovsek@fmach.it*

Silvia CARLIN silvia.carlin@fmach.it

Ph.D. Thesis of

Penghan ZHANG

penghan.zhang@unitn.it

Academic Year 2018-2021

Contents

Abstract

GC×GC is an efficient tool for the analysis of volatile compound. However, improvements are still required on VOC extraction, GC×GC setup and data processing. Different sample preparation techniques and GC×GC setup were compared based on the literature study and experimental results. Each VOC extraction technology has its own drawbacks and needs new developments. There wasn't an ideal sample preparation technique to recover all the VOCs from the beverage sample. Furthermore, the VOCs recovered by different techniques were very different. The discussion of the pros and cons of the different techniques in our study can serve as a guide for the further development and improvement of these techniques. Combining the results from different sample preparation techniques is necessary to achieve a higher coverage of global VOC profiling. For the known fermentative aromatic compounds, the best coverage can be reached by using SPME together with SPE for beer, and VALLME for wine and cider.

A fine GC×GC method development involves modulator selection, column combination and parameter optimization. Thermal modulator provides high detection sensitivity and allow exceptional trace analysis. Since the analytes coverage is the most important factor of in beverage VOC profiling, thermal modulation is a better choice. In fermented beverages, there are more polar compounds than non-polar compounds. The most suitable column combination is polar-semipolar. Same column diameters shall be used to minimize the column overloading. GC×GC parameters must be optimized. These parameters interact with each other therefore statistical prediction model is required. Response surface model is capable of doing this job while using a small number of experimental tests. The nearest neighbor distance was a suitable measurement for peak dispersion.

Column and detector saturations are unavoidable if the metabolic sample is measured at one dilution level, incorrect peak deconvolution and mass spectrum construction may happen. Data processing results can be improved by a two-stage data processing strategy that will incorporate a targeted data processing and cleaning approach upstream of the “standard” untargeted analysis. Our experiments show a significant improvement in annotation and quantification results for targeted compounds causing instrumental saturation. After subtracting the saturate signal of targeted compounds, the MS construction was improved for co-eluted compounds. Incomplete signal subtraction may occur. It leads to the detection of false positive peaks or to interferences with the construction of mass spectra of co-diluted peaks. High-resolution MS libraries and more accurate peak area detection methods should be tested for further improvement.

List of Papers

Paper 1

(in preparation)

Zhang, P., Carlin, S., Franceschi, P., Mattivi, F., & Vrhovsek, U. (2021). Review: Application of GC×GC-MS in fermented beverage VOC analysis.

Author contribution:

P.Z. wrote the paper. S.C., P.F., F.M., and U.V. revised the manuscript.

Paper 2

Zhang, P., Carlin, S., Lotti, C., Mattivi, F., & Vrhovsek, U. (2020). On sample preparation methods for fermented beverage VOCs profiling by GC×GC-TOFMS. Metabolomics, 16(10), 102. https://doi.org/10.1007/s11306-020-01718-7

Author contribution:

P.Z., S.C., and U.V. designed the study. P.Z., S.C., and C.L. performed the experiments and analyzed the data. P.Z. wrote the paper. S.C., F.M., and U.V. revised the manuscript. All authors read and approved the final manuscript.

Paper 3

(revision submitted)

Zhang, P., Carlin, S., Franceschi, .P, Mattivi, F., & Vrhovsek, U. (2021). Application of target guided data processing approach in saturated peaks correction of GC×GC analysis.

Author contribution:

P.Z., S.C., and U.V. designed the study. P.Z. performed the experiments and analyzed the data. P.Z. wrote the paper. S.C., P.F., F.M., and U.V. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Paper 4

Giannakou K., Visinoni F., Zhang P., Nathoo N., Jones P., Cotterrell M., Vrhovsek U., & Delneri D. (2021). Biotechnological exploitation of Saccharomyces jurei and its hybrids in craft beer fermentation uncovers new aroma combinations. Food Microbiology, 100, https://doi.org/10.1016/j.fm.2021.103838

Author contribution:

D.D. conceived the study; D.D. and U.V. supervised the genetic and phenotypic experiments, and the GC×GC-MS analysis, respectively. M.C. and P.J. supervised the pilot-scale fermentation. K.G. carried out the experimental work with the inputs of F.V. N.N. contributed to the construction of five hybrids. K.G. and F.V. analyzed the genetic and physiological data and KG, F.V. and P.Z. performed and analyzed the GC×GC-MS volatile compounds spectra. K.G., F.V. and D.D. wrote the manuscript with input of M.C., P.Z., U.V.. All authors contributed to the article and approved the submitted version.

Paper 5

Lin C., de la Cerda García-Caro R., Zhang P., Carlin S., Gottlieb A., Agerlin Petersen M., Vrhovsek U., & Bond U. (2021). Packing a punch: understanding how flavours are produced in lager fermentations, FEMS Yeast Research, 21(5). https://doi.org/10.1093/femsyr/foab040

Author contribution:

C.L., R.C. and P.Z. wrote the paper. S.C., A.G., M.P. U.V. and U.B. revised the manuscript.

Acronyms

1D	one dimensional
^{1}D	first dimensional
$^{1}T_R$	first dimensional retention time
2D	two dimensional
^{2}D	second dimensional
$^{2}T_R$	second dimensional retention time
AEDA	aroma extract dilution analysis
AALLME	air-assisted liquid-liquid microextraction
B.C.	before christ
Car	carboxen
CE	capillary electrophoresis
DHS	dynamic headspace
DLLME	dispersive liquid-liquid microextraction
DVB	divinylbenzene
EG silicon	ethylene glycol modified silicone
GC×GC	two dimensional comprehensive gas chromatography
HPLC	high-performance liquid chromatography
LC	liquid chromatography
LLE	liquid-liquid extraction
RTI	retention time indices
LTPRI	linear temperature programmed retention index

m/z	mass to charge ratio (ion channel)
MHE	multiple headspace extraction
MS	mass spectrometer or mass spectrum
mSBSE	multiple stir bar sorptive extraction
NND	nearest neighbor distance
NMR	nuclear magnetic resonance
OAV	odor activity value
OSME	odor-specific magnitude estimation
p.w.	peak witdth
PARAFAC	parallel factor analysis
PDMS	polydimethylsiloxane
RCF	relative centrifugal force
SELDI	surface enhanced laser desorption/ionization
SHS	switchable hydrophilic solvents
SPE	solid phase extraction
SPME	solid phase microextraction
SPS	switchable polar solvents
TD	thermal desorption
TF	thin-film
TIC	total ionic chromatogram
TOF	time-of-flight
VALLME	vortex-assisted liquid-liquid microextraction

VOC	volatile organic compound

Preface

It is important to understand the science of beverage flavor. However, the flavor of fermented beverages is complex, as several different types of volatile organic compounds(VOCs) contribute to the expression of a beverage flavor. These compounds come from four sources: (1) primary aromas, i.e., those arising from the row material, such as grape for wine and hops for beer; (2) pre-fermentative aromas, i.e., those compounds formed during operations between harvest and the beginning of alcoholic fermentation; (3) secondary aromas, i.e., those arising from the fermentation process; and (4) tertiary aromas, i.e., those arising during subsequent storage, especially for long-term storage of wine in wooden barrels or bottled wine contact with oak. These compounds contribute to the citrusy, herbal, spicy, flowery, fruity, and sometimes also to off-flavors. To obtain an accurate VOC profile of beverage, an advanced analytical platform is demanded to detect and ideally identify each of them. Extraction methods followed by two-dimensional comprehensive gas chromatography coupled with a mass spectrometer (GC×GC-MS) might be a solution.

In this study, the current state of the art and future prospects of applying GC×GC-MS into fermented beverages VOC profiling is firstly discussed. The application of GC×GC-MS was optimized based on VOCs in fermented beverages. The experimental optimization consisted of three parts: the VOC extraction, the GC×GC separation, and the data processing methods. The results of these experiments were formulated into two publications. In the following, the experimental methods and results of this thesis will be briefly described.

Chapter 1 discussed current state of the art and future prospects of applying GC×GC-MS into fermented beverages VOC profiling. The paper covers sample

preparation techniques, modulator selection, column combination, GC×GC parameter optimization, and data processing approaches.

Chapter 2 studied the impact of using different VOC extraction techniques [dynamic headspace (DHS), vortex-assisted liquid–liquid microextraction (VALLME), multiple stir bar sorptive extraction (mSBSE), solid phase extraction (SPE), and solid phase microextraction (SPME)] to figure out the most suitable sample preparation protocol for profiling the VOCs from fermented beverages. The GC×GC over parameter were optimized based on the beverage VOCs as well. The result shows that GC×GC oven parameters can be optimized with the Box–Behnken surface response model and response measure on peak dispersion. Due to the unavoidable column and detector saturation during metabolomic analysis, errors may happen during mass spectrum construction. Profiling results obtained with different sample preparation methods show considerable variance. Common findings occupy a small fraction of total annotated VOCs. For known fermentative aromas, best coverage can be reached by using SPME together with SPE for beer, and VALLME for wine and cider.

Chapter 3 proposed a data processing workflow to resolve the problems caused by detector and column saturation which commonly happens in the VOC profiling of fermented beverages. The workflow is a two-stage data processing strategy that will incorporate a targeted data processing and cleaning approach upstream of the “standard” untargeted analysis. By using the retention time and MS stored in a library, the annotation and quantification of the targeted saturated peaks have been significantly improved. After subtracting the non-perfected signals caused by saturation, peaks of co-elutes can be annotated more accurately.

Each Chapter features an introduction to the respective topic and a thorough discussion of the presented results. Chapter 1, 2, and 3, at the same time, present publications compiling results obtained during this work (see List of Papers for author contribution). The thesis is concluded by a summary and outlook in Chapter 4.

1 Application of GC×GC-MS in fermented beverage VOC profiling

Penghan Zhang, Silvia Carlin, Pietro Franceschi, Fulvio Mattivi, and Urska Vrhovsek

(in preparation)

Abstract

To meet consumer demand for fermented beverages with a wide range of flavors, it is important to characterize VOCs in the beverage and their relationship with raw materials, fermentation process, and the aging process. Sample preparation techniques coupled with GC×GC-MS is a promising method for the analysis of various VOCs in fermented beverages. Few articles discuss the application of GC×GC for the characteristics of fermented beverage VOCs and the problems faced in the experimental analysis. This review critically discusses each step of GC×GC-MS workflow in the specific context of fermented beverage VOC research, includes the most general VOC extraction techniques; modulator, column setup, and parameter optimization methods for GC×GC instrument setup; and background correction, peak detection, annotation, statistical cross-class study for data handling. Novel SPME, SBSE, and SPE have a prospective in beverage VOCs analysis. Thermal modulation provides higher detection sensitivity. Polar-semipolar column with the same inner diameter is recommended. For GC×GC parameter optimization, model response can be generated with modified nearest neighbors' distances (NND) approach according to the specific demands of each study. Local minima method is suitable for the background correction. Improving the results of peak detection requires cross validating multiple methods. Development of a template database for beverage VOCs can greatly improve robustness and accuracy of peak annotation. Mathematical feature comparisons based on raw GC data can avoid the impact of errors in prior data processing on inter-group feature extraction.

1.1 Introduction

Humans began consuming fermented beverages, including wine, beer, and other products, from the Neolithic period [1]. Wine is considered to be one of the most ancient known alcoholic beverages, and is produced by the fermentation of grape juice by yeasts. Grape-based fermented beverages were produced in China as early as the sixth century B.C. or slightly earlier; wine was also developed in the Near East c. 6000-5800 B.C. [2,3]. It played a central role in the livelihood of the ancient Greeks and Romans. In addition to its gastronomical and social uses, there are strong bonds between wine and many ancient religions, including Judaism and Christianity [4].

The principal wine producing countries are now in Europe, the Americas, China, Australia and South Africa, while European countries and Argentina head the list of wine consuming countries [5]. The earliest known barley beers can be dated back to about 3400 B.C. In the Middle Ages, beer continued to be the main beverage in the cooler climates of Northern Europe. Today, it is the third most popular beverage after water and tea [6].

The value of the global alcoholic beverages market was USD 1,439 billion in 2017 and is expected to reach USD 1,684 billion by 2025, growing at a CAGR of 2.0% from 2018 to 2025 [7]. Specifically, in terms of beer, the global beer market size was USD 693.39 billion and is expected to reach USD 962.39 billion by 2026, exhibiting a CAGR of 4.22% [8]. In the case of wine, in 2019 the market size was USD 364.25 billion and is expected to reach USD 444.93 billion by 2027, exhibiting a CAGR of 6.05% [9]. The growth of the global alcoholic beverages market is attributed to the increase in the global young adult population, in the surge in disposable income, and in the steady rise in consumer demand for premium/ultra-premium products. These young drinkers

prefer to have innovative products with more sophisticated and intense flavors as compared to traditional drinkers [10].

In order to match the production and the new demands, it is important to understand the science of flavor. The flavor of fermented beverages is determined by the volatile organic compounds (VOCs) from raw materials that are produced during the fermentation process, and the aging process. Over 1300 VOCs are found in fermented beverages [11]. Such compounds contribute to citrus, herbal, spicy, floral, fruity, and sometimes off-flavors [12]. Extraction methods and gas chromatography coupled with detectors are the traditional methods for detecting and identifying VOCs [13]. To accurately obtain a VOC profile of a beverage, an advanced analytical platform is required to detect and ideally identify each VOC. Two-dimensional gas chromatography coupled with mass spectrometry (GC×GC-MS) may be a solution [14]. When GC×GC-MS is used, the final results of VOC profiling depend on the quality of sample extraction, GC×GC separation and data processing.

This review critically discusses each step of GC×GC-MS workflow in the specific context of fermented beverage VOC research. The extraction part includes the most general techniques. Techniques targeting a specific compound or a group of compounds are not taken into consideration. GC×GC separation part covers the modulator, column setup, and parameter optimization methods. And in the last data processing part, different algorithms, their pros and cons are discussed based on the demand of VOC profiling.

1.2 Sample extraction

Direct injection of fermented beverage into GC is not suitable because the sample contains large amounts of non-volatile components which are harmful

to the instrument [15]. Besides, some VOCs need to be concentrated to reach the detection limit of the detector. However, there is no universal method of sample preparations that is suitable for the untargeted screening of different volatiles, as the concentration of beverage aromas ranges from mg/L to ng/L levels and cover a wide polarity range. The physicochemical properties of common VOCs in the beverages are listed in Table S 1. Hence different sample preparation techniques are involved before the GC injection [16]. Selection of the VOC extraction methods is crucial. Different methods vary in extraction mechanism and absorption material. They require different optimization. And, most importantly, applying different extraction methods results in different analytical sensitivity to one volatile compound and different coverage to the entire VOC profiling [17–19].

Before talk about the sampling techniques, it's necessary to clarify several types of extraction procedures: a single extraction, a repeated stepwise extraction, or a continuous extraction. In case of a single extraction step, sample is placed in a closed multi-phase system. After an enough long period, the system reaches equilibrium and the concentration of the analyte in each phase remains constant. The values of the concentrations are controlled by the equilibrium constant (distribution constant, partition coefficient, Henry's law constant). Most modern sampling techniques can be applied with this approach, for example, HS and HS-SPME, in-solution SPME, LLE and SBSE. If exhaustion sampling is the purpose, stepwise extraction or multiple extraction can be applied. In this procedure, the single extraction step may be repeated and aliquots are analyzed again. The combined information (sum of signals) of these analyses corresponds to the total amount of that analyte in the original sample, and the resulting sum of peak areas is thus independent of the unknown equilibrium constant. In continuous extraction, equilibrium is not required. Analytes is continuously removed from the liquid beverages sample by gas (DHS) or stationary phases (SPE). Analytes to try to reestablish the equilibrium state

which is never reached. Finally, all the volatile analytes are removed from the sample and collected [20].

In this review, we will talk about five common extraction methods and their possible further applications on beer, wine and other fermented beverages. Unfortunately, because of the diversity of VOCs, an ideal method that can concentrate and recover all VOC compounds does not exist. Commonly used sample preparation techniques for VOC analysis are dynamic headspace (DHS), Dispersive liquid-liquid microextraction (DLLME), multiple stir bar sorptive extraction (mSBSE), solid phase extraction (SPE), and solid phase microextraction (SPME). They can be roughly classified into two categories, headspace or in-solution. Headspace techniques are HS-SPME and DHS. In-solution techniques are DLLME, mSBSE, and SPE. SPME can also be applied as in-solution sampling. Applications of the headspace analysis benefit either from the simplicity of automation or removing some of the matrix effects. However, a study have found that even the dynamic headspace technique suffers from dependence of the calibration data on the sample matrix composition under inexhaustive extraction [21].

1.2.1 DHS

DHS was initially introduced in 1981 [22]. It eliminates the sample matrix effect during the absolute VOC quantification and it is widely applied for the advantages as an environment-friendly, easy to implement and versatile procedure. The central idea of the method is to use a continuous flow of inert gas to extract the volatile organic compounds from a liquid matrix. Extracted VOCs are further concentrated into an adsorbent or cryogenic trap. A scheme is attached in Figure S 1. Since the concentration of the analyte (VOCs) in the headspace decreases exponentially over extraction time, with proper mathematical modeling, it is possible to obtain a total peak area which is

proportional to the total amount of analyte existing in the original sample. Volatile sampling by D-HS is based on the following equation:

$$C_i = C_0 * \mathrm{e}^{-qt} \quad (1.)$$

C_i represent the concentration of an analyte remaining in the sample after a continuous sweeping process depends on the original concentration C_0 and decreases exponentially with time t, q being a constant of proportionality related to the recovery. Chromatographic peak areas A_i are in proportional (k) to analyte concentration C_0:

$$A_i = kC_0 * (1 - \mathrm{e}^{-qt}) \quad (2.)$$

This area can be then linked to concentration by using an appropriate calibration procedure. On the market, a wide variety of chemical traps with different dimensions, composition (e.g., Tenax, Silica Gel, Chromosorb, graphitized carbons (Carbotrap), or carbon molecular sieves), thermal stability and desorption characteristics are available [23]. Traps made of Tenax TA sorbent are the most common, owing to their wide absorption range, high-temperature stability, low water affinity, and long shelf life [24].

DHS sampling has been successfully applied into VOC profiling for wine, beer and other types of fermented beverages [25–27]. DHS sampling has high sensitivity, especially for extraction of highly volatile compounds. The higher recoveries are explained by the transfer of the thermodynamic equilibrium to the gas phase above the sample, the use of a larger volume of extractant gas phase, and the transfer of the entire volatile fraction of the extraction into the GC instrument. In addition, the sensitivity can be adjusted by optimizing the parameters involved in DHS sampling. The availability of automated instruments that allow volatile sampling and GC analysis to be performed simultaneously has facilitated the application of this technique for high-throughput, reproducible analysis of a wide variety of samples. However, this

flexibility requires more complex instrumentation, more maintenance, and complex optimization.

As the VOCs in fermented beverages vary from polar to apolar, and from trace to high concentration, finding suitable parameter setup for flavor profiling is not simple. For these reasons, the use DHS coupled with GC×GC is still limited and is far from general implementation for routine analysis. We hope that ongoing advances in adsorbent development and GC×GC quantification will soon help to explore the full potential of both techniques for the analysis of complex beverage VOCs.

1.2.2 HS-SPME

SPME was developed by Arthur and Pawliszyn in 1990 [28]. It is considered one of the most versatile sample preparation techniques currently available [29]. It has become very popular due to its ease of use, high sensitivity, reproducibility, and low cost. The principle of sorptive extraction can be compared to the partition process of liquid-liquid extraction, as sorbent fiber is below its glass transition point at room temperature and therefore acts as a non-miscible liquid phase. A HS-SPME sampling scheme can be found in Figure S 2. The final equilibrium sampling state can be expressed as:

$$C_O * V_s = C_s * V_s + C_G * V_G + C_F * V_F \quad (3.)$$

C_O is the original concentration of the analyte in the sample with the volume V_S, and C_S is its concentration under equilibrium conditions; C_G and C_F are the analyte concentration in the gas phase (headspace) and the fiber coating with the volumes V_G and V_F. $C_F * V_F$ represents the absorbed amount of a analyte in the fiber coating W_F:

$$C_G = \frac{W_F}{V_F * K_{F/G}} \quad (4.)$$

$$C_S = \frac{W_F}{V_F * K_{F/G} * K_{G/S}} \quad (5.)$$

$K_{G/S}$ and $K_{F/G}$ are the distribution constants between gas phase and sample phase, and between fiber phase and gas phase. Substitute the C_G and C_S in equation 3, the absorbed amount in the fiber coating W_F is:

$$W_F = \frac{C_O * V_S * V_F * K_{F/G} * K_{G/S}}{V_F * K_{F/G} * K_{G/S} + V_G * K_{G/S} + V_S} \quad (6.)$$

HS-SPME requires neither solvents nor previous sample preparation and is highly automatized [30]. Headspace-SPME-GC has been successfully applied to the determination of some flavor compounds in beer, such as organic and fatty acids, alcohols, esters, monophenols and carbonyl compounds [31–36]. A method using three-phases-fiber (Divinylbenzene/Carboxen/Polydimethylsiloxane, DVB-CAR-PDMS) was developed for beer VOC analysis by Gonçalves [37]. In a recent study [38], 329 volatiles were determined from 19 types of lager beers by using DVB-CAR-PDMS fiber for extraction and bidimensional GCMS for analysis. According to the extraction principle, there are some limitations to the SPME extraction. SPME is an equilibrium-based approach that accounts for the matrix effect in the multi-phase system [39]. Reaching equilibrium is recommended to avoid the extraction deviation of fiber exposure time. This requires sufficient sampling time. In a complex system as a fermented beverage, analytes may compete with each other, influence the absorption. Some modifications have been made to remove these weaknesses. Higher extraction temperature improves mass transfer in the multi-phase system [40]. Cold fiber SPME was developed to retain the fiber absorption while heating the sample [41]. The geometry of sorbent phases has been adopted to improve the extraction efficiency and capacity of the SPME device, like SPME Arrow and thin-film SPME (TF-SPME). By ensuring a superior volume of extractive phase and larger surface area for molecule exchange, higher analytical sensitivity and faster extraction speed were achieved [29,40]. Physical assistive technologies, like vacuum and ultrasound, have also been applied to SPME sampling. With the

proper setup, the extraction processing can be faster, milder, more selective and more sensitive [42–45].

1.2.3 DLLME

Liquid-liquid extraction (LLE) is one of the oldest techniques for analyte concentration and matrix separation. LLE also known as solvent extraction separates chemicals from one solution to another based on the different solubility of the solute chemical in two solvents. It is time consuming and requires large amounts of organic solvents. To solve this problem, the dispersive liquid-liquid microextraction (DLLME) technique was proposed in 2006 [46]. In this new concept, the extraction solvent is dispersed into the aqueous sample in the form of fine droplets, rather than keeping the solvent as a discrete phase throughout the extraction process. A further improvement, vortex-assisted liquid-liquid microextraction, was published in 2010 [47]. It attracted attention for its simplicity, rapidity, miniaturization, and low cost [48]. During sample preparation, the sample and the organic phase are mixed together. The mass transfer of analytes from the sample to the organic solvent depends on the product of the overall mass transfer coefficient at the interface by the area of the interface. By vortexing, turbulence is generated and the solvent phase is broken down into small droplets to form an emulsion. As a result, the interfacial area is greatly expanded [49]. The emulsions formed during VALLME are thermodynamically unstable. At the final stage of sample preparation, according to Stoke's Law, a higher gravitational acceleration is introduced to speed up the decomposition of emulsion to form separate phases [50]. A workflow is added as Figure S 3. These apparently simple steps of VALLME involve very complex processes that demand consideration of solvent droplet break-up and coalescence as well as mass transfer of analytes between phases. These processes have been the subject of many past and ongoing investigations in other disciplines, and despite many studies and important contributions, the mechanisms involved remain unclear. Due to this lack of

mechanistic understanding, sample preparation parameters should be empirically defined. However, the final state of this in-solution sampling is a simple equilibrium two phase system. The concentration of an analyte in sorbent phase only relies on the volume of the two phases and distribution coefficient. VALLME has been successfully applied to VOC analysis of beer and wine [19,51,52]. However, basic knowledge of these processes is still lacking, limiting its application. In recent years, other forms of dispersive liquid-liquid microextraction techniques have progressed. Air-assisted liquid-liquid microextraction (AALLME) uses much less volume of an organic solvent, as extraction solvent in the absence of a disperser solvent. Switchable hydrophilic solvents expand our understanding of liquid-liquid extraction [53]. In 2005, Philip G. Jessop and his colleagues introduced carbon dioxide (CO_2)-triggered switchable hydrophilic solvents (SHS), also known as switchable polar solvents (SPS). He used the difference in polarity between the on and off SHS to selectively extract polar and non-polar substances. All these methods belonging to DLLME are becoming increasingly popular and new developments, applications and improvements are expected in the future.

1.2.4 SBSE

The basic principles of Stir Bar Sorptive Extraction (SBSE) are identical to SPME. When PDMS was the only phase available for SPME, the quantitative analysis was only obtained for the volatile compounds whose phase/water partition ratio was larger than 10^5. SBSE was developed to enlarge the application range and boost analytical sensitivity by coating more PDMS on the sorption unit [54]. It can be applied to the headspace sampling, as HS-SPME, construct a three-phase system. But mostly, the sorbent phase is put into the liquid sample, form a two-phase equilibrium system. The workflow is illustrated in Figure S 4. Latterly, multiple SBSE was proposed to increase the application on polar analytes by adding ethylene glycol/silicone phase during the extraction [55]. Both SBSE and multiple SBSE have been successfully applied to beverage

VOC analysis [16,18,19,56]. An interesting approach of SBSE on beverages is the back extraction [56]. After the volatile extraction, analytes on the stir bars can be desorbed by liquid extraction. Back extraction offers possibility from different aspects, for example, performing the analysis with HPLC, more choices on sorbent phases [57], and downstream odour and tasty test.

1.2.5 SPE

Solid-Phase Extraction (SPE) is the most widely used method for organic compounds analysis. It is flexible in solvents, clean-up, concentration and fractionation of organic compounds from several samples. A typical four steps workflow is demonstrated in Figure S 5. It also allows the hydrolysis during the sample extraction [58]. SPE retains the molecule in three measures: partition for small nonpolar analytes [59], adsorption for larger nonpolar molecules [60], and electronic interactions, such as hydrogen bonding, ionic interactions, or π–π bonds for polar and polarizable analytes [61]. Due to these characteristics, a wide range of organic analytes, from non-polar to very polar analytes, can be absorbed with high extraction efficiency. A few studies have been published on applying SPE to beverage analysis. Most of them focus on the quantification of a certain group of VOCs [62–64]. An exciting variety of SPE uses small particles made of nanomaterials for liquid samples. Because of their small size, the nanomaterials have a fairly large surface area and large surface/volume ratio. Consequently, the extraction time is saved. In a study of wine amines analysis, a very short extraction time (1 min) was applied [65]. The major drawbacks of using SPE for beverage VOC analysis is the cumbersome extraction procedure, which can be mitigated by automation. Another expectant variety is molecular imprinting-based solid-phase extraction. Its solid phase is a polymer network with the memory of the template molecules. The polymer saves the synergy of chemical functionality and stereochemistry of target compounds [66]. Increasing the selectivity benefits the quantification of target compounds. Firstly, a larger amount of samples can be applied to improve the analytical sensitivity.

Secondly, less separation power is required, saving the analytical time molecular imprinting-based solid-phase extraction has been successfully applied on ester compounds in aqueous solution [67].

1.3 GC×GC separation

Fermented beverages contain a large number of volatile compounds. For this reason, Direct measure with detector will result in a mixed signal of all the compounds. Hence, chromatography is used to separate VOCs before the detection. Mass spectrometry (MS) can provide information for both identification and quantification in a single measurement. Because targeted compounds are volatiles, a gas chromatographic separation stage (GC) is most commonly used for VOC analysis in beverages. There are two types of approaches according to the study purpose, untargeted VOC profiling and targeted VOC quantification. A bio phenomenon modelling must take into account of many possible chemical variables. Targeted VOC quantification demands fast analytical rate at each measurement. It's not the strength of GC×GC. Non-targeted profiling focuses on the detection of as many VOCs as possible to obtain patterns or fingerprints of a specific beverage production process and for these reasons, it only requires semi-quantitative measurement of concentration. The profiling result is usually used to explore the differences between sample groups [68]. The research question could be, for example, if there are any biomarkers when a newly developed yeast is used in the fermentation [69]. In this type of study, the analytical coverage is crucial to prevent the missing of any substance. The coverage must be improved with better separation. Fermented beverages contain too many volatile compounds which overcrowd the chromatogram of a traditional one dimension GC (1DGC). When signals of VOCs overlap with each other, it's difficult to proceed with the identification and quantification. GC×GC was developed to conquer this problem [70]. In GC×GC analysis, the analytes are firstly separated by one column. The elute of 1D separation is continuously concentrated and injected

into the second column for further separation. The working principle of GC×GC is clearly demonstrated in Figure 1 [71].

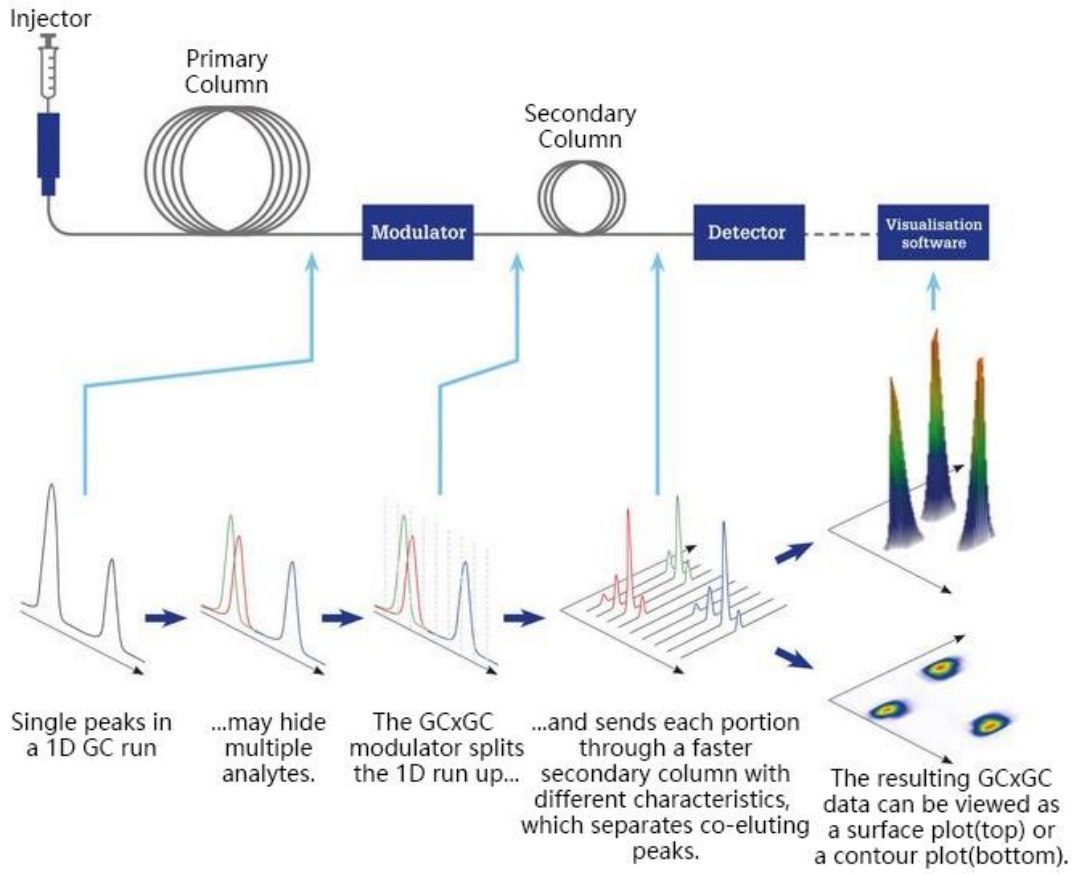


Figure 1 The basic working principle of GC×GC [71]

Method development of GC×GC is much more difficult than that of one-dimensional gas chromatography. First of all, more factors need to be considered, such as the selection of the GC×GC column set and the setting of the regulator, additionally, there are secondary or higher level interactions among these parameters. For this reason, the optimization of GC×GC is complex, and each parameter cannot be simply optimized individually. Currently, most GC×GC studies use the method adapted from 1D GC or rely on the analytical setting recommended by the instrument manufacturer. Although the high-resolution power of GC×GC allows users to get sufficient separation without proper optimization. Improper setup may lead to reduced

separation power, co-eluting peaks, and distort peaks, consequently, annotation and quantification error. The optimal use of the GC×GC relies mainly on the user's knowledge and abilities for method development.

1.3.1 Modulator

In order to achieve high quality GC×GC separations, it is critical to efficiently transfer the eluent from the first dimensional (1D) column to the second dimensional (2D) column while maintaining the separation obtained in the first dimension. The general criterion is that most peaks on 1D dimension have at least three slices, while no wrap-around occurs on 2D dimension [18]. A modulator is needed to periodically trap condensable eluent in the 1D column, refocusing it to a narrow band, and injecting it into the 2D column [72]. Different systems are commercially available, but they are mainly based on thermal or flow modulation [73]. Thermal modulators can be further divided into heater based and cryogenic types. The former collects the 1D eluting analytes at or slightly below the oven temperature and releases them into the 2D column by increasing the temperature. The latter collects analytes at the very low temperature that cryogens create and releases them at or slightly above the oven temperature. Flow based modulators utilize a short collection loop to collect the 1D column eluent, which is subsequently flushed onto the 2D column [74].

Both thermal and flow based modulator have their advantages and disadvantages in areas such as robustness, performance, volatility range, limit of detection, associated costs [75]. There is no single modulator that would be the best choice for all purposes. Since the non-target VOC profiling is the main task of GC×GC in beverage analysis, a modulator should fit the demand of this task. Due to its technical characteristics, thermal modulation is a superior platform for VOC profiling. it completely transfers all analytes from a 1D column to a 2D

column, while achieving narrow injection bands in the second dimension and fast modulation cycles provides high detection sensitivity and allow exceptional trace analysis [76]. This guarantees the coverage of profiling. Besides, the method development of thermal modulation is simpler than flow modulation. The lack of a sample loop eliminates one parameter that needs to be optimized. The thermal modulator requires fewer connections, reducing potential leakage. The major issue with cryogenic modulators is the high running cost associated with the large amount of cryogens consumed daily [77]. However, VOC profiling is required once for each biological study. This cost is affordable.

1.3.2 Column setup

Chromatographic columns are the key components of chromatographic systems. Comprehensive two-dimensional gas chromatography increases the peak capacity and associated separation capability by introducing two different dimensional separation mechanisms on the two columns [78,79]. According to this rule, the concept of separation orthogonality is established. Separation orthogonality is maximized by using independent retention mechanisms in both dimensions. Many users utilize the polarity difference between the two dimensions by using a combination of polar-nonpolar or nonpolar-polar columns. However, later studies have reported that enhanced separation resolution can be obtained by using semipolar columns in the second dimension under nonorthogonal conditions [14,80]. The reason may be that analytes that are retained excessively in one dimension are not retained at all in the other dimension because the chemistry of the two stationary phases is so different [81]. The choice of stationary phase polarity and column order must take into account the sample contained in the analytes. In fermented beverage samples, there are more polar compounds than non-polar compounds. The most suitable column combination is polar-semipolar [82].

Another important factor for column selection, especially the second dimension, is column diameter. To maintain the separation resolution achieved from 1D column (short modulation time), second dimension separation must be significantly faster [78]. A short column between 80 and 200 cm long to completely elute the analytes within 10 seconds. Narrower diameter column in the second dimension was preferred as it allows a higher column efficiency as plate height will change reversely with the column diameter [83]. However, use different column diameters induces flow-mismatch problem, two columns cannot both work under the optimal flowrate [84]. Another problem caused by the reduced column diameter in the second dimension is overloading. For a beverage sample, the contents of major and minor compounds may vary by 9 magnitudes. To cover the trace compounds during VOC profiling, major compounds are over-concentrated. Overloading may lead to distortion of analyte peaks, result in data processing error [85]. Column overload problems can be reduced by increasing the internal diameter of the secondary columns. Today, the same column diameter is typically used for both columns. In this case, both columns can be operated at near-optimum gas flowrate and peak quality will increase significantly. This will lead to more reliable semi quantitative profiling results.

1.3.3 Separation optimization

After the GC×GC hardware is ready, the GCGC parameters need to be set and optimized. The most important GC×GC parameters to optimize are the column-flow, 1st and 2nd oven temperature program, and the modulation setup. These parameters interact with each other and therefore cannot be optimized separately. The optimization requires statistical experimental design and predictive modeling [18,19]. The quality of the GC×GC separation must be measured quantitatively. The first proposed approach is orthogonality measurement [86]. Several different metrics are latterly developed to represent the orthogonality of a given two-dimensional separation, such as Asterisks

equations, Bin counting, and Convex hull [87–90]. However, good global orthogonality is not equal to the good separation suits to the typical need. Direct measuring distances among the targeted peaks reflects the most real. Nearest neighbors' distances (NND) approaches can be modified according to the specific demands of each study [91,92]. Figure 2 is an example chromatogram of common beverage VOCs obtained at optimized GC×GC condition.

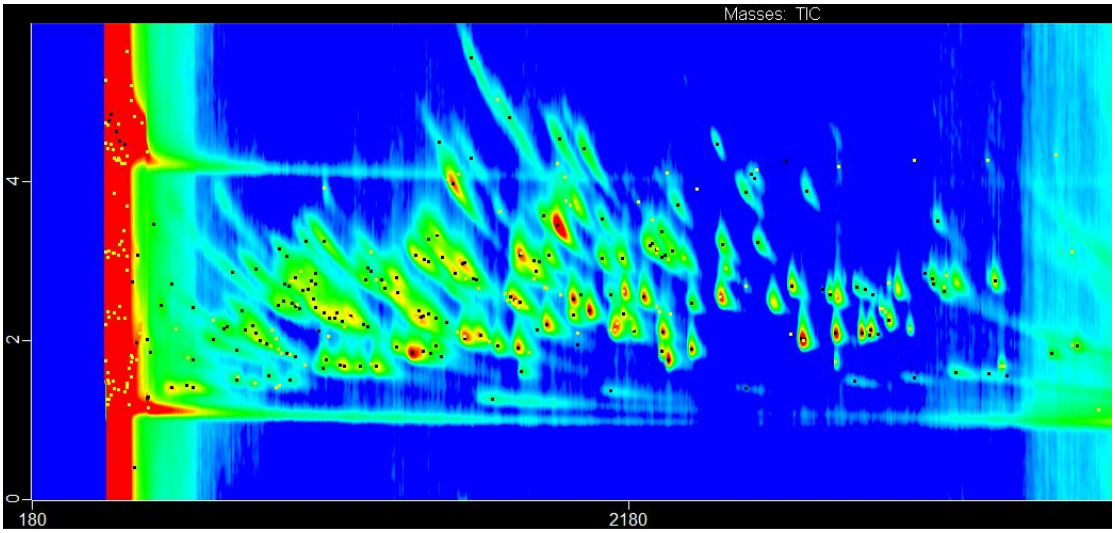


Figure 2 Chromatogram of common beverage VOCs obtained at optimized GC×GC condition. List of VOCs and GC×GC setup can be found in a previous study [19].

1.4 Data processing

At present, the biggest limitation of GC×GC applications is data processing. When a beverage sample is analyzed by GC×GC in combination with a multi-channel detector such as a time-of-flight mass spectrometer (TOFMS), huge amounts of data (hundreds megabits per measurement) are obtained. Especially when multiple sample classes or replicates are required, the size of the data files increases further. In addition to the potentially huge number of data files, the dimensionality of the data increases with additional separation dimensions and multi-channel detection. As a matter of fact, considering the

amount of information contained in a single GC measurement, it is not possible to manually interpret hundreds of different peaks within an acceptable time window. Therefore, the validation of classical chromatographic results such as peak area, resolution and sample fingerprinting is no longer a trivial step. Handling high density data requires the use of data processing techniques and minimal analyst intervention.

The analysis of chromatographic data can be performed in different ways, at the pixel level, on a peak list or peak area basis, by commercial software or by laboratory-developed code [93]. The extraction of useful chemical information from these experimental data requires image signal processing approaches which have to be coupled with linear algebra and statistical concepts to perform a multivariate analysis of the data. These data processing methods can be loosely referred to as chemometric based approaches. A general data processing workflow for GC×GC-MS data and the user input impact for each step are summarized in Figure 3. The complete data processing workflow start from the experiment design, extract the accurate peak information from the GC×GC-MS data, multivariant analysis, and statistical model validation. User has high impact to the validation of peak information and statistical model [94]. Among the different aspects of the workflow, many new ideas are under development in particular in the steps of data preprocessing, peak detection, peak deconvolution, and annotation. However, as often happens, there is no agreement on a “standard” procedure and there is no clear evidence of one approach being clearly better than others. In addition, even with the most popular commercial GC×GC software, a large number of data processing parameters need to be set by the analyst based on the experimental data. For all these reasons, it is of paramount importance for the analyst to have a basic understanding of these processes.

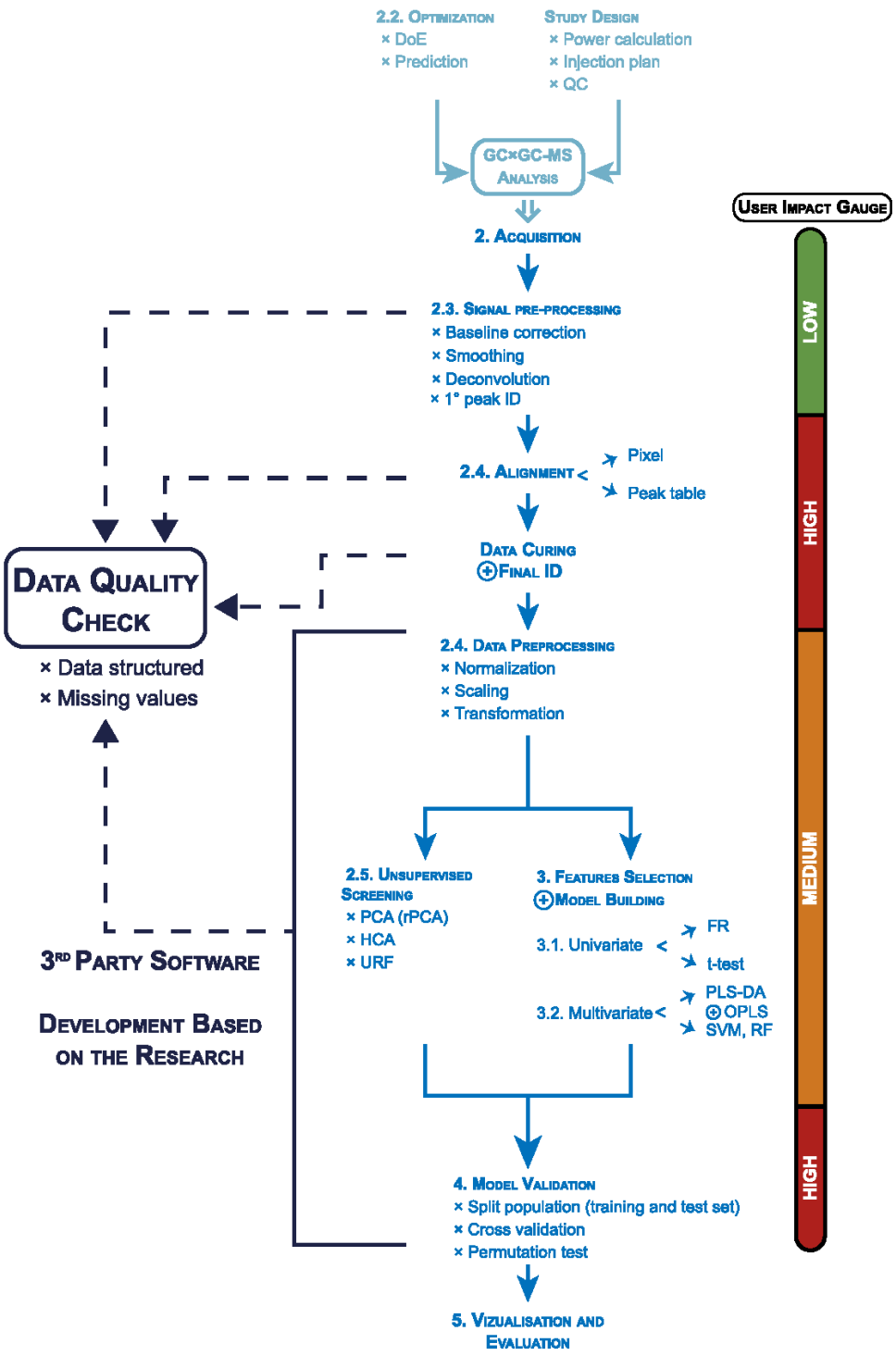


Figure 3 General data processing workflow for GC×GC-MS data and the user input impact for each step [94]

1.4.1 Background correction

To accurately identify and quantify analytes, background correction is required. This long-standing problem was first reported in the 1960s [95]. The central idea is to try to remove the signals that are not coming from the ionization of the molecules and that are present either as high frequency noise or as a smooth background signal. Background correction consists of denoising, smoothing and baseline correction to reduce baseline interference. In the denoising and smoothing procedure, low intensity signals are first removed, regardless of frequency, and then high frequency signals are removed, regardless of their intensity. In the specific case of GC×GC data, electronic noise may be dominant so it should be removed at the beginning. Next, baseline correction can be performed with the aim of subtracting the background signal from the measurement. The drift comes from changes in flow rate and temperature-induced leakage of the stationary phase. Baseline-drift corrections are usually performed by curve fitting or smoothing strategies. Both methods aim to fit a curve through assumed background data points, using loss functions such as the well-known least squares method, or through polynomial fitting [96]. Many background correction methods have been developed for GC×GC, such as direct subtraction of mean blanks, local minima, asymmetric weighted least squares, trilinear decomposition, linear least squares curve fitting methods coupled with moving average smoothing combined with robust orthogonal background correction, and singular value decomposition [97–102]. They can be roughly categorized as parametric or non-parametric. Parametric models are defined as those that assume that the background is produced by a process that can be described by a constant number of parameters and can be modeled by applying linear, quadratic, or polynomial regression. On the other hand, non-parametric methods do not presuppose the shape of the baseline and allow for a flexible number of parameters, the exact number of which depends on the data. However, all these strategies tend to rely on assumptions or premises, which in some cases can lead to incorrect conclusions. When a large number of peak clusters are present, which is common in beverage VOC analysis, the

clusters of true peaks are not easily distinguished from the background signal. Baseline correction becomes increasingly difficult and may lead to the removal of the true signal. Few papers illustrate the mathematical premise of the algorithm in terms of the experimental chemistry of the chromatographic data and test the performance of the model under different extreme conditions, such as heavy co-elution, column saturation, detection saturation, and their coexistence. This makes it difficult for the analyst to choose the most appropriate method based on the experimental chromatogram. In our opinion, the local minima approach respects the experimental data the most. This method is the most respectful of the experimental data and is more suitable for chromatographic data of complex beverage samples.

1.4.2 Peak detection

The aim of peak detection is to distinguish signals of analyte from each other and from the background signal and this step is therefore crucial for correct interpretation of an experiment. GC×GCMS peak detection methods were firstly adapted from the ones used in one dimensional GC. Regardless of the modulation, the “sample” is continuously fed to the detector. And the feedback of the detector is always a time series of the signal intensity, so 1D integration approaches were firstly applied to the vector of detector signals from GC×GC, usually with in-house developed scripts. After all peaklets are integrated, the 2D-plot is then used to manually draw/define polygons around peaklets which belong to the same peak. Finally, the 1D-peaklet areas of all peaklets inside each polygon are integrated manually [103]. These methods can successfully be applied to individual compound detection. However, for untargeted analysis, with thousands of peaklets per chromatogram these methods become unmanageable. To solve the problem, automated methods to cluster all peaklets originating from the individual compounds into 2D peaks was developed [104,105]. Although the algorithm was able to correctly cluster peaklets in most cases, errors may occur on the distorted peak, and unimodality criterion

must be carefully chosen for crowded chromatograms. Cluster error can be avoided by performing the peak detection directly on the 2D demodulated data. During demodulation, the 1D signal vector is rearranged into a 2D matrix. Consequently, multivariate and graphical methods can be applied directly on the 2D plane. Several 2D peak picking methods have been applied to GC×GCMS data, such as local maximum value combined with parallel factor analysis, watershed, and peak shape matching [106–108]. As usual, they all work well with perfect chromatographic data. Unfortunately, a perfect chromatogram cannot be guaranteed. Especially during the beverage VOC profiling, the contents of major and minor compounds may vary by 9 magnitudes. It's common to see deformed peaks as the result of column or detector saturation. When the saturation occurs, these peak picking strategies show their limitations producing false peak splitting which result in incorrect deconvolution. At this stage, these methods can be used in combination, and the results obtained are validated against each other [85].

1.4.3 Peak annotation

After peak detection, peak annotation can be performed relying on using multiple factors. GC×GCMS data provides several identification factors: two retention times, an electron impact-based fragmentation spectrum, and, in some cases, also high resolution or soft ionization MS. Combining the chromatographic retention time and MS information is however always required to achieve level 1 annotation for VOC profiling [109]. Peaks may be identified based on their known primary, secondary retention times, and retention-time windows or by using so-called 2D templates. Many software also includes linear retention indices (LRI) calculation, which can be used to cross-check the annotation with an analyte-retention-time database. By using PEG homologs together with n-alkanes, 2D retention indices system can be created [110]. Retention time and/or index modeling is an area in which a key focus is to improve the identification of unknown peaks [74,111,112]. The identification

generally relies on electron impact ionization mass spectra libraries (NIST, Wiley, etc.), sometimes also on the mass accuracy calculation (when an HRMS detector is used). Many algorithms have been published to compare the mass spectrum of a chromatographic peak with standard MS recorded in the libraries. Their performances have been extensively studied [113–118]. All this requires is a simple setting of the appropriate match thresholds, all of which provide accurate annotation results. In the majority of cases, the error in MS identification is mostly caused by inaccurate MS construction during the peak detection step. A possible solution is smart template matching [119]. Since prior knowledge is available on beverage VOCs. Besides, most VOCs are constantly presented in all types of fermented drinks. For these reasons, it's not necessary to perform a complete untargeted analysis. Building a library for beverage VOCs and matching their template to the chromatogram peak can dramatically reduce annotation error. Moreover, in case a new compound is found, its template can be simply added to the current library and recognized in future data processing.

1.4.4 Cross-class comparison

With VOC profiling results, analysts commonly seek to discover relevant chemical features that differentiate (or are common between) sample classes, and correlate them to chemical/physical properties associated with sample classes. This cross-class analysis gives rise to many applications, including sample classification, class comparison, chemical fingerprinting, and chemical/biomarker discovery. In case of beverage study, the above applications can be specified as difference on primary materials, sensory properties, microbial activity, spoilage, ageing, etc. For example, monoterpenes are important contributors to the aroma from primary materials: grape for wine and hops for beer. High levels of linalool and α-terpineol are well documented in Muscat wines [120]. Linalool , at the same time, is an indicator of using hops during beer making [121]. (Z)-rose oxide is important in Gewürztraminer wines [122]. The flavanoids have correlations to the mouth feel of red wines [123]. It is heavily

influenced by interactions with cell wall material and winemaking practices [124,125]. Furan derivatives, including furfural, 5-methylfurfural, are typically dependent on the degree of toasting and contact surface area of the oak [126]. Sotolon is a chiral furanone responsible for premature-aging flavor in dry white wines [127]. Esters and acetates contribute to sweet fruity flavor in wines. Changes in ester concentrations in wines during maturation [34]. Higher alcohols and their corresponding acetate esters are metabolites produced from the secondary metabolism of amino acids via the Ehrlich pathway [128]. Aldehydes and ketones, particularly diketones, are considered off-flavors. Diacetyl and 2, 3-pentadione – by-products from valine and isoleucine biosynthesis in yeasts [129].

As we know, a GC×GC chromatogram can easily contain hundreds of peaks. Mining the useful information manually in such “big data” can quickly become overwhelming for large scale untargeted research. Chemometric strategies that involving feature selection can improve the success of downstream exploration by removing irrelevant chromatographic information and reducing the size of the data files that need manual inspection. A common chemometric cross-class analysis consists of two steps, chromatogram/peak alignment and statistical exploration [130].

Data alignment is a crucial part of cross-class analysis as retention-time shifts between analyses are not uncommon. These shifts usually originate from column degradation, carrier gas pressure and temperature variations, and maintenance on the instrument. In large-scale studies, Unstable chromatography may cause a peak to be detected at different retention times in different measurements. Signal fluctuations at the MS detector can result in peak deconvolution difference on those peaks. Consequently, a peak has different retention times and is quantified in different ion channels. This can create problems for peak rescaling and peak area comparison within an omic

study. Hence, data alignment is particularly important preprocessing steps prior to cross-sample analysis. This alignment is performed either before peak detection on pixel/tile level or after the peak detection on peak table level. In the former case, the entire chromatogram is used for the alignment. Various alignment algorithms are available including wraparound correction, such as PARAFAC based modulation time shift correction, windowed rank minimization, indexing schemes cylindrical mapping [131–134]. The alignment can be achieved from pixel/tile perspective. With binning and tiling schemes of the chromatogram, overall size of the data files is reduced [135]. This computational approach allows to speed up the processing time and reduce computer power needs [136]. During peak table based alignment, assigning a unique identifier to each peak and assuming this to be consistent across all chromatograms. This approach can be completely manually or automatically [137,138].

Extracting the valuable information in a high dimension data set consisting of hundreds of chemical features and several sample properties is difficult. To simplify this process, statistical tools must be used to reduce the data dimension. One way of selecting the significant features is applying univariate statistics, such as Fisher ratio (FR), parametric or nonparametric testing. Univariate statistics describe the variation in a single variable, which is assumed independently from other variables. Fisher ratio approach is popular and has been widely used. Its concept was taken from ANOVA analysis. Briefly, FR is calculated by dividing the between-class variance by the within-class variance. The obtained metric can be compared with the F critical value to identify compounds which vary in the group significantly [139]. Originally, it was calculated based on the result in the peak table after the peak detection. To avoid the interference from peak detecting error, it has now evolved to image based comparison, tile-base FR [140]. It is important to mention that, many chemicals in the beverage are not independent, so simply applying the FR analysis may discard many relevant features or overlook important features. One should have a scientific background of the sample to supervise the FR test. On the other

hand, in multivariate statistical analysis, multiple correlated variables can be combined and represented by a new “latent” vector. In such a way, data dimension is reduced and the most informative variables are extracted. Still in presence of more variable than samples, correlation can arise also by "chance". Even with multivariate methods, validation and domain specific knowledge are required to confirm the results of the experiments. Principal component analysis based tools determine the sources of the greatest variance, and are widely used in GC×GC result exploration [141]. PCA can also be applied on graphic data. With proper binning size, the data density is reduced and misalignment is mitigated [142]. Recently, machine learning methods have been compared based on wine data [143]. It is expected that more artificial intelligence technologies will be applied to data mining in the field of GC×GC in the future.

1.5 Conclusion

This article provides an overview of the analytical approach of GC×GC in beverage VOC profiling. Common sample extraction method were covered. These techniques include dynamic headspace (DHS), dispersive liquid-liquid microextraction (DLLME), solid phase microextraction (SPME), stir bar extraction (SBSE), and solid phase extraction (SPE). Their basic extraction principle, (dis)advantages, and future perspectives on beverage VOC profiling were discussed. For DHS, the sensitivity can be adjusted by optimizing the of parameters involved. However, this flexibility requires more complex instrumentation, more maintenance, and complex optimization. Find a universal parameter setup for beverage VOC profiling is not simple. The theory of DLLME requires cross disciplines knowledge. Mechanisms involved remain unclear. Many new methods belongs to DLLME have progressed, such as air-assisted liquid-liquid microextraction (AALLME) and switchable hydrophilic solvents. These methods are becoming increasingly popular and new developments, applications and improvements are expected in the future. Sample extraction efficiency limits the application of SPME. Its mass transfer efficiency can be

improved by higher extraction temperature, larger contact surface and assistive technologies likes vacuum and ultrasound. With the proper setup, the extraction processing can be faster, milder, more selective and more sensitive. SBSE has only two sorbent phases are commercially available, PDMS and ethylene glycol/silicone phase. Back extraction is an interesting application of SBSE. It offers possibility of perform the analysis with HPLC, more choices on sorbent phases, and downstream odor and tasty test. SPE retains the molecule in varies mechanisms. A wide range of organic analytes, from non-polar to very polar analytes, can be absorbed with high extraction efficiency. The major drawbacks of using SPE for beverage VOC analysis is the cumbersome extraction procedure. Automation is required. Molecular imprinting-based solid-phase extraction has a bright future. It increasing the selectivity benefits the quantification of target compounds.

Extracted beverage VOCs must be well separated by the GC×GC system. A fine GC×GC method development involves modulator selection, column combination and parameter optimization. Thermal modulator provides high detection sensitivity and allow exceptional trace analysis. Since the analytes coverage is the most important factor of in beverage VOC profiling, thermal modulation is a better choice. Columns are the key components of chromatographic systems. Better column orthogonality doesn't mean enhanced separation resolution. Experimental evidence shows that semipolar columns in the second dimension is the right choice. In fermented beverages, there are more polar compounds than non-polar compounds. The most suitable column combination is polar-semipolar. Narrower diameter column in the second dimension was preferred for higher column efficiency, and short elute time. However, due to the column overload the data processing problems cause by it, use same column diameter is recommended. GC×GC parameters such as, column-flow, 1^{st} and 2^{nd} oven temperature program, and the modulation setup must be optimized. These parameters interact with each other therefore statistical prediction model is required. Good global orthogonality is not equal

to the separation suits to the typical need. Direct measuring distances among the targeted peaks reflects the most real.

Mining valuable information from a large amount of high-dimensional GC×GC data is not an easy task. Analysts need a certain background in signal processing and statistics to choose the appropriate data processing methods and adjust the parameters. Conventional data processing procedures are background correction, peak detection, peak annotation, and cross-class comparison. Background correction is used to eliminate the effects of detector noise and instrument variation on peak quantification. The complex composition of beverages makes their analytical chromatography susceptible to non-ideal conditions such as saturation. We recommend local minima approach. The aim of peak detection is to distinguishing signals of analyte from each other and from the background signal. Several methods based on different mechanisms have been developed. Given the complexity of chromatograms, the simultaneous use of different methods and their validation against each other is an appropriate approach. The accuracy of peak annotation depends heavily on the quality of the peak detection. The development of a template database for beverage VOCs can greatly improve robustness and accuracy. Cross-class comparison requires data alignment and data dimensionality reduction. In order to avoid being affected by errors in other data processing steps, comparison based on raw image data is the current trend.

References

(1) Patrick, C. H. *Alcohol, Culture, and Society*; Duke University Press, 1952.

(2) McGovern, P.; Jalabadze, M.; Batiuk, S.; Callahan, M. P.; Smith, K. E.; Hall, G. R.; Kvavadze, E.; Maghradze, D.; Rusishvili, N.; Bouby, L.; Failla, O.; Cola, G.; Mariani, L.; Boaretto, E.; Bacilieri, R.; This, P.; Wales, N.;

Lordkipanidze, D. Early Neolithic Wine of Georgia in the South Caucasus. *Proc. Natl. Acad. Sci.* **2017**, *114* (48), E10309–E10318. https://doi.org/10.1073/pnas.1714728114.

(3) Tucker, A. The Beer Archaeologist. *SMITHSONIAN MAGAZINE*. 2011.

(4) *The Oxford Companion to Wine*, Fourth edition.; Robinson, J., Harding, J., Smart, R. E., Lavigne, V., Dubourdieu, D., Eds.; Oxford University Press: Oxford, 2015.

(5) OIV. *Report on the State of the Vitiviniculture World Market*; news.reseau-concept.net, 2019.

(6) Nelson, M. *The Barbarian's Beverage: A History of Beer in Ancient Europe*, 1. publ.; Routledge: London, 2005.

(7) Prasannan, A. *Alcoholic Beverages Market by Type (Beer, Distilled Spirits, Wine, and Others) and Distribution Channel (Convenience Stores, On Premises, Liquor Stores, Grocery Shops, Internet Retailing, and Supermarkets): Global Opportunity Analysis and Industry Forecast, 2018 - 2025*; A01711; alliedmarketresearch, 2018.

(8) Fortunebusinessinsights. *Beer Market Size, Share & Industry Analysis*; FBI102489; 2020.

(9) Fortunebusinessinsights. *Wine Market Size, Share & Industry Analysis*; FBI102836; 2020.

(10) Jaeger, S. R.; Worch, T.; Phelps, T.; Jin, D.; Cardello, A. V. Preference Segments among Declared Craft Beer Drinkers: Perceptual, Attitudinal and Behavioral Responses Underlying Craft-Style vs. Traditional-Style Flavor Preferences. *Food Qual. Prefer.* **2020**, *82*, 103884. https://doi.org/10.1016/j.foodqual.2020.103884.

(11) Nykänen, L.; Suomalainen, H. *Aroma of Beer, Wine, and Distilled Alcoholic Beverages*; Handbuch der Aromaforschung; D. Reidel ; Distributors

for the U.S.A. and Canada, Kluwer Boston: Dordrecht, Holland ; Boston, U.S.A. : Hingham, MA, 1983.

(12) Ravasio, D.; Carlin, S.; Boekhout, T.; Groenewald, M.; Vrhovsek, U.; Walther, A.; Wendland, J. Adding Flavor to Beverages with Non-Conventional Yeasts. *Fermentation* **2018**, *4* (1), 15. https://doi.org/10.3390/fermentation4010015.

(13) Kobayashi, M.; Shimizu, H.; Shioya, S. Beer Volatile Compounds and Their Application to Low-Malt Beer Fermentation. *J. Biosci. Bioeng.* **2008**, *106* (4), 317–323. https://doi.org/10.1263/jbb.106.317.

(14) Cordero, C.; Kiefl, J.; Schieberle, P.; Reichenbach, S. E.; Bicchi, C. Comprehensive Two-Dimensional Gas Chromatography and Food Sensory Properties: Potential and Challenges. *Anal. Bioanal. Chem.* **2015**, *407* (1), 169–191. https://doi.org/10.1007/s00216-014-8248-z.

(15) Andrés-Iglesias, C.; Montero, O.; Sancho, D.; Blanco, C. A. New Trends in Beer Flavour Compound Analysis. *J. Sci. Food Agric.* **2015**, *95* (8), 1571–1576. https://doi.org/10.1002/jsfa.6905.

(16) Horák, T.; Čulík, J.; Jurková, M.; Čejka, P.; Kellner, V. Determination of Free Medium-Chain Fatty Acids in Beer by Stir Bar Sorptive Extraction. *J. Chromatogr. A* **2008**, *1196–1197*, 96–99. https://doi.org/10.1016/j.chroma.2008.05.014.

(17) Liu, Y.; Su, H.; Song, H.-L. Comparison of Four Extraction Methods, SPME, DHS, SAFE, Versus SDE, for the Analysis of Flavor Compounds in Natto. *Food Anal. Methods* **2018**, *11* (2), 343–354. https://doi.org/10.1007/s12161-017-1005-0.

(18) Stefanuto, P.-H.; Perrault, K. A.; Dubois, L. M.; L'Homme, B.; Allen, C.; Loughnane, C.; Ochiai, N.; Focant, J.-F. Advanced Method Optimization for Volatile Aroma Profiling of Beer Using Two-Dimensional Gas Chromatography Time-of-Flight Mass Spectrometry. *J. Chromatogr. A* **2017**, *1507*, 45–52. https://doi.org/10.1016/j.chroma.2017.05.064.

(19) Zhang, P.; Carlin, S.; Lotti, C.; Mattivi, F.; Vrhovsek, U. On Sample Preparation Methods for Fermented Beverage VOCs Profiling by GCxGC-TOFMS. *Metabolomics* **2020**, *16* (10), 102. https://doi.org/10.1007/s11306-020-01718-7.

(20) Kolb, B.; Ettre, L. S. *Static Headspace-Gas Chromatography: Theory and Practice*, 2nd ed.; Wiley: Hoboken, N.J, 2006.

(21) Dunn, W. B.; Townshend, A.; Dunn, W. B.; Green, J. D. Comparison of Total Vaporisation and Dynamic Headspace Techniques Combined with Direct Mass Spectrometric Detection for the On-Line Analysis of Liquid Process Streams. *The Analyst* **1998**, *123* (2), 343–348. https://doi.org/10.1039/a705516a.

(22) Kolb, B.; Auer, M.; Pospisil, P. *Angewandte Chromatographie-Applied Chromatography*; No. 35E; Bodenseewerk Perkin-Elmer & Co.: Überlingen, FRG, 1981.

(23) Ross, C. F. Headspace Analysis. In *Comprehensive Sampling and Sample Preparation*; Elsevier, 2012; pp 27–50. https://doi.org/10.1016/B978-0-12-381373-2.00036-3.

(24) Soria, A. C.; García-Sarrió, M. J.; Sanz, M. L. Volatile Sampling by Headspace Techniques. *TrAC Trends Anal. Chem.* **2015**, *71*, 85–99. https://doi.org/10.1016/j.trac.2015.04.015.

(25) Wen, Y.; Lopez, R.; Ferreira, V. An Automated Gas Chromatographic-Mass Spectrometric Method for the Quantitative Analysis of the Odor-Active Molecules Present in the Vapors Emanated from Wine. *J. Chromatogr. A* **2018**, *1534*, 130–138. https://doi.org/10.1016/j.chroma.2017.12.064.

(26) Aberl, A.; Coelhan, M. Determination of Volatile Compounds in Different Hop Varieties by Headspace-Trap GC/MS—In Comparison with Conventional Hop Essential Oil Analysis. *J. Agric. Food Chem.* **2012**, *60* (11), 2785–2792. https://doi.org/10.1021/jf205002p.

(27) Liu, J.; Toldam-Andersen, T. B.; Petersen, M. A.; Zhang, S.; Arneborg, N.; Bredie, W. L. P. Instrumental and Sensory Characterisation of Solaris White Wines in Denmark. *Food Chem.* **2015**, *166*, 133–142. https://doi.org/10.1016/j.foodchem.2014.05.148.

(28) Arthur, C. L.; Pawliszyn, Janusz. Solid Phase Microextraction with Thermal Desorption Using Fused Silica Optical Fibers. *Anal. Chem.* **1990**, *62* (19), 2145–2148. https://doi.org/10.1021/ac00218a019.

(29) Reyes-Garcés, N.; Gionfriddo, E.; Gómez-Ríos, G. A.; Alam, Md. N.; Boyacı, E.; Bojko, B.; Singh, V.; Grandy, J.; Pawliszyn, J. Advances in Solid Phase Microextraction and Perspective on Future Directions. *Anal. Chem.* **2018**, *90* (1), 302–360. https://doi.org/10.1021/acs.analchem.7b04502.

(30) Zhu, J.; Chai, X.-. Some Recent Developments in Headspace Gas Chromatography. *Curr. Anal. Chem.* **2005**, *1* (1), 79–83. https://doi.org/10.2174/1573411052948488.

(31) Cajka, T.; Riddellova, K.; Tomaniova, M.; Hajslova, J. Recognition of Beer Brand Based on Multivariate Analysis of Volatile Fingerprint. *J. Chromatogr. A* **2010**, *1217* (25), 4195–4203. https://doi.org/10.1016/j.chroma.2009.12.049.

(32) Charry-Parra, G.; DeJesus-Echevarria, M.; Perez, F. J. Beer Volatile Analysis: Optimization of HS/SPME Coupled to GC/MS/FID. *J. Food Sci.* **2011**, *76* (2), C205–C211. https://doi.org/10.1111/j.1750-3841.2010.01979.x.

(33) Kobayashi, M.; Shimizu, H.; Shioya, S. Beer Volatile Compounds and Their Application to Low-Malt Beer Fermentation. *J. Biosci. Bioeng.* **2008**, *106* (4), 317–323. https://doi.org/10.1263/jbb.106.317.

(34) Riu-Aumatell, M.; Miró, P.; Serra-Cayuela, A.; Buxaderas, S.; López-Tamames, E. Assessment of the Aroma Profiles of Low-Alcohol Beers Using HS-SPME–GC-MS. *Food Res. Int.* **2014**, *57*, 196–202. https://doi.org/10.1016/j.foodres.2014.01.016.

(35) Saison, D.; De Schutter, D. P.; Delvaux, F.; Delvaux, F. R. Optimisation of a Complete Method for the Analysis of Volatiles Involved in the Flavour

Stability of Beer by Solid-Phase Microextraction in Combination with Gas Chromatography and Mass Spectrometry. *J. Chromatogr. A* **2008**, *1190* (1–2), 342–349. https://doi.org/10.1016/j.chroma.2008.03.015.

(36) Silva, G. A. da; Augusto, F.; Poppi, R. J. Exploratory Analysis of the Volatile Profile of Beers by HS–SPME–GC. *Food Chem.* **2008**, *111* (4), 1057–1063. https://doi.org/10.1016/j.foodchem.2008.05.022.

(37) Gonçalves, J. L.; Figueira, J. A.; Rodrigues, F. P.; Ornelas, L. P.; Branco, R. N.; Silva, C. L.; Câmara, J. S. A Powerful Methodological Approach Combining Headspace Solid Phase Microextraction, Mass Spectrometry and Multivariate Analysis for Profiling the Volatile Metabolomic Pattern of Beer Starting Raw Materials. *Food Chem.* **2014**, *160*, 266–280. https://doi.org/10.1016/j.foodchem.2014.03.065.

(38) Martins, C.; Brandão, T.; Almeida, A.; Rocha, S. M. Enlarging Knowledge on Lager Beer Volatile Metabolites Using Multidimensional Gas Chromatography. *Foods* **2020**, *9* (9), 1276. https://doi.org/10.3390/foods9091276.

(39) Lashgari, M.; Yamini, Y. An Overview of the Most Common Lab-Made Coating Materials in Solid Phase Microextraction. *Talanta* **2019**, *191*, 283–306. https://doi.org/10.1016/j.talanta.2018.08.077.

(40) *Handbook of Solid Phase Microextraction*; Elsevier, 2012. https://doi.org/10.1016/C2011-0-04297-7.

(41) Wang, Y.-C.; Wang, J.-L.; Shu, Y.-Y. Purge-Assisted and Temperature-Controlled Headspace Solid-Phase Microextraction Combined with Gas Chromatography–Mass Spectrometry for Determination of Six Common Phthalate Esters in Aqueous Samples. *J. Food Meas. Charact.* **2020**, *14* (4), 1833–1841. https://doi.org/10.1007/s11694-020-00430-3.

(42) Mihaljević Žulj, M.; Maslov, L.; Tomaz, I.; Jeromel, A. Determination of 2-Aminoacetophenone in White Wines Using Ultrasound Assisted SPME

Coupled with GC-MS. *J. Anal. Chem.* **2015**, *70* (7), 814–818. https://doi.org/10.1134/S1061934815070102.

(43) Orazbayeva, D.; Kenessov, B.; Psillakis, E.; Nassyrova, D.; Bektassov, M. Determination of Transformation Products of Unsymmetrical Dimethylhydrazine in Water Using Vacuum-Assisted Headspace Solid-Phase Microextraction. *J. Chromatogr. A* **2018**, *1555*, 30–36. https://doi.org/10.1016/j.chroma.2018.04.048.

(44) Vakinti, M.; Mela, S.-M.; Fernández, E.; Psillakis, E. Room Temperature and Sensitive Determination of Haloanisoles in Wine Using Vacuum-Assisted Headspace Solid-Phase Microextraction. *J. Chromatogr. A* **2019**, *1602*, 142–149. https://doi.org/10.1016/j.chroma.2019.03.047.

(45) Yang, Y.; Chu, G.; Zhou, G.; Jiang, J.; Yuan, K.; Pan, Y.; Song, Z.; Li, Z.; Xia, Q.; Lu, X.; Xiao, W. Rapid Determination of the Volatile Components in Tobacco by Ultrasound-Microwave Synergistic Extraction Coupled to Headspace Solid-Phase Microextraction with Gas Chromatography-Mass Spectrometry: Sample Preparation. *J. Sep. Sci.* **2016**, *39* (6), 1173–1181. https://doi.org/10.1002/jssc.201501185.

(46) Rezaee, M.; Assadi, Y.; Milani Hosseini, M.-R.; Aghaee, E.; Ahmadi, F.; Berijani, S. Determination of Organic Compounds in Water Using Dispersive Liquid–Liquid Microextraction. *J. Chromatogr. A* **2006**, *1116* (1–2), 1–9. https://doi.org/10.1016/j.chroma.2006.03.007.

(47) Yiantzi, E.; Psillakis, E.; Tyrovola, K.; Kalogerakis, N. Vortex-Assisted Liquid–Liquid Microextraction of Octylphenol, Nonylphenol and Bisphenol-A. *Talanta* **2010**, *80* (5), 2057–2062. https://doi.org/10.1016/j.talanta.2009.11.005.

(48) Psillakis, E.; Kalogerakis, N. Developments in Single-Drop Microextraction. *TrAC Trends Anal. Chem.* **2002**, *21* (1), 54–64. https://doi.org/10.1016/S0165-9936(01)00126-1.

(49) *Turbulence Phenomena*; Elsevier, 1972. https://doi.org/10.1016/B978-0-12-206070-0.X5001-9.

(50) Gupta, A.; Eral, H. B.; Hatton, T. A.; Doyle, P. S. Nanoemulsions: Formation, Properties and Applications. *Soft Matter* **2016**, *12* (11), 2826–2841. https://doi.org/10.1039/C5SM02958A.

(51) Russo, M. V.; Avino, P.; Perugini, L.; Notardonato, I. Fast Analysis of Nine PAHs in Beer by Ultrasound-Vortex-Assisted Dispersive Liquid–Liquid Micro-Extraction Coupled with Gas Chromatography-Ion Trap Mass Spectrometry. *RSC Adv.* **2016**, *6* (17), 13920–13927. https://doi.org/10.1039/C5RA24873F.

(52) Ortega, C.; López, R.; Cacho, J.; Ferreira, V. Fast Analysis of Important Wine Volatile Compounds. *J. Chromatogr. A* **2001**, *923* (1–2), 205–214. https://doi.org/10.1016/S0021-9673(01)00972-4.

(53) Jessop, P. G.; Heldebrant, D. J.; Li, X.; Eckert, C. A.; Liotta, C. L. Reversible Nonpolar-to-Polar Solvent. *Nature* **2005**, *436* (7054), 1102–1102. https://doi.org/10.1038/4361102a.

(54) Baltussen, E.; Sandra, P.; David, F.; Cramers, C. Stir Bar Sorptive Extraction (SBSE), a Novel Extraction Technique for Aqueous Samples: Theory and Principles. *J. Microcolumn Sep.* **1999**, *11* (10), 737–747.

(55) Berrou, K.; Dunyach-Remy, C.; Lavigne, J.-P.; Roig, B.; Cadiere, A. Multiple Stir Bar Sorptive Extraction Combined with Gas Chromatography-Mass Spectrometry Analysis for a Tentative Identification of Bacterial Volatile and/or Semi-Volatile Metabolites. *Talanta* **2019**, *195*, 245–250. https://doi.org/10.1016/j.talanta.2018.11.042.

(56) Castro, L. F.; Ross, C. F. Determination of Flavour Compounds in Beer Using Stir-Bar Sorptive Extraction and Solid-Phase Microextraction: Determination of Flavor Compounds in Beer. *J. Inst. Brew.* **2015**, *121* (2), 197–203. https://doi.org/10.1002/jib.219.

(57) Gilart, N.; Marcé, R. M.; Borrull, F.; Fontanals, N. New Coatings for Stir-Bar Sorptive Extraction of Polar Emerging Organic Contaminants. *TrAC Trends Anal. Chem.* **2014**, *54*, 11–23. https://doi.org/10.1016/j.trac.2013.10.010.

(58) Allen, D.; Bui, A. D.; Cain, N.; Rose, G.; Downey, M. Analysis of Free and Bound Phenolics in Wine and Grapes by GC–MS after Automated SPE. *Anal. Bioanal. Chem.* **2013**, *405* (30), 9869–9877. https://doi.org/10.1007/s00216-013-7405-0.

(59) Martire, D. E.; Boehm, R. E. Unified Theory of Retention and Selectivity in Liquid Chromatography. 2. Reversed-Phase Liquid Chromatography with Chemically Bonded Phases. *J. Phys. Chem.* **1983**, *87* (6), 1045–1062. https://doi.org/10.1021/j100229a025.

(60) Gritti, F.; Guiochon, G. Adsorption Mechanism in Reversed-Phase Liquid Chromatography. *J. Chromatogr. A* **2006**, *1115* (1–2), 142–163. https://doi.org/10.1016/j.chroma.2006.02.095.

(61) Pap, T.; Horváth, V.; Tolokán, A.; Horvai, G.; Sellergren, B. Effect of Solvents on the Selectivity of Terbutylazine Imprinted Polymer Sorbents Used in Solid-Phase Extraction. *J. Chromatogr. A* **2002**, *973* (1–2), 1–12. https://doi.org/10.1016/S0021-9673(02)01084-1.

(62) Babaee, S.; Daneshfar, A.; Khezeli, T. Determination of Carboxylic Acids in Non-Alcoholic Beer Samples by an Ultrasonic-Assisted Dispersive Micro-Solid Phase Extraction Based on Ni/Cu-Al Layered Double Hydroxide Nanocomposites Followed by Gas Chromatography. *Ultrason. Sonochem.* **2017**, *34*, 847–855. https://doi.org/10.1016/j.ultsonch.2016.07.023.

(63) Deng, Z.; Hu, K.; Zhang, Y.; Zhao, W.; Wang, F.; Guo, L.; Zhang, W.; He, J.; Huang, Y.; Zhang, S. On-Cartridge Derivatisation Using a Calixarene Solid-Phase Extraction Sorbent for Facile, Sensitive and Fast Determination of Formaldehyde in Beer. *Food Chem.* **2016**, *211*, 314–319. https://doi.org/10.1016/j.foodchem.2016.05.069.

(64) Zheng, S.-J.; Wang, Y.-L.; Liu, P.; Zhang, Z.; Yu, L.; Yuan, B.-F.; Feng, Y.-Q. Stable Isotope Labeling-Solid Phase Extraction-Mass Spectrometry Analysis for Profiling of Thiols and Aldehydes in Beer. *Food Chem.* **2017**, *237*, 399–407. https://doi.org/10.1016/j.foodchem.2017.05.090.

(65) Milheiro, J.; Ferreira, L. C.; Filipe-Ribeiro, L.; Cosme, F.; Nunes, F. M. A Simple Dispersive Solid Phase Extraction Clean-up/Concentration Method for Selective and Sensitive Quantification of Biogenic Amines in Wines Using Benzoyl Chloride Derivatisation. *Food Chem.* **2019**, *274*, 110–117. https://doi.org/10.1016/j.foodchem.2018.08.116.

(66) Pichon, V.; Chapuis-Hugon, F. Role of Molecularly Imprinted Polymers for Selective Determination of Environmental Pollutants—A Review. *Anal. Chim. Acta* **2008**, *622* (1–2), 48–61. https://doi.org/10.1016/j.aca.2008.05.057.

(67) Yang, R.; Liu, Y.; Yan, X.; Liu, S. Simultaneous Extraction and Determination of Phthalate Esters in Aqueous Solution by Yolk-Shell Magnetic Mesoporous Carbon-Molecularly Imprinted Composites Based on Solid-Phase Extraction Coupled with Gas Chromatography–Mass Spectrometry. *Talanta* **2016**, *161*, 114–121. https://doi.org/10.1016/j.talanta.2016.08.037.

(68) Cevallos Cevallos, J. M.; Reyes-De-Corcuera, J. I.; Etxeberria, E.; Danyluk, M. D.; Rodrick, G. E. Metabolomic Analysis in Food Science: A Review. *Trends Food Sci. Technol.* **2009**, *20* (11–12), 557–566. https://doi.org/10.1016/j.tifs.2009.07.002.

(69) Giannakou, K.; Visinoni, F.; Zhang, P.; Nathoo, N.; Jones, P.; Cotterrell, M.; Vrhovsek, U.; Delneri, D. Biotechnological Exploitation of Saccharomyces Jurei and Its Hybrids in Craft Beer Fermentation Uncovers New Aroma Combinations. *Food Microbiol.* **2021**, *100*, 103838. https://doi.org/10.1016/j.fm.2021.103838.

(70) Ong, R. C. Y.; Marriott, P. J. A Review of Basic Concepts in Comprehensive Two-Dimensional Gas Chromatography. *J. Chromatogr. Sci.* **2002**, *40* (5), 276–291.

(71) McGregor, L.; Barden, D. *Easy as GCxGC*; The Cannabis Scientist, 2019.

(72) Edwards, M.; Mostafa, A.; Górecki, T. Modulation in Comprehensive Two-Dimensional Gas Chromatography: 20 Years of Innovation. *Anal. Bioanal. Chem.* **2011**, *401* (8), 2335–2349. https://doi.org/10.1007/s00216-011-5100-6.

(73) Prebihalo, S. E.; Berrier, K. L.; Freye, C. E.; Bahaghighat, H. D.; Moore, N. R.; Pinkerton, D. K.; Synovec, R. E. Multidimensional Gas Chromatography: Advances in Instrumentation, Chemometrics, and Applications. *Anal. Chem.* **2018**, *90* (1), 505–532. https://doi.org/10.1021/acs.analchem.7b04226.

(74) Seeley, J. V.; Micyus, N. J.; Bandurski, S. V.; Seeley, S. K.; McCurry, J. D. Microfluidic Deans Switch for Comprehensive Two-Dimensional Gas Chromatography. *Anal. Chem.* **2007**, *79* (5), 1840–1847. https://doi.org/10.1021/ac061881g.

(75) Duhamel, C.; Cardinael, P.; Peulon-Agasse, V.; Firor, R.; Pascaud, L.; Semard-Jousset, G.; Giusti, P.; Livadaris, V. Comparison of Cryogenic and Differential Flow (Forward and Reverse Fill/Flush) Modulators and Applications to the Analysis of Heavy Petroleum Cuts by High-Temperature Comprehensive Gas Chromatography. *J. Chromatogr. A* **2015**, *1387*, 95–103. https://doi.org/10.1016/j.chroma.2015.01.095.

(76) Bahaghighat, H. D.; Freye, C. E.; Synovec, R. E. Recent Advances in Modulator Technology for Comprehensive Two Dimensional Gas Chromatography. *TrAC Trends Anal. Chem.* **2019**, *113*, 379–391. https://doi.org/10.1016/j.trac.2018.04.016.

(77) Boswell, H.; Chow, H.-Y.; Gorecki, T. Modulators. In *Separation Science and Technology*; Elsevier, 2020; Vol. 12, pp 89–140. https://doi.org/10.1016/B978-0-12-813745-1.00004-0.

(78) Giddings, J. C. Concepts and Comparisons in Multidimensional Separation. *J. High Resolut. Chromatogr.* **1987**, *10* (5), 319–323. https://doi.org/10.1002/jhrc.1240100517.

(79) Liu, Z.; Phillips, J. B. Comprehensive Two-Dimensional Gas Chromatography Using an On-Column Thermal Modulator Interface. *J. Chromatogr. Sci.* **1991**, *29* (6), 227–231. https://doi.org/10.1093/chromsci/29.6.227.

(80) Giddings, J. C. Sample Dimensionality: A Predictor of Order-Disorder in Component Peak Distribution in Multidimensional Separation. *J. Chromatogr. A* **1995**, *703* (1–2), 3–15. https://doi.org/10.1016/0021-9673(95)00249-M.

(81) Stefanuto, P.-H.; Focant, J.-F. Columns and Column Configurations. In *Separation Science and Technology*; Elsevier, 2020; Vol. 12, pp 69–88. https://doi.org/10.1016/B978-0-12-813745-1.00003-9.

(82) Zhu, S. Effect of Column Combinations on Two-Dimensional Separation in Comprehensive Two-Dimensional Gas Chromatography: Estimation of Orthogonality and Exploring of Mechanism. *J. Chromatogr. A* **2009**, *1216* (15), 3312–3317. https://doi.org/10.1016/j.chroma.2009.01.107.

(83) Golay, M. J. E. Height Equivalent to a Theoretical Plate of Tubular Gas Chromatographic Columns Lined with a Porous Layer. *Nature* **1963**, *199* (4891), 370–371. https://doi.org/10.1038/199370a0.

(84) Peroni, D.; Janssen, H.-G. Comprehensive Two-Dimensional Gas Chromatography under High Outlet Pressure Conditions: A New Approach to Correct the Flow-Mismatch Issue in the Two Dimensions. *J. Chromatogr. A* **2014**, *1332*, 57–63. https://doi.org/10.1016/j.chroma.2014.01.051.

(85) Reichenbach, S. E.; Tao, Q.; Cordero, C.; Bicchi, C. A Data-Challenge Case Study of Analyte Detection and Identification with Comprehensive Two-Dimensional Gas Chromatography with Mass Spectrometry (GC×GC-MS). *Separations* **2019**, *6* (3), 38. https://doi.org/10.3390/separations6030038.

(86) Liu, Zaiyou.; Patterson, D. G.; Lee, M. L. Geometric Approach to Factor Analysis for the Estimation of Orthogonality and Practical Peak Capacity in Comprehensive Two-Dimensional Separations. *Anal. Chem.* **1995**, *67* (21), 3840–3845. https://doi.org/10.1021/ac00117a004.

(87) Camenzuli, M.; Schoenmakers, P. J. A New Measure of Orthogonality for Multi-Dimensional Chromatography. *Anal. Chim. Acta* **2014**, *838*, 93–101. https://doi.org/10.1016/j.aca.2014.05.048.

(88) Gilar, M.; Olivova, P.; Daly, A. E.; Gebler, J. C. Orthogonality of Separation in Two-Dimensional Liquid Chromatography. *Anal. Chem.* **2005**, *77* (19), 6426–6434. https://doi.org/10.1021/ac050923i.

(89) Davis, J. M.; Stoll, D. R.; Carr, P. W. Dependence of Effective Peak Capacity in Comprehensive Two-Dimensional Separations on the Distribution of Peak Capacity between the Two Dimensions. *Anal. Chem.* **2008**, *80* (21), 8122–8134. https://doi.org/10.1021/ac800933z.

(90) Semard, G.; Peulon-Agasse, V.; Bruchet, A.; Bouillon, J.-P.; Cardinaël, P. Convex Hull: A New Method to Determine the Separation Space Used and to Optimize Operating Conditions for Comprehensive Two-Dimensional Gas Chromatography. *J. Chromatogr. A* **2010**, *1217* (33), 5449–5454. https://doi.org/10.1016/j.chroma.2010.06.048.

(91) Nowik, W.; Héron, S.; Bonose, M.; Nowik, M.; Tchapla, A. Assessment of Two-Dimensional Separative Systems Using Nearest-Neighbor Distances Approach. Part 1: Orthogonality Aspects. *Anal. Chem.* **2013**, *85* (20), 9449–9458. https://doi.org/10.1021/ac4012705.

(92) Nowik, W.; Bonose, M.; Héron, S.; Nowik, M.; Tchapla, A. Assessment of Two-Dimensional Separative Systems Using the Nearest Neighbor Distances Approach. Part 2: Separation Quality Aspects. *Anal. Chem.* **2013**, *85* (20), 9459–9468. https://doi.org/10.1021/ac4012717.

(93) de Juan, A.; Tauler, R. Comparison of Three-Way Resolution Methods for Non-Trilinear Chemical Data Sets. *J. Chemom.* **2001**, *15* (10), 749–771. https://doi.org/10.1002/cem.662.

(94) Stefanuto, P. H.; Smolinska, A.; Focant, J. F. Advanced Chemometric and Data Handling Tools for GC×GC-TOF-MS: Application of Chemometrics and Related Advanced Data Handling in Chemical Separations. *TrAC - Trends Anal. Chem.* **2021**, *139*, 116251. https://doi.org/10.1016/j.trac.2021.116251.

(95) Wilson, J. D.; McInnes, C. A. J. The Elimination of Errors Due to Baseline Drift in the Measurement of Peak Areas in Gas Chromatography. *J. Chromatogr. A* **1965**, *19*, 486–494. https://doi.org/10.1016/S0021-9673(01)99489-0.

(96) Eilers, P. H. C. A Perfect Smoother. *Anal. Chem.* **2003**, *75* (14), 3631–3636. https://doi.org/10.1021/ac034173t.

(97) Reichenbach, S. E.; Ni, M.; Zhang, D.; Ledford, E. B. Image Background Removal in Comprehensive Two-Dimensional Gas Chromatography. *J. Chromatogr. A* **2003**, *985* (1–2), 47–56. https://doi.org/10.1016/S0021-9673(02)01498-X.

(98) Boelens, H. F. M.; Dijkstra, R. J.; Eilers, P. H. C.; Fitzpatrick, F.; Westerhuis, J. A. New Background Correction Method for Liquid Chromatography with Diode Array Detection, Infrared Spectroscopic Detection and Raman Spectroscopic Detection. *J. Chromatogr. A* **2004**, *1057* (1–2), 21–30. https://doi.org/10.1016/j.chroma.2004.09.035.

(99) Zhang, Y.; Wu, H.-L.; Xia, A.-L.; Hu, L.-H.; Zou, H.-F.; Yu, R.-Q. Trilinear Decomposition Method Applied to Removal of Three-Dimensional Background Drift in Comprehensive Two-Dimensional Separation Data. *J. Chromatogr. A* **2007**, *1167* (2), 178–183. https://doi.org/10.1016/j.chroma.2007.08.055.

(100) Zeng, Z.-D.; Chin, S.-T.; Hugel, H. M.; Marriott, P. J. Simultaneous Deconvolution and Re-Construction of Primary and Secondary Overlapping Peak Clusters in Comprehensive Two-Dimensional Gas Chromatography. *J. Chromatogr. A* **2011**, *1218* (16), 2301–2310. https://doi.org/10.1016/j.chroma.2011.02.028.

(101) Filgueira, M. R.; Castells, C. B.; Carr, P. W. A Simple, Robust Orthogonal Background Correction Method for Two-Dimensional Liquid Chromatography. *Anal. Chem.* **2012**, *84* (15), 6747–6752. https://doi.org/10.1021/ac301248h.

(102) Allen, R. C.; John, M. G.; Rutan, S. C.; Filgueira, M. R.; Carr, P. W. Effect of Background Correction on Peak Detection and Quantification in Online

Comprehensive Two-Dimensional Liquid Chromatography Using Diode Array Detection. *J. Chromatogr. A* **2012**, *1254*, 51–61. https://doi.org/10.1016/j.chroma.2012.07.034.

(103) Korytár, P.; van Stee, L. L. P.; Leonards, P. E. G.; de Boer, J.; Brinkman, U. A. Th. Attempt to Unravel the Composition of Toxaphene by Comprehensive Two-Dimensional Gas Chromatography with Selective Detection. *J. Chromatogr. A* **2003**, *994* (1–2), 179–189. https://doi.org/10.1016/S0021-9673(03)00438-2.

(104) Peters, S.; Vivó-Truyols, G.; Marriott, P. J.; Schoenmakers, P. J. Development of an Algorithm for Peak Detection in Comprehensive Two-Dimensional Chromatography. *J. Chromatogr. A* **2007**, *1156* (1–2), 14–24. https://doi.org/10.1016/j.chroma.2006.10.066.

(105) Vivó-Truyols, G. Bayesian Approach for Peak Detection in Two-Dimensional Chromatography. *Anal. Chem.* **2012**, *84* (6), 2622–2630. https://doi.org/10.1021/ac202124t.

(106) Min Chen; Reichenbach, S. E.; Jiazheng Shi. Automated Unmixing of Comprehensive Two-Dimensional Chemical Separations with Mass Spectrometry. In *2005 IEEE International Conference on Electro Information Technology*; IEEE: Lincoln, NE, USA, 2005; pp 1–6. https://doi.org/10.1109/EIT.2005.1627049.

(107) Reichenbach, S. E.; Ni, M.; Kottapalli, V.; Visvanathan, A. Information Technologies for Comprehensive Two-Dimensional Gas Chromatography. *Chemom. Intell. Lab. Syst.* **2004**, *71* (2), 107–120. https://doi.org/10.1016/j.chemolab.2003.12.009.

(108) Kim, S.; Ouyang, M.; Jeong, J.; Shen, C.; Zhang, X. A New Method of Peak Detection for Analysis of Comprehensive Two-Dimensional Gas Chromatography Mass Spectrometry Data. *Ann. Appl. Stat.* **2014**, *8* (2). https://doi.org/10.1214/14-AOAS731.

(109) Viant, M. R.; Kurland, I. J.; Jones, M. R.; Dunn, W. B. How Close Are We to Complete Annotation of Metabolomes? *Curr. Opin. Chem. Biol.* **2017**, *36*, 64–69. https://doi.org/10.1016/j.cbpa.2017.01.001.

(110) Veenaas, C.; Haglund, P. A Retention Index System for Comprehensive Two-Dimensional Gas Chromatography Using Polyethylene Glycols. *J. Chromatogr. A* **2018**, *1536*, 67–74. https://doi.org/10.1016/j.chroma.2017.08.062.

(111) Seeley, J. V.; Seeley, S. K. Model for Predicting Comprehensive Two-Dimensional Gas Chromatography Retention Times. *J. Chromatogr. A* **2007**, *1172* (1), 72–83. https://doi.org/10.1016/j.chroma.2007.09.058.

(112) Jiang, M.; Kulsing, C.; Nolvachai, Y.; Marriott, P. J. Two-Dimensional Retention Indices Improve Component Identification in Comprehensive Two-Dimensional Gas Chromatography of Saffron. *Anal. Chem.* **2015**, *87* (11), 5753–5761. https://doi.org/10.1021/acs.analchem.5b00953.

(113) Abramson, F. P. Automated Identification of Mass Spectra by the Reverse Search. *Anal. Chem.* **1975**, *47* (1), 45–49. https://doi.org/10.1021/ac60351a028.

(114) Stein, S. E.; Scott, D. R. Optimization and Testing of Mass Spectral Library Search Algorithms for Compound Identification. *J. Am. Soc. Mass Spectrom.* **1994**, *5* (9), 859–866. https://doi.org/10.1016/1044-0305(94)87009-8.

(115) Kim, S.; Koo, I.; Jeong, J.; Wu, S.; Shi, X.; Zhang, X. Compound Identification Using Partial and Semipartial Correlations for Gas Chromatography–Mass Spectrometry Data. *Anal. Chem.* **2012**, *84* (15), 6477–6487. https://doi.org/10.1021/ac301350n.

(116) Kim, S.; Zhang, X. Comparative Analysis of Mass Spectral Similarity Measures on Peak Alignment for Comprehensive Two-Dimensional

Gas Chromatography Mass Spectrometry. *Comput. Math. Methods Med.* **2013**, *2013*, 1–12. https://doi.org/10.1155/2013/509761.

(117) Samokhin, A.; Sotnezova, K.; Lashin, V.; Revelsky, I. Evaluation of Mass Spectral Library Search Algorithms Implemented in Commercial Software: Evaluation of Library Search Algorithms. *J. Mass Spectrom.* **2015**, *50* (6), 820–825. https://doi.org/10.1002/jms.3591.

(118) Zhang, J.; Xia, Y.; Zheng, C. H.; Wang, B.; Zhang, X.; Chen, P. Combine Multiple Mass Spectral Similarity Measures for Compound Identification. *Int. J. Data Min. Bioinforma.* **2016**, *15* (1), 84. https://doi.org/10.1504/IJDMB.2016.076018.

(119) Reichenbach, S. E.; Carr, P. W.; Stoll, D. R.; Tao, Q. Smart Templates for Peak Pattern Matching with Comprehensive Two-Dimensional Liquid Chromatography. *J. Chromatogr. A* **2009**, *1216* (16), 3458–3466. https://doi.org/10.1016/j.chroma.2008.09.058.

(120) Lee, S.-J.; Noble, A. C. Use of Partial Least Squares Regression and Multidimensional Scaling on Aroma Models of California Chardonnay Wines. *Am. J. Enol. Vitic.* **2006**, *57* (3), 363.

(121) Almaguer, C.; Schönberger, C.; Gastl, M.; Arendt, E. K.; Becker, T. *Humulus Lupulus* - a Story That Begs to Be Told. A Review: *Humulus Lupulus* - a Story That Begs to Be Told. *J. Inst. Brew.* **2014**, n/a-n/a. https://doi.org/10.1002/jib.160.

(122) Guth, H. Quantitation and Sensory Studies of Character Impact Odorants of Different White Wine Varieties. *J. Agric. Food Chem.* **1997**, *45* (8), 3027–3032. https://doi.org/10.1021/jf970280a.

(123) Gawel, R. Red Wine Astringency: A Review. *Aust. J. Grape Wine Res.* **1998**, *4* (2), 74–95. https://doi.org/10.1111/j.1755-0238.1998.tb00137.x.

(124) Bindon, K. A.; Smith, P. A.; Kennedy, J. A. Interaction between Grape-Derived Proanthocyanidins and Cell Wall Material. 1. Effect on

Proanthocyanidin Composition and Molecular Mass. *J. Agric. Food Chem.* **2010**, *58* (4), 2520–2528. https://doi.org/10.1021/jf9037453.

(125) Sacchi, K. L.; Bisson, L. F.; Adams, D. O. A Review of the Effect of Winemaking Techniques on Phenolic Extraction in Red Wines. *Am. J. Enol. Vitic.* **2005**, *56* (3), 197.

(126) Fernández de Simón, B.; Cadahía, E.; del Álamo, M.; Nevares, I. Effect of Size, Seasoning and Toasting in the Volatile Compounds in Toasted Oak Wood and in a Red Wine Treated with Them. *Anal. Chim. Acta* **2010**, *660* (1–2), 211–220. https://doi.org/10.1016/j.aca.2009.09.031.

(127) Pons, A.; Lavigne, V.; Landais, Y.; Darriet, P.; Dubourdieu, D. Identification of a Sotolon Pathway in Dry White Wines. *J. Agric. Food Chem.* **2010**, *58* (12), 7273–7279. https://doi.org/10.1021/jf100150q.

(128) Hazelwood, L. A.; Daran, J. M.; Van Maris, A. J. A.; Pronk, J. T.; Dickinson, J. R. The Ehrlich Pathway for Fusel Alcohol Production: A Century of Research on Saccharomyces Cerevisiae Metabolism (Applied and Environmental Microbiology (2008) 74, 8, (2259-2266)). *Appl. Environ. Microbiol.* **2008**, *74* (12), 3920. https://doi.org/10.1128/AEM.00934-08.

(129) Pires, E. J.; Teixeira, J. A.; Brányik, T.; Brandão, T.; Vicente, A. A. Continuous Beer Fermentation - Diacetyl as a Villain: Diacetyl as a Villain. *J. Inst. Brew.* **2015**, *121* (1), 55–61. https://doi.org/10.1002/jib.205.

(130) Berrier, K. L.; Prebihalo, S. E.; Synovec, R. E. Advanced Data Handling in Comprehensive Two-Dimensional Gas Chromatography. In *Separation Science and Technology*; Elsevier, 2020; Vol. 12, pp 229–268. https://doi.org/10.1016/B978-0-12-813745-1.00007-6.

(131) Skov, T.; Hoggard, J. C.; Bro, R.; Synovec, R. E. Handling within Run Retention Time Shifts in Two-Dimensional Chromatography Data Using Shift Correction and Modeling. *J. Chromatogr. A* **2009**, *1216* (18), 4020–4029. https://doi.org/10.1016/j.chroma.2009.02.049.

(132) Johnson, K. J.; Prazen, B. J.; Young, D. C.; Synovec, R. E. Quantification of Naphthalenes in Jet Fuel with GC×GC/Tri-PLS and Windowed Rank Minimization Retention Time Alignment. *J. Sep. Sci.* **2004**, *27* (5–6), 410–416. https://doi.org/10.1002/jssc.200301640.

(133) Pierce, K. M.; Wood, L. F.; Wright, B. W.; Synovec, R. E. A Comprehensive Two-Dimensional Retention Time Alignment Algorithm To Enhance Chemometric Analysis of Comprehensive Two-Dimensional Separation Data. *Anal. Chem.* **2005**, *77* (23), 7735–7743. https://doi.org/10.1021/ac0511142.

(134) Weusten, J. J. A. M.; Derks, E. P. P. A.; Mommers, J. H. M.; van der Wal, S. Alignment and Clustering Strategies for GC×GC–MS Features Using a Cylindrical Mapping. *Anal. Chim. Acta* **2012**, *726*, 9–21. https://doi.org/10.1016/j.aca.2012.03.009.

(135) Harvey, P. McA.; Shellie, R. A. Data Reduction in Comprehensive Two-Dimensional Gas Chromatography for Rapid and Repeatable Automated Data Analysis. *Anal. Chem.* **2012**, *84* (15), 6501–6507. https://doi.org/10.1021/ac300664h.

(136) Cordero, C.; Liberto, E.; Bicchi, C.; Rubiolo, P.; Reichenbach, S. E.; Tian, X.; Tao, Q. Targeted and Non-Targeted Approaches for Complex Natural Sample Profiling by GCxGC-QMS. *J. Chromatogr. Sci.* **2010**, *48* (4), 251–261. https://doi.org/10.1093/chromsci/48.4.251.

(137) Stadler, S.; Stefanuto, P.-H.; Brokl, M.; Forbes, S. L.; Focant, J.-F. Characterization of Volatile Organic Compounds from Human Analogue Decomposition Using Thermal Desorption Coupled to Comprehensive Two-Dimensional Gas Chromatography–Time-of-Flight Mass Spectrometry. *Anal. Chem.* **2013**, *85* (2), 998–1005. https://doi.org/10.1021/ac302614y.

(138) Bean, H. D.; Hill, J. E.; Dimandja, J.-M. D. Improving the Quality of Biomarker Candidates in Untargeted Metabolomics via Peak Table-Based Alignment of Comprehensive Two-Dimensional Gas Chromatography–Mass

Spectrometry Data. *J. Chromatogr. A* **2015**, *1394*, 111–117. https://doi.org/10.1016/j.chroma.2015.03.001.

(139) Pierce, K. M.; Hoggard, J. C.; Hope, J. L.; Rainey, P. M.; Hoofnagle, A. N.; Jack, R. M.; Wright, B. W.; Synovec, R. E. Fisher Ratio Method Applied to Third-Order Separation Data To Identify Significant Chemical Components of Metabolite Extracts. *Anal. Chem.* **2006**, *78* (14), 5068–5075. https://doi.org/10.1021/ac0602625.

(140) Marney, L. C.; Christopher Siegler, W.; Parsons, B. A.; Hoggard, J. C.; Wright, B. W.; Synovec, R. E. Tile-Based Fisher-Ratio Software for Improved Feature Selection Analysis of Comprehensive Two-Dimensional Gas Chromatography–Time-of-Flight Mass Spectrometry Data. *Talanta* **2013**, *115*, 887–895. https://doi.org/10.1016/j.talanta.2013.06.038.

(141) Hotelling, H. Analysis of a Complex of Statistical Variables into Principal Components. *J. Educ. Psychol.* **1933**, *24* (6), 417–441. https://doi.org/10.1037/h0071325.

(142) Sudol, P. E.; Gough, D. V.; Prebihalo, S. E.; Synovec, R. E. Impact of Data Bin Size on the Classification of Diesel Fuels Using Comprehensive Two-Dimensional Gas Chromatography with Principal Component Analysis. *Talanta* **2020**, *206*, 120239. https://doi.org/10.1016/j.talanta.2019.120239.

(143) Reichenbach, S. E.; Zini, C. A.; Nicolli, K. P.; Welke, J. E.; Cordero, C.; Tao, Q. Benchmarking Machine Learning Methods for Comprehensive Chemical Fingerprinting and Pattern Recognition. *J. Chromatogr. A* **2019**, *1595*, 158–167. https://doi.org/10.1016/j.chroma.2019.02.027.

(144) *Comprehensive Sampling and Sample Preparation: Analytical Techniques for Scientists*; Pawliszyn, J., Ed.; Elsevier: Amsterdam, 2012.

(145) David, F.; Ochiai, N.; Sandra, P. Two Decades of Stir Bar Sorptive Extraction: A Retrospective and Future Outlook. *TrAC Trends Anal. Chem.* **2019**, *112*, 102–111. https://doi.org/10.1016/j.trac.2018.12.006.

(146) Andrade-Eiroa, A.; Canle, M.; Leroy-Cancellieri, V.; Cerdà, V. Solid-Phase Extraction of Organic Compounds: A Critical Review (Part I). *TrAC Trends Anal. Chem.* **2016**, *80*, 641–654. https://doi.org/10.1016/j.trac.2015.08.015.

(147) Culleré, L.; López, R.; Ferreira, V. The Instrumental Analysis of Aroma-Active Compounds for Explaining the Flavor of Red Wines. In *Red Wine Technology*; Elsevier, 2019; pp 283–307. https://doi.org/10.1016/B978-0-12-814399-5.00020-7.

Supplementary

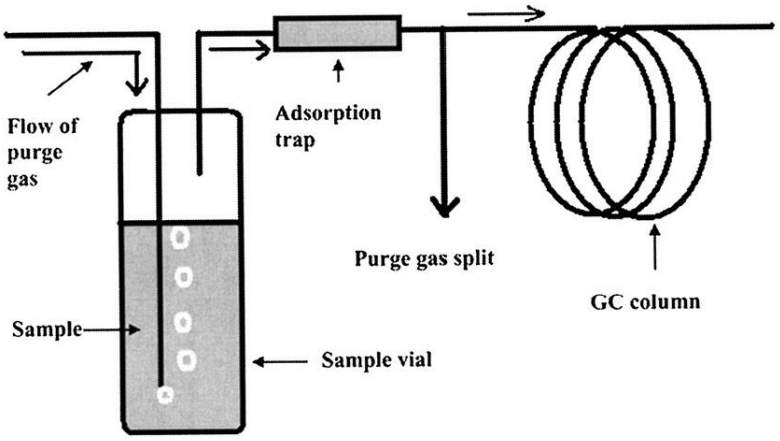


Figure S 1 the scheme of a dynamic headspace system [144].

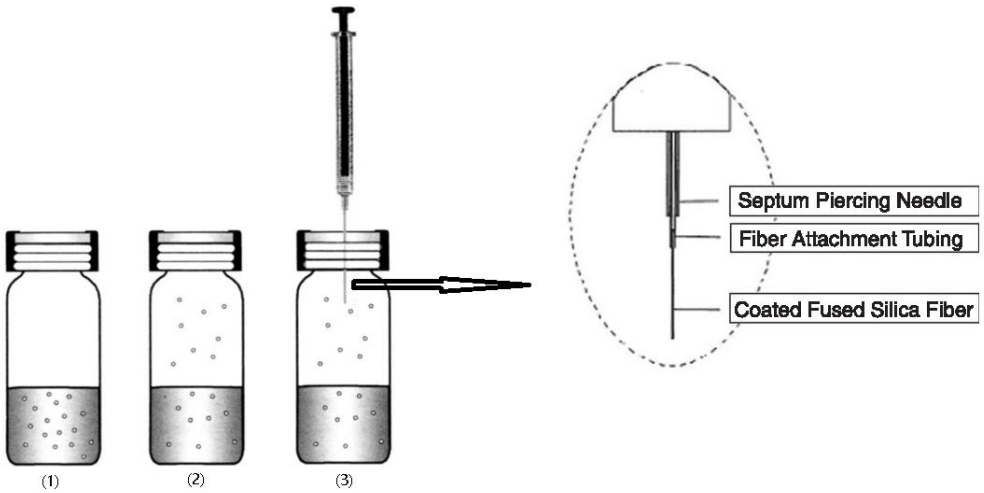


Figure S 2 the scheme of a HS-SPME. Adapted from book [144].

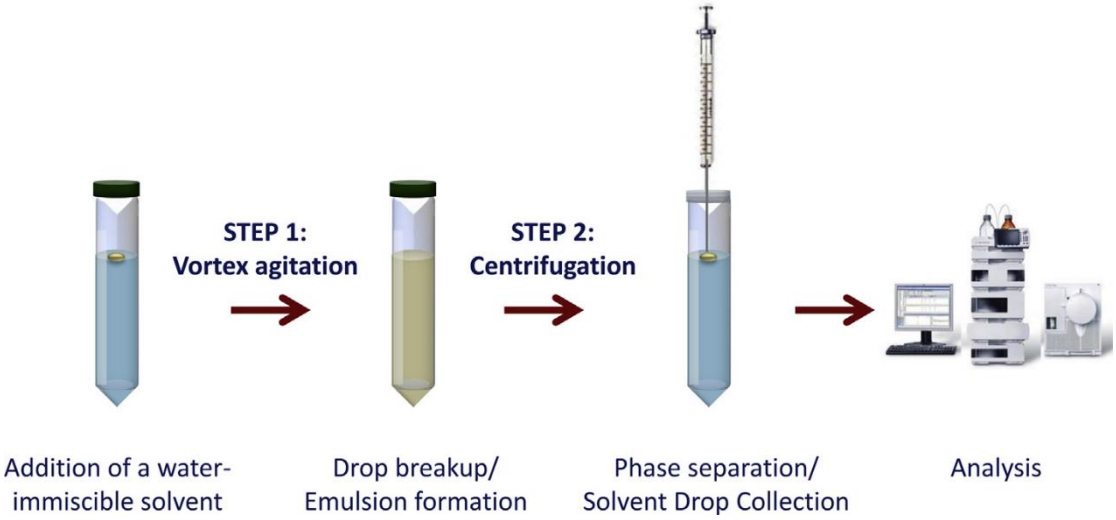


Figure S 3 schematic representation of the VALLME procedure [47].

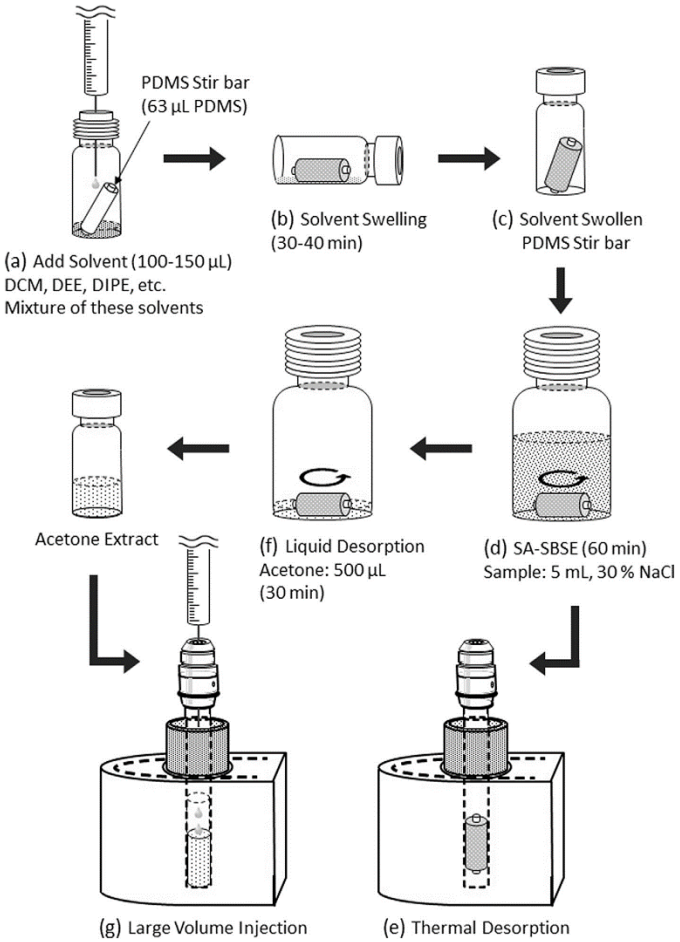


Figure S 4 Schematic of a SBSE sample preparation [145].

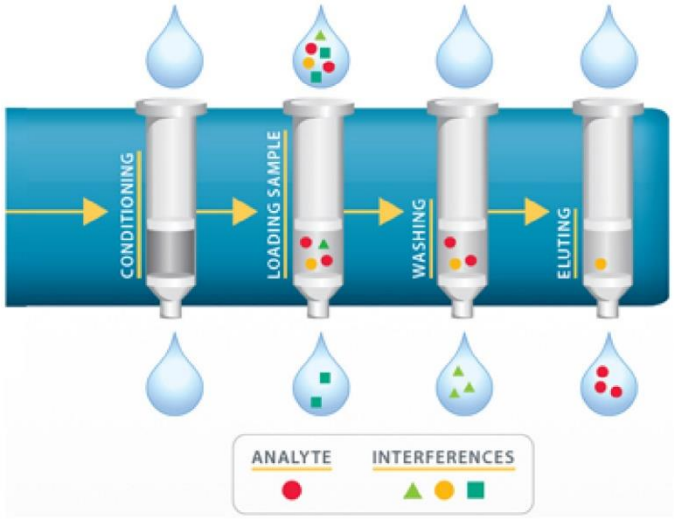


Figure S 5 a scheme of typical four step SPME [146].

Table S 1 physicochemical properties of common VOCs in the beverages, Odor threshold data are from

book chapter [147]. Other data are from PubChem.

Compound	**Odor threshold * µg/L**	**Aromatic descriptor**	**LogP**	**Henry constant (10^{-5} mol/atm, at 25°C)**	**Solubility (Water, mg/L)**	**Solubility (70% ethanol, v/v)**
Linalool	25	Flowery, muscat	2.97	2.15	1590	0.67
Geraniol	20	Flowery	3.56	1.15	100	0.25
P-Damascenone	0.05	Baked apple	3.2			
P-Ionone	0.09	Violets	4	8.1	169	0.25
Ethyl cinnamate	1.1	Flowery	2.99			miscible
Ethyl dihydrocinnamate	1.6	Plum, flowery	2.73		220	miscible
Z-3-Hexenol	400	Green	1.61	1.6	16000	soluble
Ethanol	24900	Alcohol	-0.31	5	1000000	
Isobutanol	40000	Fusel, alcohol	0.76	0.987	85000	miscible
Isoamyl alcohol	30000	Harsh, fusel	1.16	1.41	26700	soluble
P-Phenylethanol	14000	Rose	1.36	0.0256	22200	miscible
Methionol	1000	Plastic, green	0.5			soluble
Isoamyl acetate	30	Banana	2.25	58.7	2000	miscible
Phenylehtyl acetate	250	Honey, flowery	2.3		insoluble	0.33
Acetic acid	300000	Vinegar	-0.17	0.0143	1000000	miscible
Butyric acid	173	Cheese, Rancid	0.79	0.0535	60000	miscible
Hexanoic acid	420	Fatty, cheese	1.92	0.0758	10300	miscible
Octanoic acid	500	Rancid, harsh	3.05	0.0829	789	miscible
Decanoic acid	1000	Fatty	4.09	0.134	61.8	very soluble
Ethyl acetate	12300	Solvent, fruity	0.73	13.4	80000	miscible
Ethyl butyrate	125	Fruity	1.85	43.6	4900	miscible
Ethyl hexanoate	62	Green, apple	2.4		629	0.33
Ethyl octanoate	580	Sweet, fruit soap	3.5		70.1	0.2
Ethyl decanoate	200	Sweet, fruit	4.6		15.9	miscible
Isobutyric acid	2300	Acid, cheese	0.94	0.0885	167000	miscible
2-Methylbutyric acid	33	Acid, cheese	1.18		45000	miscible
3-Methylbutyric acid	33	Acid, cheese	1.16	0.0833	40700	soluble
Ethyl isobutyrate	15	Sweet fruit	1.5			miscible
Ethyl 2-methylbutyrate	18	Sweet fruit	1.9			soluble
Ethyl 3-methylbutyrate	3	Fruity	1.7		2000	miscible
Diacetyl	100	Cream, butter				
2,3-Pentanedione	900	Cream, butter	-0.85	0.0262	66700	miscible
Ethyl lactate	154000		-0.18	0.058	1000000	miscible
Z-Whiskylactone	67	Coconut	2.5			
Guaiacol	9.5	Smoky, spicy	1.3	0.12	18700	miscible
Eugenol	6	Clove	2.27	0.192	2460	miscible
E-Isoeugenol	6	Flowery	3.04	0.36	810	miscible

Vanillin	995	Vanilla	1.21	0.00022	11020	miscible
Acetovanillone	1000	Vanilla, caramel				
4-Ethylphenol	35	Animal, leather	2.58	0.0773	4900	miscible
4-Ethylguaiacol	33	Phenolic	1.7			miscible
4-Propylguaiacol	10	Clove, phenolic	2.2			miscible
y-Octalactone	7	Coconut	1.6			
y-Nonalactone	25	Peach	2.08	17	1201	soluble
y-Decalactone	0.7	Peach	2.72			soluble
4-Vinylphenol	180	Medicinal	2.4		soluble	soluble
4-Vinylguaiacol	40	Phenolic, pleasant	2.4		insoluble	miscible
Furaneol	5	Cotton candy	0.95	1.5	315000	soluble
Homofuraneol	125	Cotton candy				
Sotolon	5	Curry, spicy	0.4			
Ethyl 2-hydroxy-4-methylpentanoate	300	Blackberry	1.6			

2 On sample preparation methods for fermented beverage VOCs profiling by GC×GC‐TOFMS

Penghan Zhang, Silvia Carlin, Cesare Lotti, Fulvio Mattivi, Urska Vrhovsek

Published in Metabolomics, volume 16, issue 102, September 2020, DOI: 10.1007/s11306-020-01718-7

Abstract

Aromas and tastes have crucial influences on the quality of fermented beverages. The determination of aromatic compounds requires global non-targeted profiling of the volatile organic compounds (VOCs) in the beverages. However, experimental VOC profiling result depends on the chosen VOC collection method. This study aims to observe the impact of using different sample preparation techniques [dynamic headspace (DHS), vortex-assisted liquid‐liquid microextraction (VALLME), multiple stir bar sorptive extraction (mSBSE), solid phase extraction (SPE), and solid phase micro-extraction (SPME)] to figure out the most suitable sample preparation protocol for profiling the VOCs from fermented beverages. Five common sample preparation methods were studied with beer, cider, red wine, and white wine samples. After the sample preparation, collected VOCs were analyzed by comprehensive two-dimensional gas chromatography coupled with time of flight mass spectrometry (GC×GC-TOFMS). GC×GC oven parameters can be optimized with the Box‐Behnken surface response model and response measure on peak dispersion. Due to the unavoidable column and detector saturation during metabolomic analysis, errors may happen during mass spectrum construction. Profiling results obtained with different sample preparation methods show considerable variance. Common findings occupy a small fraction of total annotated VOCs. For known fermentative aromas, best coverage can be reached by using SPME together with SPE for beer, and VALLME for wine and cider. GC×GC-TOFMS is a promising tool for non-targeted profiling on VOCs from fermented

beverages. However, a proper data processing protocol is lacking for metabolomic analysis. Each sample preparation method has a specific profiling spectrum on VOC profiling. The coverage of the VOC metabolome can be improved by combining complementary methods.

2.1 Introduction

Aromas and tastes are the major factors influencing the consumer's perception of the quality of the wine, beer, and other fermented beverages [1]. Determining the aromatic compounds in those beverages and understanding their effects on the human organoleptic effect is important [2]. The determination of aromatic compounds usually starts with the characterization and quantification of the volatiles [3]. In the past, due to the insufficient analytical capability, understanding the flavor perception was supported by models with simplified chemical inputs. Important volatile compounds were preselected based on their odor activity values (OAVs) [4]. Widely used approaches are aroma extract dilution analysis (AEDA) and odor-specific magnitude estimation (OSME) [5,6]. According to the OAVs, about 75 aromatic compounds were found and listed in a review study [7].

With the development of modern chromatography, potential aromatic compounds can be selected based on the metabolomic scenario [8]. During this approach, a wider spectrum of chemical compounds are studied by chemometrics tools, aimed to explore the chemical composition of food to its sensory property. With the potential to reveal previously ignored aromatic compounds and relationships. Eventually, a better perception model can be constructed. To apply a chemometric approach, complete profiling of volatile organic compounds (VOCs) is desirable. One of the challenges on the mapping work is that the number of volatile compounds in a fermented beverage is over 1300 [9]. An advanced analytical platform is demanded to detect and ideally identify each of them. Two-dimensional gas chromatography coupled with a mass spectrometer (GC×GC-MS) might be a solution [10]. However, the results of VOC profiling are not only separation and detection dependent, but also sample preparation dependent [11]. Unfortunately, because of the diversity of VOCs, an ideal method that can concentrate and recover all VOC compounds does not exist. Commonly used sample preparation techniques for VOC

analysis are dynamic headspace (DHS), Vortex-assisted liquid-liquid microextraction (VALLME), multiple stir bar sorptive extraction (mSBSE), solid phase extraction (SPE), and solid phase microextraction (SPME).

DHS is widely applied for the advantages as environment-friendly, easy to implement and versatile procedure. It uses a continuous flow of inert gas to extract the volatile organic compounds from a liquid matrix. Extracted VOCs are further concentrated into an adsorbent or cryogenic trap. In the ideal case, the signal intensity of an analyte is negatively linearly correlated to its remaining concentration in the sample, which depends on its original concentration and decreases exponentially with sampling time and evaporative constant [12]. Traps made of Tenax TA sorbent are the most common, owing to their wide absorption range, high-temperature stability, low water affinity, and long shelf life [13].

VALLME draws attention by its simplicity, speed, miniaturization and low-cost [14]. During the sample preparation, the sample and the organic phase are mixed together. Mass transfer of an analyte from sample to organic solvent phase depends on the product of the overall mass transfer coefficient on the interface and the interfacial area. By vortexing, a turbulence flow is generated to breakdown the solvent phase into small droplets, forming an emulsion. Consequently, the interfacial area is enlarged [15]. The emulsions formed during VALLME are thermodynamically unstable. At the final stage of sample preparation, according to Stoke's Law, a higher gravitational acceleration is introduced to speed up the decomposition of emulsion to form separate phases [16].

The mSBSE and SPME are appreciated for their features of solvent-less extraction, predictable extraction efficiencies, multiple uses for determinations

in gaseous and aqueous matrices, and excellent repeatability and reproducibility. They also share the same basic principle. VOCs are extracted from a sample based on their affinity for and solubility in a viscous liquid phase. The extraction is a kinetic process with theoretical recovery depending on the phase ratio and distribution coefficient. Certain sampling time is required to reach the final equilibrium state. One of the most important factors for mSBSE and SPME analysis is sorbent. Thanks to the low glass transition temperature (<40 ºC), very low bleeding profile at high temperatures (>350 ºC) after appropriate conditioning, and excellent diffusivity and permeability, PDMS is by far the best material for non-polar analytes [17]. To cover more polar analytes in one analysis, supplementary sorbents are used in combination with PDMS. For SBSE, the only commercially available product is an ethylene glycol modified silicone material (EG silicon) [18]. For SPME, the most common three-phases sorbent is DVB/Car/PDMS [19].

Solid-phase extraction (SPE) is the most widely used and flexible method for the extraction, changing of solvents, clean-up, concentration and fractionation of organic compounds from several samples. SPE can absorb a wide range of organic analytes (from non-polar to very polar analytes) and has high extraction efficiency. Besides, analytes absorbed onto the SPE cartridge/column/disc are stable, can be stored for a long time [20]. The retention and elution mechanism of SPE is similar to the high-performance liquid chromatography. Solid-phase retains the molecule in three measures: partition for small nonpolar analytes [21], adsorption for larger nonpolar molecules [22], and electronic interactions, such as hydrogen bonding, ionic interactions, or π–π bonds for polar and polarizable analytes [23]. Many sorbents are available for SPE analysis. Among all, ENV+, hyper-crosslinked hydroxylated polystyrene divinylbenzene copolymer, shows high specific surface area and an extraordinary adsorption capability [24].

Different sample preparation techniques have been studied based on the analytical results obtained from 1DGC-MS [25,26]. In some studies, GC×GC-MS have been applied to fermented alcoholic samples analysis with one or more sample preparation techniques [11,27]. However, so far, to our knowledge, there has been no study on selecting the most suitable sample preparation techniques, coupled with untargeted GC×GC-MS, for the measurement of VOCs from fermented beverages. We studied the above mentioned sample preparation techniques followed by GC×GC-MS analysis to compare their capability of recovering VOCs from different fermented beverages (beer, white and red wine and cider).

2.2 Material and methods

2.2.1 Samples, reagents and supplies

Four types of fermented beverages, white wine, red wine, cider, and beer were studied. For each type of beverage, a pooled sample was prepared by mixing commercial products from local markets. Pooled white wine was obtained by mixing Riesling, Műller, Manzoni Bianco, Sauvignon blanc, Veltliner, and Gewürztraminer to cover as many volatile compounds as possible. Pooled red wine was made from single variety *Blaufrankisch* of different quality levels. Pooled cider sample covers various varieties from Trentino Alto Adige (Golden Delicious, Stark Delicious, Granny Smith, Royal Gala, Winesap, Morgenduft, Fuji, and Braeburn). Pooled beer was made of pale lager beers produced by Heineken. Beer samples were degassed by sonication for 10 seconds before the mixing.

To optimize the GC×GC separation, a stock solution of 10 mg/L in ethanol was prepared. It consists of 131 chemical compounds which are commonly present in fermented beverages (Table S 2). Alkane standards (C9 to C30), internal

standards (2-octanol and 1-heptanol), sodium chloride, ammonium sulfate, and dichloromethane were purchased from Sigma Aldrich.

2.2.2 Sample preparation

Dynamic headspace:

Tenax TA sorbent tube was conditioned at 315 ºC for 1 hour in TDS Tube Conditioner (Gerstel). Two milliliters of the sample were dosed into a 20 mL headspace vial which contains 2 g of NaCl. Before the sampling, 25 µL of 2-octanol solution (2 mg/L in ethanol) was added as internal standard. DHS sampling was carried out with Gerstel MPS2 auto-sampler (Mülheimander Ruhr, Germany), including the sampling, thermal desorption (ThermoDesorption Unit, Gerstel) and cryo focusing (Cooling Injection System, Gerstel). Detailed experimental parameters are listed in Table S 3.

Vortex-assisted liquid-liquid microextraction:

The method was in accordance with a previous study [28] with minor modification. 2.5 mL of sample was placed together with 3.5 mL deionized water and 2.5g $(NH_4)_2SO_4$ in a 15 mL glass vortex tube. After completely dissolve, 20 µL of internal standard solution (100 mg/L 1-heptanol in EtOH) and 0.2 mL of dichloromethane were added. The mixture was vortexed at 3200 rpm for 1 minute and then centrifuged at 5 °C, 1174 RCF for 10 minutes. Once the phases were separated, the organic phase was transferred into a 0.3 mL vial by a 250 µL syringe (Hamilton). Collected volatile extracts were stored at -20 °C until the analysis. From each extract, 1 µL was taken for the GC×GC analysis. Liquid injections were performed with Gerstel MPS2 auto-sampler monitored by ChromaToF (Leco, St Joseph, MI, USA).

Multiple stir bar sorptive extraction:

Stir bars (Twister) coated with 24 µL PDMS and 32µL EG Silicone were soaked in acetonitrile for 12 hours and then thermal conditioned by Gerstel TDS Tube Conditioner. PDMS stir was conditioned at 300 ºC in a nitrogen flow for 60 minutes. For EG Silicone stir bar, the conditioning temperature and time were 220 ºC and 30 minutes. After conditioning, they were placed in a 10 mL headspace vial [29]. 10 mL of sample was dosed and sampled at room temperature for 1 hour. The stirring speed was 800 rpm. After the sampling, the stir bar was washed by deionized water, dried with lint-free tissue and place in thermal desorption liner. Thermal desorption and cryo-focusing parameters are shown in Table S 4.

Solid phase extraction:

Sample preparation and extraction protocols were modified from a previous study [30]. ENV+ cartridges (1 g Biotage, Sweden), used during solid-phase extraction, were pre-conditioned with 15 mL methanol followed by 20 mL of deionized water. For each sample, 50 mL of the sample (25 mL for degassed beer sample) was mixed with 100 µL n-heptanol solution (250 mg/L in ethanol) and loaded in the SPE cartridge. After the extraction, the solid phase was washed with 15 mL of deionized water. Analytes were recovered by 30 mL of dichloromethane and then concentrated to 250 µL.

Solid-phase microextraction:

SPME fiber of 2 cm (50/30 DVB/CAR/PDMS), from Supelco, was conditioned according to the manual before use. For each pooled sample, 1 mL of the matrix was mixed with 1 g NaCl in a 20 mL headspace vial. Before the analysis, 25 µL of internal standard solution (2-octanol in ethanol at concentration of 2 mg/L) was added. Instrumental setup can be found in Table S 5. Other details can be found in the literature [27].

2.2.3 GC×GC separation and detection

GC×GC system was the Agilent 7890 A (Agilent Technologies, Santa Clara, CA). Injections were performed with splitless mode. Equipped columns were VF-Wax column (100% polyethylene glycol; 30m x 0.25mm x 0.25 µm, Agilent J&WScientific Inc., Folsom, CA) as the 1st dimension and Rxi-17Sil MS 1.50 m x 0.15mm x 0.15 µm, Restek Bellefonte, USA) as the 2nd dimension. A non-moving quad jet dual-stage thermal modulator was used to coupling the two columns. MS signal was obtained with Pegasus IV time-of-flight mass spectrometer (Leco Corporation, St. Joseph, MI).

GC×GC separation parameters were optimized according to the Box-Behnken experiment design (JMP, SAS Institute, Cary, NC, USA). Used column setup was a polar column followed by a medium polar column, recommended by a previous study [31]. Considered independent variables were column flow, temperature program, 2nd oven temperature offset, modulation temperature offset, modulation time, hot pulse time (% of the entire modulation time). The dependent variable was the median value of Nearest Neighbor Distance (NND) [32], which is calculated only based on the annotated peaks. Modulation time was adjusted to avoid the wrap-around phenomena. After the optimization, same standard solution was injected and separated under the model suggested condition to confirm the real separation measure.

The optimized column flow was 0.8 mL/min. The oven temperature was a program from 40 °C (2 min holding time) to 250 °C at a rate of 3 °C/min. After reaching the final temperature of 250 °C, 15 min holding time was applied. The temperature offsets of second oven and modulator were fixed at +5 °C and +8 °C respectively. Within the 6-second-modulation time, 2.1 seconds are used for hot pulse. The transfer line was kept at 250 °C. The TOFMS was operated in electron ionization mode at 70 eV. The ion source temperature was 230 °C.

The acquisition frequency was 200 Hz within a mass scan range from 35 m/z to 450 m/z. The detector voltage was -1341 V. In the case of the liquid injection, an acquisition delay of 7 min was applied. This setup was applied to the measurements of all types of beverages.

2.2.4 Data processing and Peak annotation

GC×GC-MS data acquisition and processing were achieved with LECO ChromaTOF (Version 4.22). The processing consists of peak picking, peak annotation, and statistic confirmation. During the peak picking, signals which were just above the noise were taken into account (baseline offset = 1). Minimal expected peak width (on 2^{nd} dimension) for deconvolution was 0.8 s. A peak was defined when at least 5 ions whose signal to noise ratio is above 100 [11], can be grouped. A picked peak was annotated by matching its mass spectrum (MS) to the reference spectrum in the database. In this study, used MS databases were NIST/EPA/NIH 11, Wiley 8 and the FFNSC 2. The MS similarity threshold for the peak annotation was 700. To determine this value, a standard MS library was created by recording the real measured mass spectrums of chemical standards (Table S 2) under a well-separated condition. The same chemicals were mixed and analyzed under different GC×GC separation conditions (Table S 6). A reversed match mode (match the library mass spectrums to the measured mass spectrum) was used to identify these chemical peaks. The similarity of the mass spectrum of each peak under each separation condition was collected and to be used to evaluate the similarity threshold value. For pooled samples, to minimize the peak detection and annotation error, each sample was analyzed with three technical replications. Inter-measurements peak alignment was performed based on the retention times (both $^{1}t_R$ and $^{2}t_R$) and mass spectrum. A minimal MS similarity of 600 was required. An analyte was further examined, only if it can be detected in all the technical replications. An inter-class comparison was performed between sample class and blank class. Fisher ratio thresholding was used to eliminate

artifact compounds (sorbent bleeding, column bleeding and other possible interferences). The applied significance level was 0.05. Peak identification was then completed by checking the linear temperature programmed retention index (LTPRI), which is available in the NIST RI database.

2.3 Result and discussion

2.3.1 Separation optimization

Many methods have been published to measure the separation degree for two-dimensional chromatography. A brief review of each type of method was published in a previous study [33]. Among all the methods, nearest neighbor distance was chosen because it measures the absolute dispersion among all the peaks. For non-targeted profiling, sample composition is usually complex and often requires an extended analytical time. If the separation space is limited (in the time domain), approaches based on the orthogonality measure are more suitable. All the major oven parameters for GC×GC separation were taken into account during the optimization (Table S 6).

The experimental data were analysis by a Box-Behnken surface response model, which allows efficient estimation of the first- and second-order coefficients for input variables [34]. The model is perfectly fitted with adjust R^2 is nearly 1 (Table S 7). Apart from the quadratic term made of column flow and temperature program rate, all the model components have significant influences on the separation result. The model supposes that the oven temperature program rate and modulation time are the most important parameters (Table S 8 & Table S 9). This result is expected. Oven temperature directly links to the retention time, which linearly impacts the separation resolution. Modulator act as a collector for elute compounds. Elongate the modulation time reduces the separation resolution of the 1st dimension. A lower

value for oven temperature and modulation time is always desired unless the wrap-around effect happens. When a compromise has to be made, according to the absolute value of estimates of each parameter, higher priority has to be given to the temperature rate. Other parameters were optimized within the range in which a valid chromatogram can be obtained (Table S 6). Chosen values are not always at apices. Optimized separation parameters were: Column flow, 0.8 mL/min; Modulation time, 6s; 2nd oven temperature offset, 5 °C; Temperature program, 3 °C; Modulation temperature offset, 8 °C; Hot pulse time, 35% of the entire modulation time. Model predicted median NND under the optimized oven condition was 14.6s (**Fig S 1**). The experimental confirmation shows the true median NND was 12s, which is 20% lower than the model predicted value.

2.3.2 MS similarity

MS matching is the most important step for peak annotation. A suitable similarity threshold value should concern both the annotation accuracy and recovery of the not-well-resolved peaks. A previous study [35] recommended using a moderate MS matching score, between 500 and 700, to obtain the highest proportion of reproducible peak. In our study, a standard solution, which contains 131 common volatile compounds from wine, was analyzed under the diverse separation conditions. Sum-up the analytical results under all the separation conditions, 2715 annotation were made by reserved matching. Matching scores of each annotation were collected and the distribution is shown in **Fig S 2**. The majority of the peaks, 2149 out of 2715, matched to the library mass spectrum with a relatively high score (>800). Below 800, peak identification was improved by 2‰ for every 20 unit reduction on matching score. Eventually, a sharp drop in the improvement was observed starting from 720. It is reasonable to set the similarity threshold value at 700.

The deviation of the MS is the result of peak picking error which is caused by over deconvolution. As demonstrated in **Fig S 3**. The MS signals of 1-phenylethyl acetate were grouped into two peaks. The combination of the constructed two mass spectra is more similar to the reference mass spectrum from the NIST library. Over deconvolution can be partially eliminated - in theory - by preventing the peak from the non-Gaussian shape and increasing the expected peak width value during the data processing. However, in the case of the non-targeted metabolomics study, the content of the compounds can vary up to 10^6 magnitudes. It's not possible to find a sample preparation approach that allows the presence of trace components in the chromatogram, and at the meantime, keeps the peak shapes Gaussian-like for major compounds; or define an ideal peak width for the entire chromatogram. Tolerance from the MS matching score is necessary.

With a better separation (2.3.1), peaks are better picked. Consequently, the constructed mass spectrum should be more similar to the reference spectrum in the library. A violin plot (**Fig 1**) was created to visualize the relationship between the similarity distribution and the median of nearest neighbor distance, which is a global measure of peak separation [32]. From 4 to 12 NND units, peaks are more likely to have the MS matching score at 900 and less likely to be at 700. However, the improvement is not remarkable. A reason could be that the improvement of the global separation measure is not equivalent to the better separation at the non-well separated region. Compared to a real natural sample, the standard solution used for the test only contains a limited number of compounds. A threshold value would be lower than 700 if peaks are heavily overlapping with each other. There is still some room for the improvement on the peak picking. A better data processing protocol would be desired for the complex sample measurement in the metabolomics study.

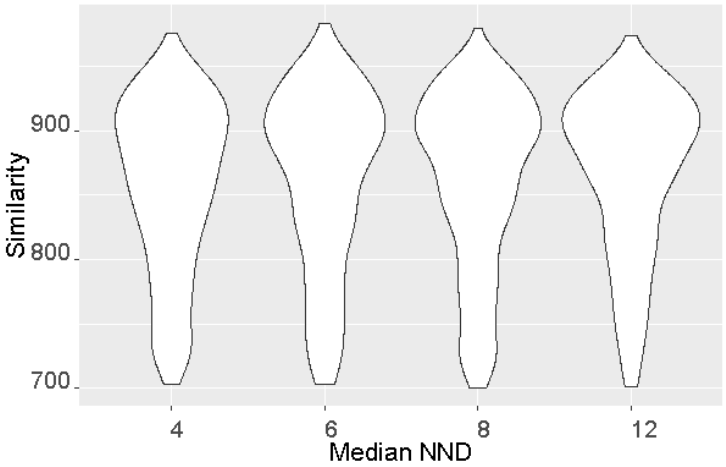


Fig 1 Violin plot of similarity distribution and Median NND

2.3.3 Comparing the analysis of five sampling techniques on four types fermented beverages

On pooled beer, red wine, cider, and white wine sample, DHS, VALLME, mSBSE, SPE and SPME sampling techniques were tested. VALLME does not apply to the beer sample, since after sample preparation, the organic phase is mixed with a gel phase. It's impossible for the further GC injection.

Only considering the annotated peak number, the best technique for the volatile compounds profiling was: SPE for beer (166) and red wine (433), DHS for cider (330) and VALLME for white wine (256). Details are listed in Table 1. If the automation level is taken into account, SPME would be the best choice for the beer sample, because SPE can only be operated manually. The experimental properties of each sample preparation technique are presented in Table 2.

Table 1: Numbers of annotated peaks obtained from different sample preparation techniques

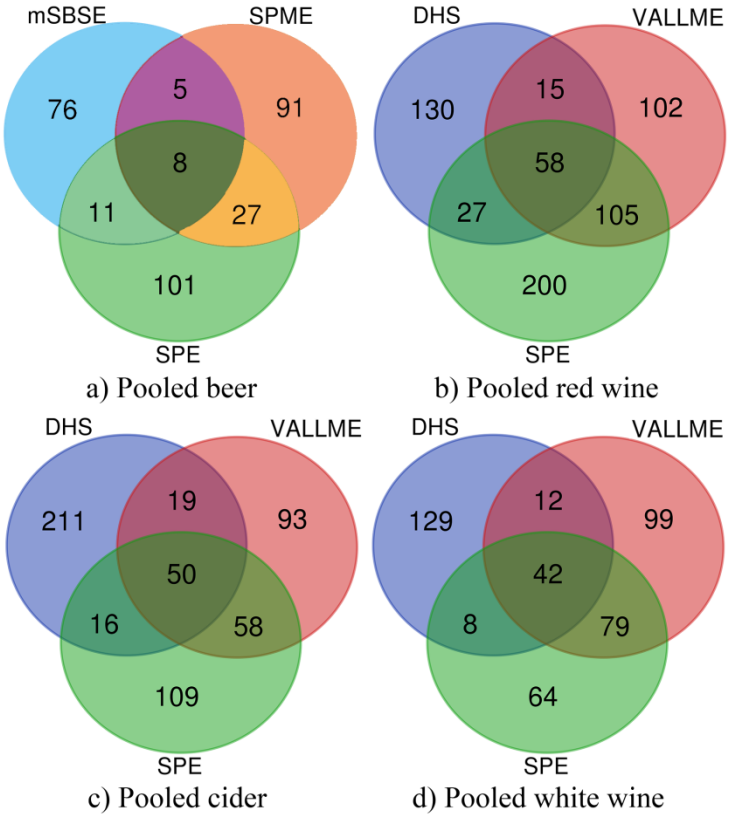
	DHS	VALLME	mSBSE	SPE	SPME
Pooled beer	53	N.A.	130	166	142
Pooled red wine	257	303	239	433	150
Pooled cider	330	257	272	280	223
Pooled white wine	215	256	168	218	145

Table 2: Comparison of 5 sampling techniques

	DHS	VALLME	mSBSE	SPE	SPME
Sample quantity	2 ml	2.5 ml	10ml or 1ml(10x*)	50 ml (10x*)	1 ml
Preparation time	≈30 min	≈10 min	≈1 h	≈2 h	≈20 min
Labor cost	auto	manual	manual	manual	auto
Drawbacks	need opt. for different sample	not applicable for beer sample	incomplete thermal desorption	solvent and time cost	

* 10 times dilution (see the sections2.3.4, 2.3.5)

For each type of sample, the best three techniques are further compared by the Venn plot (**Fig 2**). Surprisingly, each technique has its own unique profiling spectrum. Only a small fraction of the annotated peaks can be found by all three techniques. On the other hand, this means that a combination of different sample preparation techniques has the potential to improve global coverage dramatically. To be noticed, the data presented here are not all the features contained in the chromatogram but only the annotated ones.



a) Pooled beer b) Pooled red wine

c) Pooled cider d) Pooled white wine

Fig 2 Venn plot for the selected best three sample preparation techniques on four types of fermented beverages

To have a more detailed look, for each technique, identified compounds were classified according to their chemical categories. Bar plots were generated for beer (**Fig 3**), red wine (**Fig S 4**), cider (**Fig S 5**), and white wine (**Fig S 6**). Taking the measurements on beer as an example: SPE has better sampling performance for the volatile compounds in most chemical classes, except ethers and hydrocarbons (including terpenes). SPME is good at collecting and concentrating hydrocarbons (including terpenes), esters, and alcohols. DHS, probably, is not an ideal technique for VOCs profiling for beer. For the red wine sample, with SPE, fewer peaks were found in aldehyde class and in ether class compared to DHS and mSBSE respectively. It's hard to reach an agreement on which technique or even the combination of the techniques is the most suitable for the VOCs profiling on red wine, cider, and white wine. The choice will be easier to make with a specific study focus.

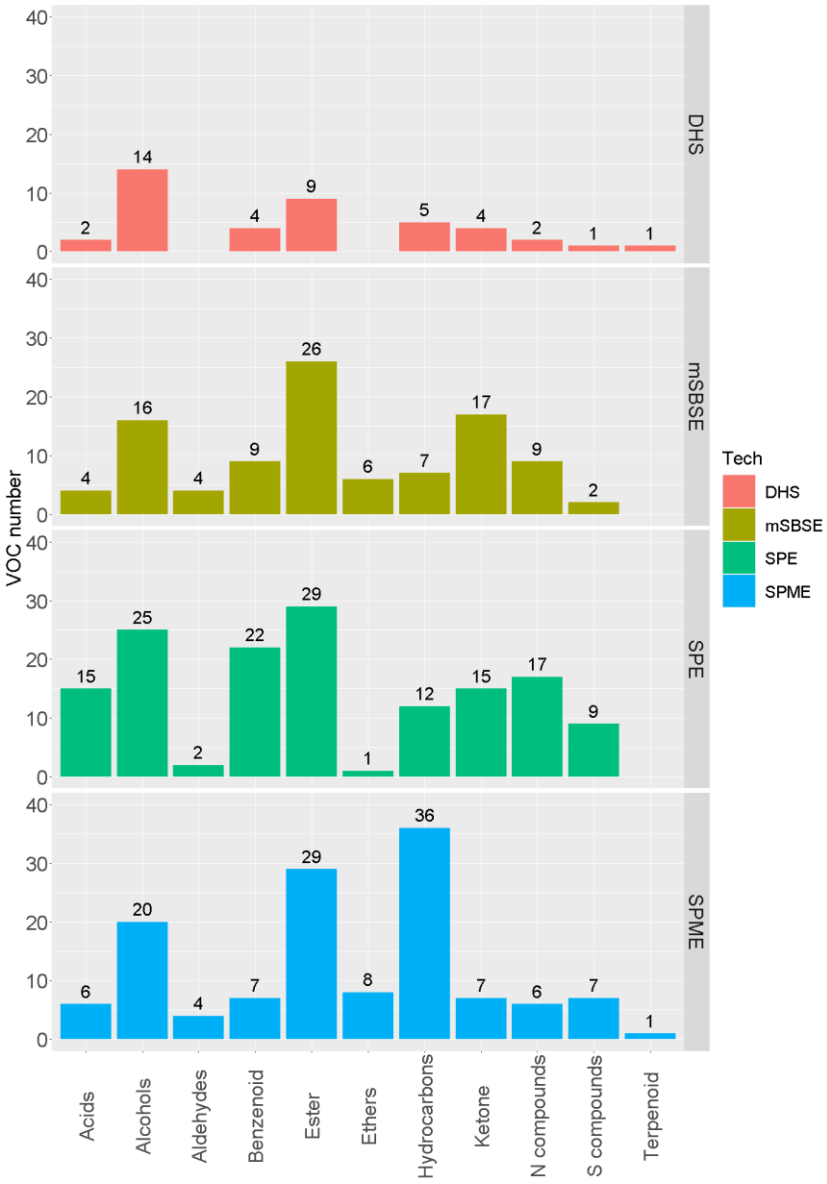


Fig 3 Comparison of the aligned peak number by chemical classes for applying different sample preparation techniques on pooled beer

2.3.4 SPE sample dose optimization

For the non-target profiling analysis, analytical techniques and protocols which allow higher sample dose, are preferred to concentrate the trace compounds above the detection limit of the instrument. Taking into consideration that several key odorants usually present at trace levels (ng to µg/L) are recognized by the human olfactory system. However, in the case of the GC×GC separation followed by automatic peak picking and annotation, drawbacks of the higher sample dose must be noticed. The influence of peak intensity and resolution on alignment has been studied before [35]. A negative result will be obtained when major chemicals saturate the column or the detector. Besides, overdose interferes with the measurement of trace compounds by decreasing the overall peak resolution. SPE was performed with 50 mL of pooled red wine sample. To study the influence of sample dose, a dilution series was made from 10 to 1000 times. With 100 times dilution, most inter-measurement aligned features (938) were detected. With 10 times dilution, only 30 peaks can be aligned. A possible strategy to improve sampling would be: use relatively less quantity of sample, make a dilution to reach the quantity demanded by the protocol and use excessive sorbent or a combination of different sorbents to increase the recovery.

2.3.5 mSBSE limitations and potential

Compare to SPME, the performance of mSBSE is not very satisfying. For most chemical classes, the peak number found with mSBSE is only slightly higher than the SPME. If we look at hydrocarbons measurement on pooled beer (**Fig 3**), more compounds were found by SPME (36) than mSBSE (7). However, in both techniques, PDMS was used to collect the non-polar compounds. The amount of sorbent coated on the stir bar is at least 50 times higher than the amount on SPME fiber. Besides, the amount of sample used during the SPE sampling was 2.5 times (adjusted by dilution before the GC injection) of SPME.

A possible explanation for fewer peaks found with mSBSE sampling is there was incomplete thermal desorption for PDMS stir. According to the manual, the recommended thermal desorption (TD) temperatures for PDMS stir is 300 °C. However, during desorption, the temperature can rise until 220 °C, limited by the thermal tolerance of EG silicon stir. The retaining effect was confirmed by a second injection after the TD-GC×GC-MS analysis (**Fig S 7**). To obtain better desorption, two types of stirs can be injected separately. This solution is not applicable to the analysis with GC×GC. Because of automatic peak annotation requires technical replications on one sample for applying fisher ratio thresholding. Even with three as minimum number of replication, a separate thermal desorption for two types of stir bars means six times TD-GC×GC-MS measurements for a single sample. The time cost is extremely expensive. However, the application potential of mSBSE cannot be denied. Even with the incomplete desorption, more compounds can be found with mSBSE sampling compare to SPME. The final state of the sampling is equilibrium. It means that the optimization for mSBSE is easier than DHS where breakthrough effect takes place. The protocol of mSBSE is less complex than VALLME and SPE.

2.3.6 Determine the fermentative aromatic compounds

Important fermentative aromatic compounds in beverages have been reviewed by Ferreira and his group [7]. According to this study, the analytical capability of each sample preparation technique followed by GC×GC-MS analysis has been evaluated. Each technique was applied to four types of beverages to check if targeted aromatic compounds can be found and how much signal intensity (peak area) can be obtained. Heat maps were used to visualize the result for white wine (**Fig 4**), beer (**Fig S 8**), red wine (**Fig S 9**), and cider (**Fig S 10**). Based on the result of white wine (**Fig 4**), most aromatic compounds can be found with all the techniques. Comparing the most convenient technique SPME to other techniques, SPME is stronger on determining butyric esters and weaker

on determining lactone and vinyl containing VOCs. If liquid-liquid extraction is applicable, extracting the sample with SPME and VALLME can reach the maximum analytical coverage on those fermentative aromatic compounds. In the beer case, when VALLME is not applicable, SPME together with SPE will provide the maximum analytical coverage. For earlier elutes, such as ethyl acetate, ethyl isobutyrate, diacetyl, 2, 3-pentanedione, solvent delay time must be carefully set. The automatic annotation for lactone containing compounds was not successful. Manually annotation was performed. The difficulty is that many lactone-containing compounds have similar mass spectrums. Using a forward MS matching strategy will easily result in misidentification. Besides, one lactone compound has dozens of names. Even with a correct annotation, linking the result to what is looking for is difficult. A possible solution would be defining a library of interesting compounds, process the reverse matching according to the retention index of the selected compounds.

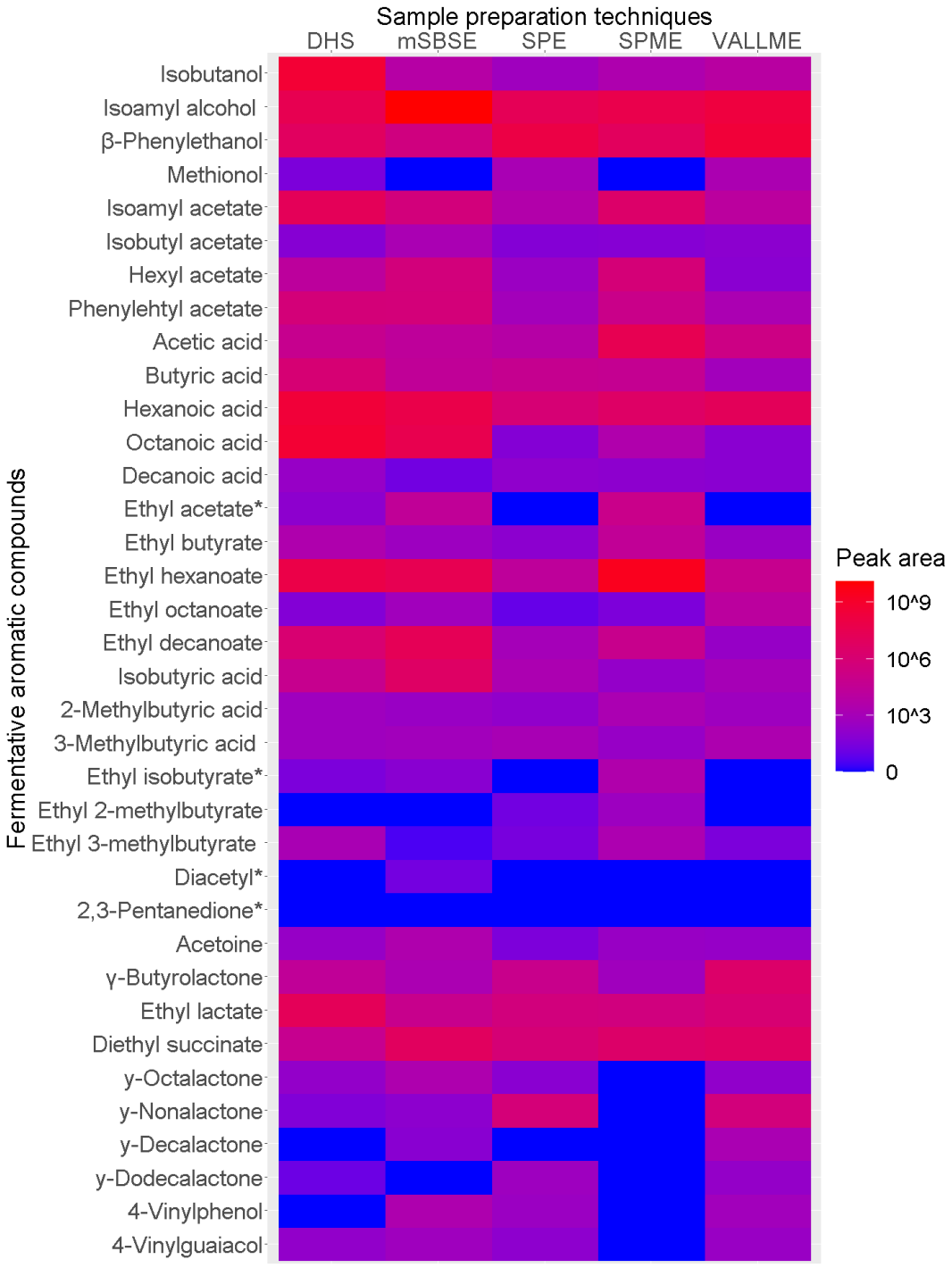


Fig 4 Heat plot for fermentative aromatic compounds determination with different sampling techniques on pooled white wine

2.4 Conclusion

GC×GC-TOFMS is an efficient tool for the analysis of volatile compound. GC×GC oven separation can be efficiently optimized by response surface model. In the case of non-target profiling for complex metabolic samples, the nearest neighbor distance was a suitable measurement for peak dispersion. When the column and detector were saturated by the major metabolites, which is unavoidable if the metabolic sample is measured at one dilution level, incorrect peak deconvolution and mass spectrum construction may happen. This limited the application of GC×GC-TOFMS to metabolites screening. Different sample preparation techniques were compared based on the results of VOC collection for fermented beverages. There wasn't an ideal sample preparation technique to recover all the VOCs from the sample. Furthermore, the VOCs recovered by different techniques were very different. The discussion of the pros and cons of the different techniques in our study can serve as a guide for the further development and improvement of these techniques. Combining the results from different sample preparation techniques is necessary to achieve a higher coverage of global VOC profiling. For the known fermentative aromatic compounds, the best coverage can be reached by using SPME together with SPE for beer, and VALLME for wine and cider.

References

(1) García-Muñoz, S.; Muñoz-Organero, G.; Fernández-Fernández, E.; Cabello, F. Sensory Characterisation and Factors Influencing Quality of Wines Made from 18 Minor Varieties (Vitis Vinifera L.). Food Qual. Prefer. 2014, 32, 241–252. https://doi.org/10.1016/j.foodqual.2013.09.005.

(2) Humia, B. V.; Santos, K. S.; Barbosa, A. M.; Sawata, M.; Mendonça, M. da C.; Padilha, F. F. Beer Molecules and Its Sensory and Biological Properties: A Review. Molecules 2019, 24 (8), 1568. https://doi.org/10.3390/molecules24081568.

(3) Pinu, F. Grape and Wine Metabolomics to Develop New Insights Using Untargeted and Targeted Approaches. Fermentation 2018, 4 (4), 92. https://doi.org/10.3390/fermentation4040092.

(4) Patton, S.; Josephson, D. V. A Method for Determining Significance of Volatile Flavor Compounds in Foods. J. Food Sci. 1957, 22 (3), 316–318. https://doi.org/10.1111/j.1365-2621.1957.tb17017.x.

(5) Ullrich, F.; Grosch, W. Identification of the Most Intense Volatile Flavour Compounds Formed during Autoxidation of Linoleic Acid. Z. Für Lebensm.-Unters. -Forsch. 1987, 184 (4), 277–282. https://doi.org/10.1007/BF01027663.

(6) MR, M.; R, M.-L.; BT, W.; NJ, M.; LM, L. Pinot Noir Aroma: A Sensory/Gas Chromatographic Approach. Dev. Food Sci. 1990, 24, 23–36.

(7) Culleré, L.; López, R.; Ferreira, V. The Instrumental Analysis of Aroma-Active Compounds for Explaining the Flavor of Red Wines. In Red Wine Technology; Elsevier, 2019; pp 283–307. https://doi.org/10.1016/B978-0-12-814399-5.00020-7.

(8) A. Reineccius, G. Flavoromics – the next Frontier?; Marriott Convention Center, 2008.

(9) Nykänen, L.; Suomalainen, H. Aroma of Beer, Wine, and Distilled Alcoholic Beverages; Handbuch der Aromaforschung; D. Reidel ; Distributors for the U.S.A. and Canada, Kluwer Boston: Dordrecht, Holland ; Boston, U.S.A. : Hingham, MA, 1983.

(10) Cordero, C.; Kiefl, J.; Schieberle, P.; Reichenbach, S. E.; Bicchi, C. Comprehensive Two-Dimensional Gas Chromatography and Food Sensory Properties: Potential and Challenges. Anal. Bioanal. Chem. 2015, 407 (1), 169–191. https://doi.org/10.1007/s00216-014-8248-z.

(11) Stefanuto, P.-H.; Perrault, K. A.; Dubois, L. M.; L'Homme, B.; Allen, C.; Loughnane, C.; Ochiai, N.; Focant, J.-F. Advanced Method Optimization for Volatile Aroma Profiling of Beer Using Two-Dimensional Gas Chromatography Time-of-Flight Mass Spectrometry. J. Chromatogr. A 2017, 1507, 45–52. https://doi.org/10.1016/j.chroma.2017.05.064.

(12) Kolb, B.; Ettre, L. S. Theory and Practice of Multiple Headspace Extraction. Chromatographia 1991, 32 (11–12), 505–513. https://doi.org/10.1007/BF02327895.

(13) Soria, A. C.; García-Sarrió, M. J.; Sanz, M. L. Volatile Sampling by Headspace Techniques. TrAC Trends Anal. Chem. 2015, 71, 85–99. https://doi.org/10.1016/j.trac.2015.04.015.

(14) Psillakis, E.; Kalogerakis, N. Developments in Single-Drop Microextraction. TrAC Trends Anal. Chem. 2002, 21 (1), 54–64. https://doi.org/10.1016/S0165-9936(01)00126-1.

(15) Turbulence Phenomena; Elsevier, 1972. https://doi.org/10.1016/B978-0-12-206070-0.X5001-9.

(16) Gupta, A.; Eral, H. B.; Hatton, T. A.; Doyle, P. S. Nanoemulsions: Formation, Properties and Applications. Soft Matter 2016, 12 (11), 2826–2841. https://doi.org/10.1039/C5SM02958A.

(17) Baltussen, E.; Cramers, C.; Sandra, P. Sorptive Sample Preparation – a Review. Anal. Bioanal. Chem. 2002, 373 (1–2), 3–22. https://doi.org/10.1007/s00216-002-1266-2.

(18) David, F.; Ochiai, N.; Sandra, P. Two Decades of Stir Bar Sorptive Extraction: A Retrospective and Future Outlook. TrAC Trends Anal. Chem. 2019, 112, 102–111. https://doi.org/10.1016/j.trac.2018.12.006.

(19) Bojko, B.; Reyes-Garcés, N.; Bessonneau, V.; Goryński, K.; Mousavi, F.; Souza Silva, E. A.; Pawliszyn, J. Solid-Phase Microextraction in Metabolomics. TrAC Trends Anal. Chem. 2014, 61, 168–180. https://doi.org/10.1016/j.trac.2014.07.005.

(20) Andrade-Eiroa, A.; Canle, M.; Leroy-Cancellieri, V.; Cerdà, V. Solid-Phase Extraction of Organic Compounds: A Critical Review (Part I). TrAC Trends Anal. Chem. 2016, 80, 641–654. https://doi.org/10.1016/j.trac.2015.08.015.

(21) Martire, D. E.; Boehm, R. E. Unified Theory of Retention and Selectivity in Liquid Chromatography. 2. Reversed-Phase Liquid Chromatography with Chemically Bonded Phases. J. Phys. Chem. 1983, 87 (6), 1045–1062. https://doi.org/10.1021/j100229a025.

(22) Gritti, F.; Guiochon, G. Adsorption Mechanism in Reversed-Phase Liquid Chromatography. J. Chromatogr. A 2006, 1115 (1–2), 142–163. https://doi.org/10.1016/j.chroma.2006.02.095.

(23) Pap, T.; Horváth, V.; Tolokán, A.; Horvai, G.; Sellergren, B. Effect of Solvents on the Selectivity of Terbutylazine Imprinted Polymer Sorbents Used in Solid-Phase Extraction. J. Chromatogr. A 2002, 973 (1–2), 1–12. https://doi.org/10.1016/S0021-9673(02)01084-1.

(24) Castaldo, R.; Gentile, G.; Avella, M.; Carfagna, C.; Ambrogi, V. Microporous Hyper-Crosslinked Polystyrenes and Nanocomposites with High Adsorption Properties: A Review. Polymers 2017, 9 (12), 651. https://doi.org/10.3390/polym9120651.

(25) Liu, Y.; Su, H.; Song, H.-L. Comparison of Four Extraction Methods, SPME, DHS, SAFE, Versus SDE, for the Analysis of Flavor Compounds in Natto. Food Anal. Methods 2018, 11 (2), 343–354. https://doi.org/10.1007/s12161-017-1005-0.

(26) Picard, M.; Franc, C.; de Revel, G.; Marchand, S. Dual Solid-Phase and Stir Bar Sorptive Extraction Combined with Gas Chromatography-Mass Spectrometry Analysis Provides a Suitable Tool for Assaying Limonene-Derived Mint Aroma Compounds in Red Wine. Anal. Chim. Acta 2018, 1001, 168–178. https://doi.org/10.1016/j.aca.2017.11.074.

(27) Carlin, S.; Vrhovsek, U.; Franceschi, P.; Lotti, C.; Bontempo, L.; Camin, F.; Toubiana, D.; Zottele, F.; Toller, G.; Fait, A.; Mattivi, F. Regional Features of Northern Italian Sparkling Wines, Identified Using Solid-Phase Micro Extraction and Comprehensive Two-Dimensional Gas Chromatography Coupled with Time-of-Flight Mass Spectrometry. Food Chem. 2016, 208, 68–80. https://doi.org/10.1016/j.foodchem.2016.03.112.

(28) Ortega, C.; López, R.; Cacho, J.; Ferreira, V. Fast Analysis of Important Wine Volatile Compounds. J. Chromatogr. A 2001, 923 (1–2), 205–214. https://doi.org/10.1016/S0021-9673(01)00972-4.

(29) Ochiai, N.; Sasamoto, K.; Ieda, T.; David, F.; Sandra, P. Multi-Stir Bar Sorptive Extraction for Analysis of Odor Compounds in Aqueous Samples. J. Chromatogr. A 2013, 1315, 70–79. https://doi.org/10.1016/j.chroma.2013.09.070.

(30) Vrhovsek, U.; Lotti, C.; Masuero, D.; Carlin, S.; Weingart, G.; Mattivi, F. Quantitative Metabolic Profiling of Grape, Apple and Raspberry Volatile Compounds (VOCs) Using a GC/MS/MS Method. J. Chromatogr. B 2014, 966, 132–139. https://doi.org/10.1016/j.jchromb.2014.01.009.

(31) Welke, J. E.; Manfroi, V.; Zanus, M.; Lazarotto, M.; Alcaraz Zini, C. Characterization of the Volatile Profile of Brazilian Merlot Wines through Comprehensive Two Dimensional Gas Chromatography Time-of-Flight Mass

Spectrometric Detection. J. Chromatogr. A 2012, 1226, 124–139. https://doi.org/10.1016/j.chroma.2012.01.002.

(32) Nowik, W.; Héron, S.; Bonose, M.; Nowik, M.; Tchapla, A. Assessment of Two-Dimensional Separative Systems Using Nearest-Neighbor Distances Approach. Part 1: Orthogonality Aspects. Anal. Chem. 2013, 85 (20), 9449–9458. https://doi.org/10.1021/ac4012705.

(33) Mommers, J.; van der Wal, S. Two Metrics for Measuring Orthogonality for Two-Dimensional Chromatography. J. Chromatogr. A 2019, 1586, 101–105. https://doi.org/10.1016/j.chroma.2018.11.081.

(34) Bezerra, M. A.; Santelli, R. E.; Oliveira, E. P.; Villar, L. S.; Escaleira, L. A. Response Surface Methodology (RSM) as a Tool for Optimization in Analytical Chemistry. Talanta 2008, 76 (5), 965–977. https://doi.org/10.1016/j.talanta.2008.05.019.

(35) Bean, H. D.; Hill, J. E.; Dimandja, J.-M. D. Improving the Quality of Biomarker Candidates in Untargeted Metabolomics via Peak Table-Based Alignment of Comprehensive Two-Dimensional Gas Chromatography–Mass Spectrometry Data. J. Chromatogr. A 2015, 1394, 111–117. https://doi.org/10.1016/j.chroma.2015.03.001.

Supplementary

Table S 2: Standard solution of 131 common aromatic compounds

LTPRI	Supplier	Compound	CAS	Unique Mass	Classes
1003	Fluka	2-butanol	78-92-2	59	alcohol
1033	Aldrich	butanoic acid, ethyl ester	105-54-4	71	ester
1042	Fluka	-pinene	127-91-3	93	terpenes
1084	Fluka	2-hexanone	591-78-6	77	ketone
1092	SAFC	*trans*-2-methyl-2-butenal	497-03-0	84	aldehyde
1100	Aldrich	3-methyl-2-butanol	26184-62-3	55	alcohol
1112	Fluka	3-pentanol	584-02-1	59	alcohol
1115	Supelco	n-butanol	71-36-3	59	alcohol
1116	Fluka	-pinene	127-91-3	93	terpenes
1119	SAFC	2-methylbutyl-acetate	624-41-9	70	ester
1119	Fluka	isopentyl acetate	123-92-2	70	ester
1130	SAFC	2-methyl-2-pentenal	623-36-9	98	aldehyde
1138	Fluka	ethyl valerate	539-82-2	85	ester
1139	Aldrich	propanoic acid, butyl ester	590-01-2	75	ester
1171	SAFC	amyl acetate	123-92-2	61	ester
1173	SAFC	2-methyl-1-butanol	137-32-6	56	alcohol
1184	Fluka	methyl caproate	106-70-7	74	ester
1186	Aldrich	(R)-(-)-α-phellandrene	99-83-2	91	terpenes
1187	Fluka	2-heptanone	110-43-0	58	ketone
1211	SAFC	2-pentylfuran	3777-69-3	81	ethers
1213	Fluka	limonene	138-86-3	68	terpenes
1218	SAFC	butanoic acid, butyl ester	109-21-7	71	ester
1220	Aldrich	ethyl hexanoate	123-66-0	88	ester
1234	Fluka	1,8-cineole	470-82-6	81	terpenes
1234	Aldrich	butyl 2-methylbutanoate	15706-73-7	103	ester
1238	Fluka	1-pentanol	71-41-0	42	alcohol
1250	SAFC	acetoin	513-86-0	43	ketone
1269	SAFC	hexyl acetate	142-92-7	61	ester
1276	Aldrich	octanal	124-13-0	57	aldehyde
1284	Fluka	2-heptanol	543-49-7	45	alcohol
1287	SAFC	isoamyl butyrate	106-27-4	71	ester
1291	Fluka	2-octanone	111-13-7	58	ketone
1295	Fluka	p-cymene	99-87-6	119	terpenes

1312	SAFC	terpinolene	586-62-9	121	terpenes
1316	SAFC	hexyl propionate	2445-76-3	57	ester
1319	Supelco	1-hexanol	111-27-3	56	alcohol
1325	SAFC	*trans*-2-hexenyl acetate	2497-18-9	43	ester
1340	Aldrich	6-methyl-5-hepten-2-one	110-93-0	43	ketone
1341	SAFC	3-methyl-1-pentanol	589-35-5	56	alcohol
1352	Fluka	*cis*/*trans* rose oxide	16409-43-1	139	terpenes
1367	SAFC	*trans*-3-hexenol	928-97-2	56	alcohol
1369	Aldrich	*cis*-3-hexen-1-ol	928-96-1	67	alcohol
1369	SAFC	heptyl acetate	112-06-1	61	ester
1380	SAFC	*trans*-2-hexen-1-ol	928-95-0	54	alcohol
1382	Fluka	2-nonanone	821-55-6	58	ketone
1385	Aldrich	2-octanol(IS)	123-96-6	45	alcohol
1400	Aldrich	2-isopropyl-3-methoxypyrazine	93905-03-4	137	pyrazine
1402	SAFC	*trans*,*trans*-2,4-hexadienal	142-83-6	65	aldehyde
1407	SAFC	hexyl butyrate	2639-63-6	71	ester
1410/ 1471	Aldrich	*cis* and *trans* linalool oxide (Ox A+Ox B)	60047-17-8	59	terpenes
1413	SAFC	hexanoate butyl-	626-82-4	99	ester
1418	SAFC	hexyl 2-methylbutanoate	10032-15-2	103	ester
1424	Fluka	1-heptanol	111-70-6	57	alcohol
1424	Fluka	acetic acid	64-19-7	60	acid
1428	Aldrich	ethyl caprylate	106-32-1	88	ester
1430	SAFC	*trans*-2-octenal	2548-87-0	59	aldehyde
1462	Fluka	sabinene hydrate	80-56-8	93	terpenes
1465	Fluka	2-ethyl-1-hexanol	104-76-7	57	alcohol
1471	SAFC	6-methyl-5-hepten-2-ol	1569-60-4	95	alcohol
1472	SAFC	octyl acetate	103-09-3	61	ester
1475	Aldrich	2-sec-butyl-3-methoxypyrazine	24168-70-5	138	pyrazine
1480	SAFC	*trans*,*trans*-2,4-heptadienal	4313-03-5	81	aldehyde
1491	SAFC	citronellal	106-23-0	71	terpenes
1497	Aldrich	decanal	112-31-2	111	aldehyde
1507	Aldrich	theaspirane A	36431-72-8	138	terpenes
1507	Aldrich	theaspirane B	36431-72-8	138	terpenes
1518	SAFC	2-isobutyl-3-methoxypyrazine	24683-00-9	124	pyrazine
1520	Aldrich	benzaldehyde	100-52-7	106	aldehyde
1532	Fluka	(-)-linalool	78-70-6	71	terpenes
1532	Aldrich	camphor	76-22-2	81	terpenes

1560	SAFC	nonyl acetate	143-13-5	61	ester
1572	Aldrich	5-methylfurfural	620-02-0	110	aldehyde
1590	SAFC	*trans*-2-octen-1-ol	18409-17-1	70	alcohol
1593	Aldrich	terpinen-4-ol	562-74-3	71	terpenes
1598	Aldrich	hexanoic acid, hexyl ester	6378-65-0	99	ester
1600	Aldrich	undecylicaldehyde	112-44-7	82	aldehyde
1615	SAFC	*cis*-5-octen-1-ol	64275-73-6	57	alcohol
1621	Aldrich	decanoate <ethyl->	110-38-3	88	ester
1637	Aldrich	2-methyl-2,4-pentanediol	107-41-5	59	alcohol
1638	Aldrich	phenylacetaldehyde	122-78-1	120	aldehyde
1642	Aldrich	*trans*-caryophyllene	87-44-5	91	terpenes
1656	Fluka	acetophenone	98-86-2	105	ketone
1673	SAFC	1-phenylethyl acetate	93-92-5	104	ester
1685	Fluka	2-methylbutyric acid	116-53-0	74	acid
1693	Aldrich	-humulene	6753-98-6	93	terpenes
1695	SAFC	dodecanal	112-54-9	71	aldehyde
1698	Fluka	(+)-α-terpineol	98-55-5	59	terpenes
1723	Aldrich	1-decanol	112-30-1	70	alcohol
1726	Aldrich	-citral	141-27-5	84	terpenes
1730	Fluka	methyl salicylate	119-36-8	120	ester
1735	Aldrich	geranyl acetate	105-87-3	68	terpenes
1788	SAFC	acetic acid, 2 phenylethyl ester	103-45-7	104	ester
1791	Fluka	myrtenol	515-00-4	108	alcohol
1795	SAFC	hexyl octanoate	1117-55-1	145	ester
1811	Bedoukian research, inc.	dihydro -ionone	31499-72-6	95	terpenes
1814	SAFC	-damascenone	23726-93-4	121	terpenes
1815	SAFC	geraniol	106-24-1	41	terpenes
1833	Fluka	-ionone	127-41-3	121	terpenes
1853	SAFC	linalyl isovalerate	1118-27-0	93	terpenes
1862	Aldrich	benzyl alcohol	100-51-6	108	alcohol
1862	Aldrich	guaiacol	90-05-1	109	alcohol
1880	Fluka	hexanoic acid	142-62-1	60	acid
1888	Fluka	-ionol	25312-34-9	95	terpenes
1890	Aldrich	2-phenylethanol	60-12-8	91	alcohol
1965	Fluka	(-)-caryophyllene oxide	1139-30-6	79	terpenes
1975	Aldrich	-ionone	79-77-6	177	terpenes
2027	Fluka	ethyl myristate	124-06-1	88	ester
2100	Fluka	octanoic acid	124-07-2	60	acid
2102	Aldrich	ethyl cinnamate	4610-69-9	131	ester

2123	Aldrich	methyl N-methylanthranilate	85-91-6	104	ester
2162	SAFC	geranic acid	459-80-3	69	terpenes
2168	SAFC	eugenol	97-53-0	164	alcohol
2170	Fluka	1-tetradecanol	112-72-1	97	alcohol
2180	Aldrich	2-methoxy-4-vinylphenol	7786-61-0	164	phenol
2181	Fluka	methyl anthranilate	134-20-3	119	ester
2187	SAFC	2-aminoacetophenone	551-93-9	120	ketone
2187	SAFC	decalactone	705-86-2	99	ketone
2200	Fluka	nonanoic acid	112-05-0	60	acid
2243	Fluka	decanoic acid	334-48-5	60	acid
2345	SAFC	methyl jasmonate	42536-97-0	83	ester
2354	Aldrich	*trans*-*trans*-farnesol	16106-95-9	69	terpenes
2410	Fluka	benzoic acid	65-85-0	74	acid
2410	SAFC	benzophenone	119-61-9	105	ketone
2486	SAFC	5-(hydroxymethyl) furfural	67-47-0	97	aldehyde
2566	Carlo erba	vanillin	99-83-2	151	aldehyde
2640	SAFC	acetovanillon	498-02-2	151	ketone
2691	Fluka	myristic acid	544-63-8	60	acid
2771	SAFC	vanillylacetone	122-48-5	137	ketone
2880	Aldrich	palmitic acid	57-10-3	60	acid
2061	Fluka	epiglobulol	88728-58-9	82	terpenes
2280	synthesized	rotundone	18374-76-0	147	terpenes

Table S 3: Experimental parameters of DHS sampling

DHS		TDU	
Incubation time (min)	20	Initial temperature (ºC)	50
Incubation temperature (ºC)	40	End temperature (ºC)	300
Agitator speed (rpm)	500	Rate (ºC/min)	120
Trapping volume (mL)	1250	Hold time (min)	6
Trapping flow (mL/min)	100	Transfer temperature (ºC)	300
Trapping temperature (ºC)	40	CIS	
Dry volume (mL)	50	Initial temperature (ºC)	0
Dry flow (mL/min)	10	End temperature (ºC)	250
Dry temperature (ºC)	25	Rate (ºC/s)	12
		Hold time (min)	10

Table S 4: TDU and CIS parameters for mSBSE analysis

TDU		CIS	
Initial temperature (ºC)	30	Initial temperature (ºC)	0

End temperature (ºC)	220	End temperature (ºC)	250
Rate (ºC/min)	120	Rate (ºC/s)	12
Hold time (min)	10	Hold time (min)	10
Transfer temperature (ºC)	300		

Table S 5: Experimental parameters of SPME sampling

Sample preparation		Fiber conditioning	
Incubation Temperature (ºC)	30	Bakeout Temperature (ºC)	270
Incubation Time (min)	5	Pre bakeout time (min)	4.50
Agitator speed (rpm)	500	Bakeout Penetration (mm)	67
Desorption time (min)	4		
Desorption Temperature (ºC)	250		

Table S 6: Optimization of GC×GC separation condition. Results of experiment 6, 12, 25 or 26 were not used because of the heavily wrap-around

Exp	Column flow (mL/min)	Temp program (ºC/min)	2nd oven temp offset (ºC)	Modulation temp offset (ºC)	Modulation time (s)	Hot pulse time (s)	Median NND (s)
1	1.6	7	1	10	4	0.6	4.03
2	1.6	3	1	10	8	1.2	8.22
3	1.2	5	3	10	4	1	8.01
4	0.8	7	5	6	4	0.6	4.02
5	1.6	7	1	10	8	2.8	8.00
6	0.8	3	5	6	4	1.4	
7	0.8	5	1	6	8	2.8	8.00
8	1.2	7	5	6	8	2.8	8.00
9	1.6	3	5	6	8	2	8.28
10	0.8	3	1	10	6	2.1	12.01
11	1.2	5	5	8	6	0.9	6.02
12	1.6	3	1	8	4	1.4	
13	0.8	7	5	10	4	1.4	4.02
14	1.6	7	5	8	4	1	4.07
15	1.2	3	3	8	8	2.8	12.18
16	1.6	5	3	6	6	2.1	6.03
17	1.6	7	5	10	8	1.2	8.00
18	1.6	3	5	10	6	2.1	12.04
19	0.8	7	1	10	8	1.2	8.00
20	0.8	5	1	8	4	0.6	8.00
21	1.2	3	1	6	6	1.5	12.04
22	1.6	7	1	6	8	1.2	8.00
23	0.8	7	3	8	6	1.5	6.00
24	1.2	7	1	6	4	1.4	4.01
25	0.8	3	5	10	4	0.6	
26	1.6	3	3	6	4	0.6	

27	0.8	5	5	10	8	2	8.00
28	0.8	3	3	6	8	1.2	8.12

Table S 7: Model fitting summary

RSquare	1
RSquare Adj	0.999997
Root Mean Square Error	0.0042
Mean of Response	7.629552
Observations (or Sum Wgts)	24

Table S 8: Model effect summary

Source	LogWorth		PValue
temp program(3,7)	6.238		0.00000
modulation time*temp program	5.940		0.00000
temp program*temp program	5.466		0.00000
modulation time*modulation time	5.315		0.00000
modulation temp*modulation temp	5.046		0.00001
temp program*hot pulse time	5.031		0.00001
column flow(0.8,1.6)	4.909		0.00001
column flow*2nd oven temp offset	4.873		0.00001
hot pulse time(0.15,0.35)	4.845		0.00001
modulation time(4,8)	4.771		0.00002
modulation temp*hot pulse time	4.488		0.00003
2nd oven temp offset*modulation temp	4.363		0.00004
hot pulse time*hot pulse time	3.959		0.00011
modulation temp(6,10)	3.451		0.00035
2nd oven temp offset*hot pulse time	3.365		0.00043
modulation time*hot pulse time	3.243		0.00057
column flow*modulation temp	3.106		0.00078
column flow*modulation time	3.086		0.00082
column flow*hot pulse time	2.902		0.00125
2nd oven temp offset(1,5)	2.801		0.00158
modulation time*2nd oven temp offset	2.532		0.00293

Table S 9: Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	6.8892815	0.005133	1342.1	<.0001*
column flow(0.8,1.6)	-0.357395	0.001255	-284.8	<.0001*
modulation time(4,8)	-0.568177	0.002339	-243.0	<.0001*
2nd oven temp offset(1,5)	-0.032752	0.001304	-25.13	0.0016*
temp program(3,7)	-3.566881	0.002712	-1315	<.0001*
modulation temp(6,10)	0.0809492	0.001523	53.16	0.0004*
hot pulse time(0.15,0.35)	0.6585843	0.002489	264.57	<.0001*
column flow*modulation time	-0.053087	0.001521	-34.89	0.0008*
modulation time*modulation time	1.5732576	0.003463	454.31	<.0001*
column flow*2nd oven temp offset	-0.525036	0.001922	-273.1	<.0001*
modulation time*2nd oven temp offset	-0.027709	0.001504	-18.42	0.0029*
modulation time*temp program	2.8814911	0.003087	933.53	<.0001*

Term	Estimate	Std Error	t Ratio	Prob>|t|
temp program*temp program	2.3224086	0.004293	540.92	<.0001*
column flow*modulation temp	0.0902685	0.002527	35.72	0.0008*
2nd oven temp offset*modulation temp	0.3464299	0.002281	151.86	<.0001*
modulation temp*modulation temp	-1.105736	0.003314	-333.6	<.0001*
column flow*hot pulse time	0.0402426	0.001425	28.24	0.0013*
modulation time*hot pulse time	0.1404141	0.003357	41.82	0.0006*
2nd oven temp offset*hot pulse time	-0.075921	0.001577	-48.15	0.0004*
temp program*hot pulse time	-0.739247	0.002257	-327.6	<.0001*
modulation temp*hot pulse time	-0.339576	0.001936	-175.4	<.0001*
hot pulse time*hot pulse time	-0.260668	0.002732	-95.40	0.0001*

Fig S 1 Prediction of median NND by the Box-Behnken surface response model under optimized oven condition

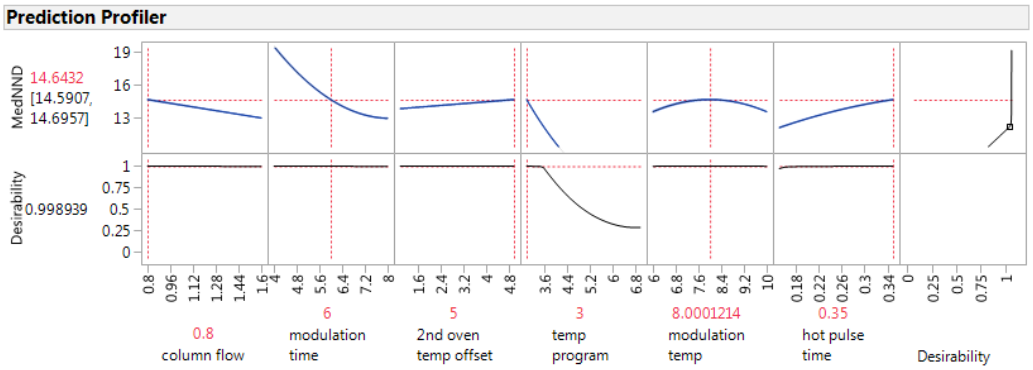


Fig S 2 MS similarity (matching score) distribution

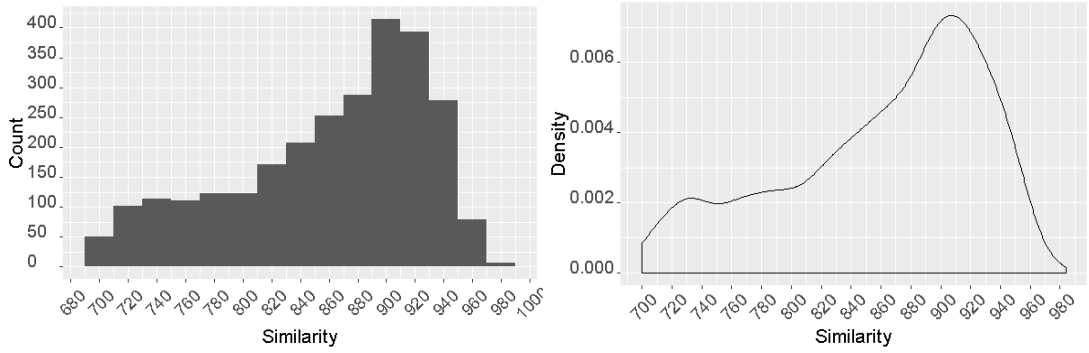


Fig S 3 Over deconvolution of peak 1-phenylethyl acetate. a) GC×GC chromatogram of peak 1-phenylethyl acetate, b) constructed mass spectrum of peak 1-phenylethyl acetate, c) library mass spectrum of 1-phenylethyl acetate (Nist), d) constructed mass spectrum of unknown peak

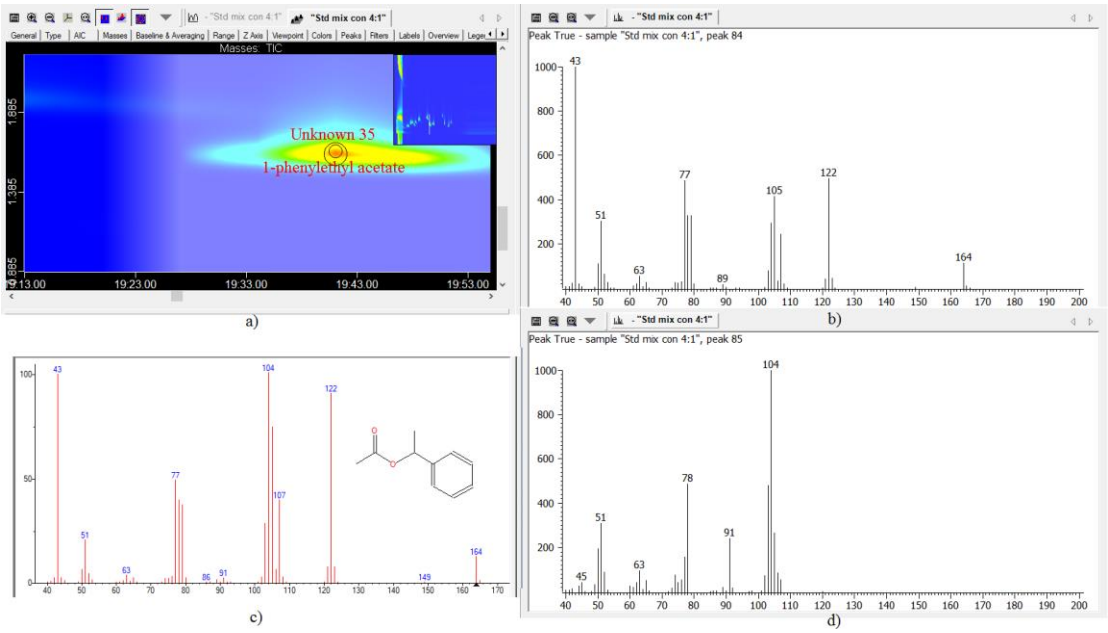


Fig S 4 Comparison of the aligned peak number by chemical classes for applying different sampling techniques on pooled red wine

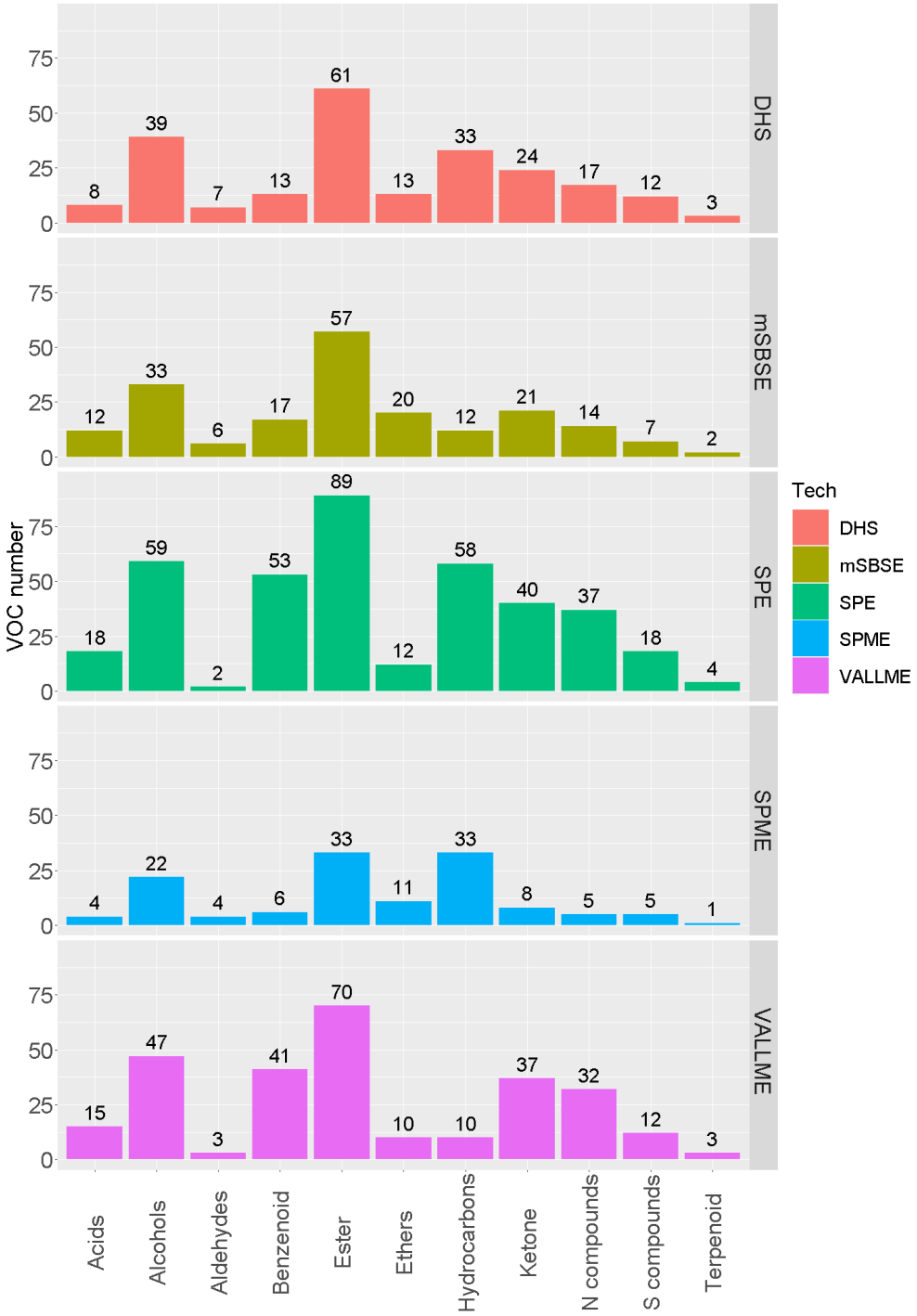


Fig S 5 Comparison of the aligned peak number by chemical classes for applying different sampling techniques on pooled cider

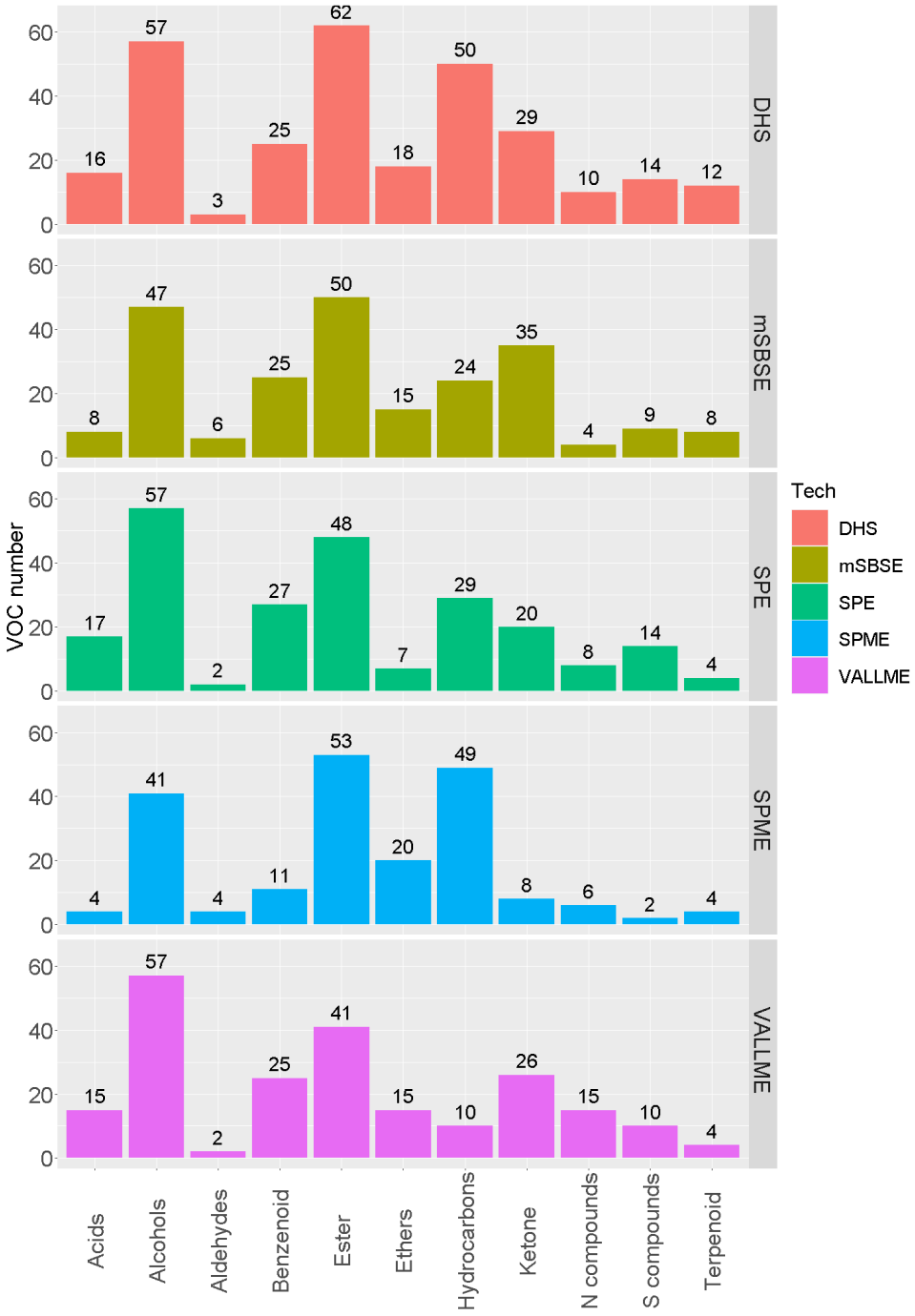


Fig S 6 Comparison of the aligned peak number by chemical classes for applying different sampling techniques on pooled white wine

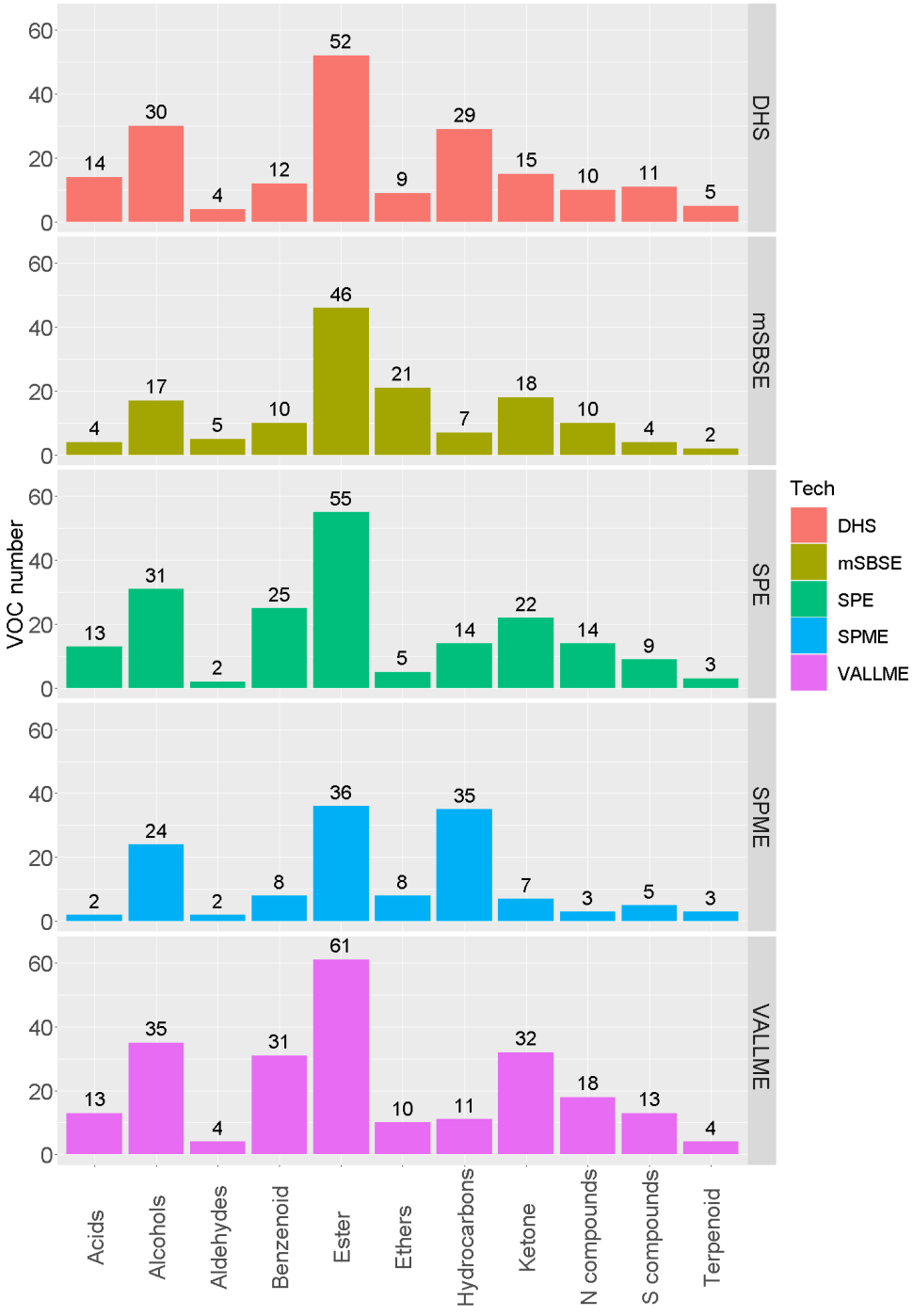
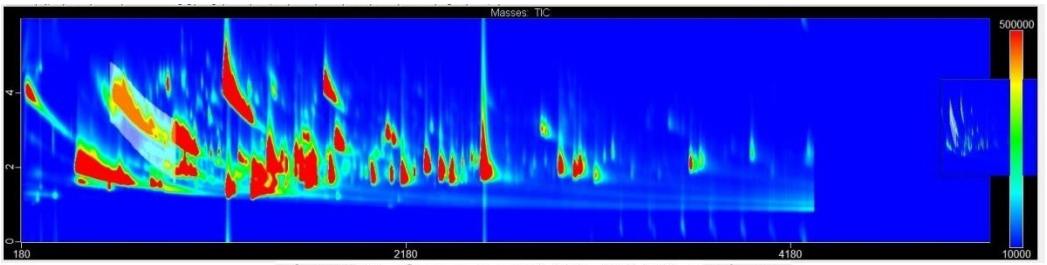
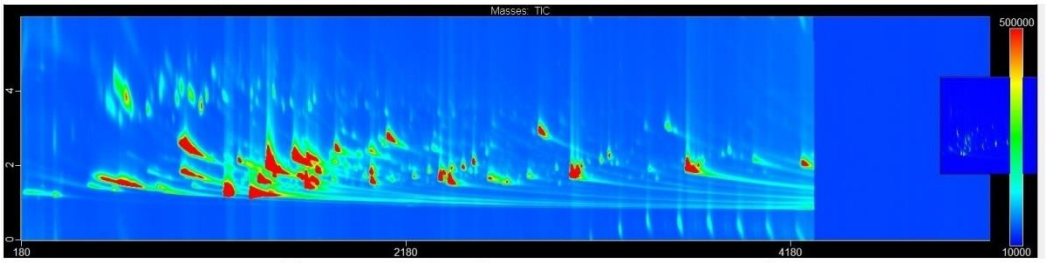


Fig S 7 mSBSE insufficient TD



a) mSBSE first time TD GCxGC-MS analysis



b) mSBSE second time TD GCxGC-MS analysis

Fig S 8 Heatplot for fermentative aromatic compounds determination with different sampling techniques on pooled beer

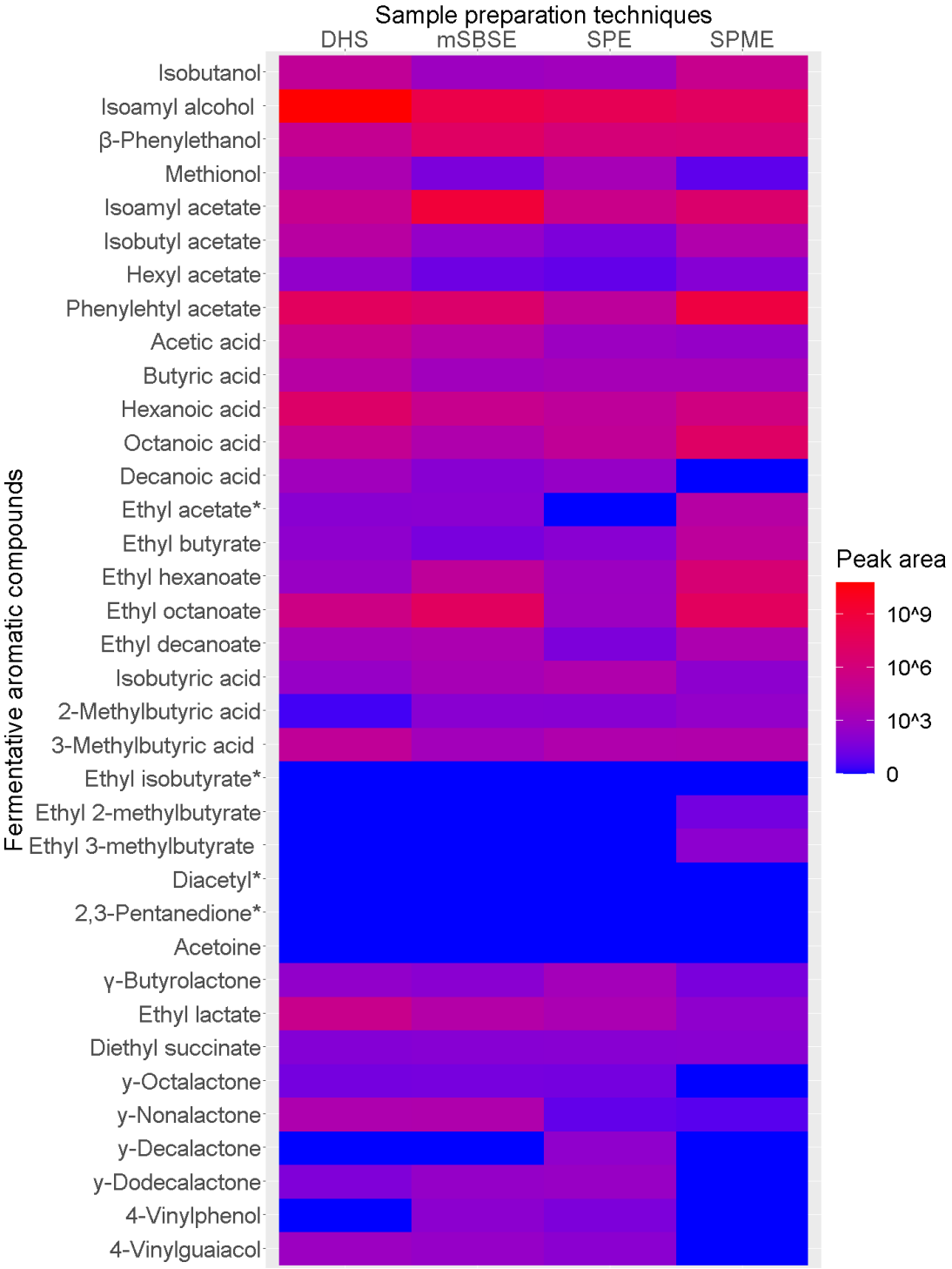


Fig S 9 Heatplot for fermentative aromatic compounds determination with different sampling techniques on pooled red wine

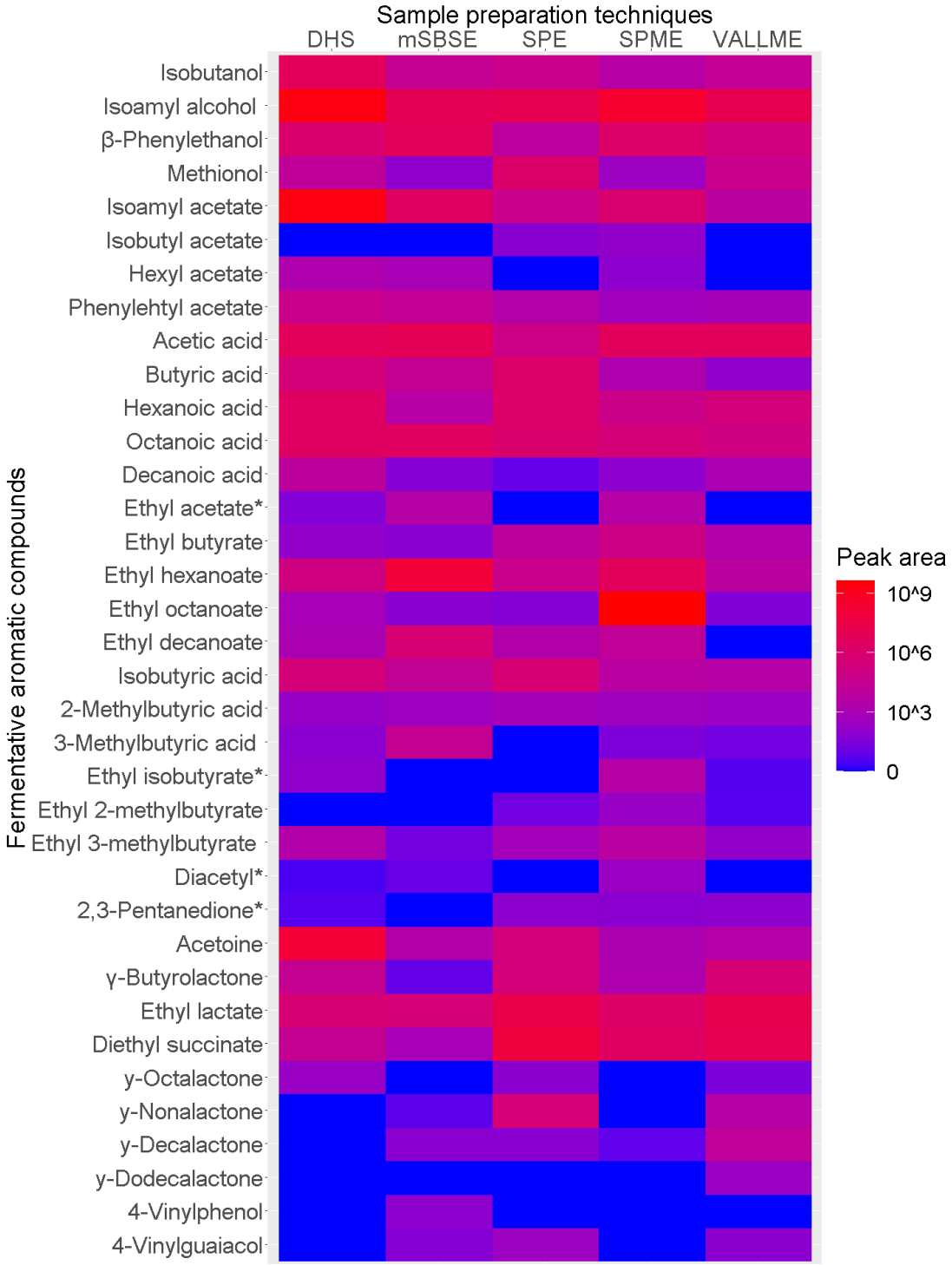
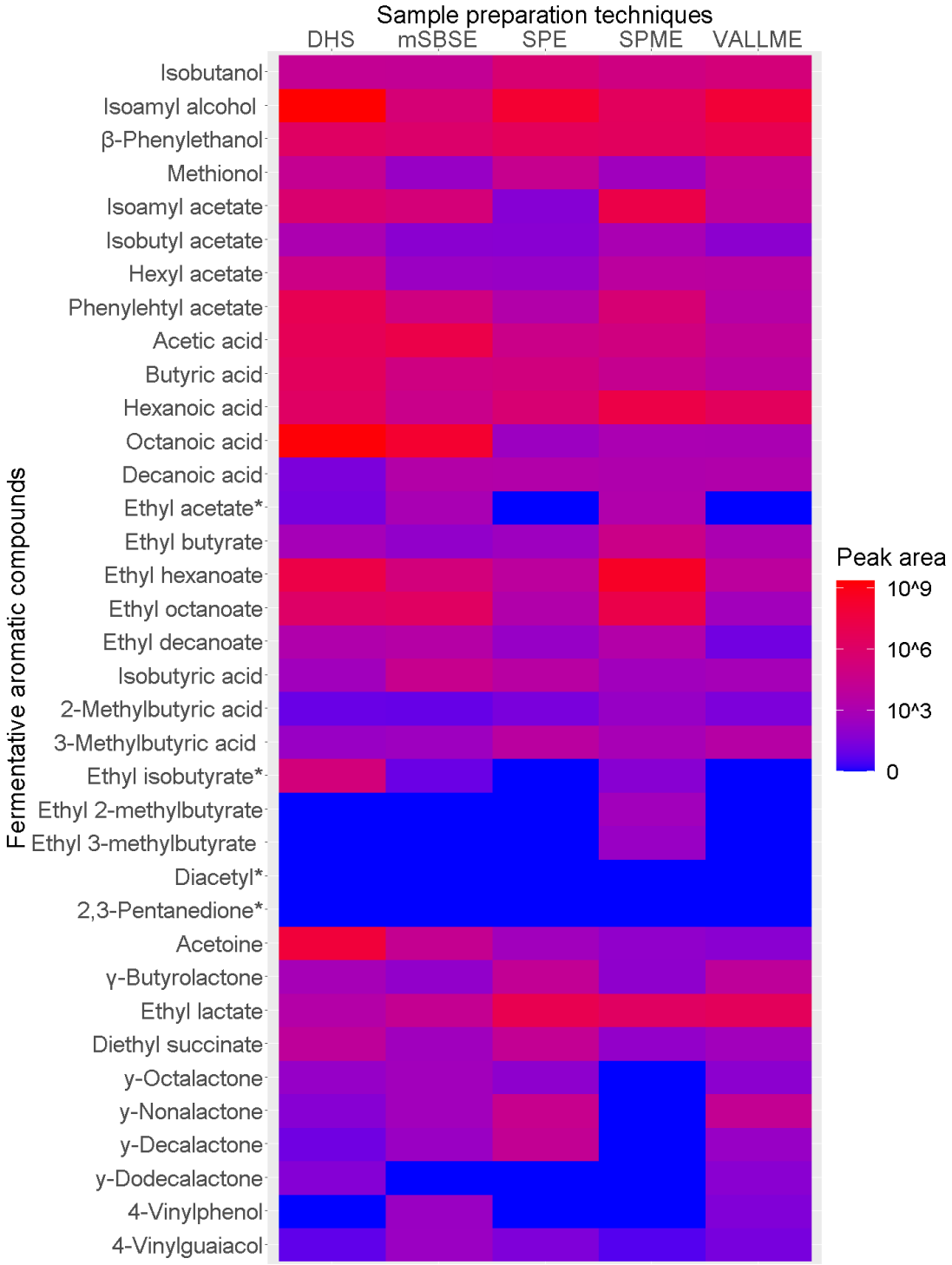


Fig S 10 Heatplot for fermentative aromatic compounds determination with different sampling techniques on pooled cider



3 Application of a target-guided data processing approach in saturated peaks correction of GC×GC analysis

Penghan Zhang, Silvia Carlin, Pietro Franceschi, Fulvio Mattivi, and Urska Vrhovsek

(revision submitted)

Abstract

Detector and column saturation are problematic in GC×GC data analysis. This limits the application of GC×GC to metabolomics research. To address the problems caused by detector and column saturation, we propose a two-stage data processing strategy that will incorporate a targeted data processing and cleaning approach upstream of the “standard” untargeted analysis. By using the retention time and MS stored in a library, the annotation and quantification of the targeted saturated peaks have been significantly improved. After subtracting the non-perfected signals caused by saturation, peaks of co-elutes can be annotated more accurately. Our research shows that the target-guided method has broad application prospects in the data analysis of GC×GC chromatograms of complex samples.

3.1 Introduction

Metabolomics technologies and the data generated by them result in a better understanding of the metabolism of many biological systems. It helps to reveal the biological linkage between genetic sequence and physiological characterization. Consequently, the biological outcome can be controlled with higher accuracy and reproducibility. The metabolomic approach has been widely applied in many fields and demonstrated successful results: microbes, food quality and crop production, animals, and human health.[1–4]

Compared to other analytical tools, such as nuclear magnetic resonance (NMR), liquid chromatography-mass spectrometry (LC-MS), and capillary electrophoresis–mass spectrometry (CE-MS), GC-MS has the best overall performance in efficiency, sensitivity, and reliability.[5] Hence, it's widely used in metabolomics.[6] However, the chromatogram of a 1DGC of a complex biosample is usually overcrowded. And an analytical method that can provide higher resolution and more separation space is usually desirable. Comprehensive two-dimensional gas chromatography (GC×GC) provides much increased separation capacity, chemical selectivity, and sensitivity for the analysis of metabolites present in complex samples. By adding one more dimension, the peak capacity is typically increased by 30 times.[7] Welthagen demonstrated the power of GC×GC in the profiling of mouse spleen tissue metabolites. The GC×GC analysis identified almost three times as many metabolites as 1DGC.[8] This increased detection potential is extremely promising, but it also implies an additional layer of complexity in terms of data processing, in particular when relying on automatic approaches.

The complete data processing pipeline is similar to the one typically found in MS-based metabolomics and includes noise cleaning, peak picking, peak alignment, and annotation. Among the previous steps, peak picking is the most

crucial one and it becomes more challenging in the presence of two-dimensional separation. In particular, considering that during metabolomic analysis the complexity of the chemical composition and the variation in the chemical content result in detector and column saturations for the major compounds. Column saturation will become more serious, especially when a narrower column diameter is used for the second dimensional separation. The narrow column effect can be seen from Figure S1 to S8. Playing with the column parameter is definitely a solution for saturation during targeted analysis. However, in a semi-quantitative untargeted study, increasing the sample concentration can improve the potential analytical sensitivity for minor compounds if the saturation of the major compound is not causing serious errors. Some chromatograms of different sample preparation techniques are given in the supplementary, Figure S9 to S13. Chromatograms of common beverage VOCs obtained under different GC×GC conditions are also given, Figure S14 to S19.

In the case of GC×GC, several peak-picking strategies have been proposed: matched filtering, local maximum followed by PARAFAC unmixing, and continuous wavelet transform.[9–11] However, when saturation occurs, these peak-picking strategies show their limitations, producing false peak splitting which results in incorrect deconvolution, and, finally, in incorrect annotation and quantification.[12] Take the example of PARAFAC, which is recognized as the most suitable analysis method for deconvolution of GC×GC-MS data in the field.[13] Errors may occur at the second step of PARAFAC approach, peak apex locating, due to the saturation.[10] Considering that for a biosample, the contents of major and minor compounds may vary by 9 magnitudes, concentration techniques are generally applied to allow the detection of trace compounds in a reliable way.[14] Unfortunately, concentration techniques do not selectively act on minor compounds. If the minor compounds are well sampled, major compounds are over-concentrated leading to column/detector saturation. The experimental workaround to this problem is to perform dilutions and measure

major and trace compounds at different dilution levels, but this choice turns out to be impractical for technical and economic reasons in the case of large scale investigations (three to five technical replications would be required for each sample). The core idea of our paper is that, fortunately, for each metabolomic analysis, the major compounds which mainly induce the column and/or detector saturation are often known, because the major components of the biosample are almost invariably known. It would then be possible to propose a “targeted” optimization of the data processing to minimize the impact of column saturation on the quantification of this targeted list of most relevant compounds. To address the problems caused by detector and column saturation, we propose a two-stage data processing strategy that will incorporate a targeted data processing and cleaning approach upstream of the “standard” untargeted analysis. To the best of our knowledge, even if many untargeted approaches to analyze GC×GC experiments have been proposed[15], all of them focus on the improvement and refinement of a pure untargeted approach which does not take into account the saturation effect. In our proposal, with a predefined library, the annotation of saturation peaks can be achieved more robustly and accurately. Then, signals of saturated peaks are subtracted from the chromatogram. Finally, the remaining unannotated signal in the chromatogram can be processed with a general untargeted approach.

3.2 Experimental section

3.2.1 Samples and reagents

To benchmark the outcome of the targeted analysis and data subtraction approach, a dilution series of standard solutions (2, 20, 200, and 2000 mg/L) was prepared. The column saturation was simulated with concentrated standard solution (200, 2000 mg/L). According to our experience, annotations for esters are vulnerable to saturations. Hence, the prepared standard solution consisted of 5 esters, from apolar to polar, according to their retention time

indices (RTI): *cis*-3-hexenyl acetate (RTI=1311), hexyl 2-methylbutanoate (RTI=1418), 2-phenethyl acetate (RTI=1788), ethyl phenylacetate (RTI=1823), ethyl cinnamate (RTI=2102). The RTI for standard polar columns was obtained from PubChem. All chemicals were purchased from Sigma Aldrich. A pooled white wine (mixture of Riesling, Müller-Thurgau, Manzoni Bianco, Sauvignon Blanc, Veltliner, and Gewürztraminer) was used to confirm the performance of this targeted analysis approach in the real-life analysis. The pooled white wine was diluted 10 times to obtain the result of the unsaturated condition. To create the saturated condition, for each 0.5 mL diluted white wine, 10 µL of 200 mg/L standard solution was added.

3.2.2 Samples preparation and injection

Standard solutions were analyzed by the liquid injection mode. 1 µL was injected for each sample. Liquid injections were performed with Gerstel MPS2 auto-sampler monitored by ChromaTOF (Leco, St Joseph, MI, USA). Pooled white wine samples were analyzed by Solid-phase microextraction. 0.5 mL of sample was mixed with 0.5 g NaCl in a 20 mL headspace vial. Before the analysis, 25 µL of internal standard solution (2-octanol in ethanol at a concentration of 2 mg/L) was added. A 2 cm 50/30 µm DVB/CAR/PDMS fiber was used (Supelco), conditioned according to the manual. Other details can be found in the literature.[16] Each sample was analyzed in 3 replications.

3.2.3 Instrument and data processing

Samples were injected into the GC × GC system (Agilent 7890 A, Agilent Technologies, Santa Clara, CA) in splitless mode. Separation was performed by VF-Wax column (100% polyethylene glycol; 30 m x 0.25 mm x 0.25 µm, Agilent J&W Scientific Inc., Folsom, CA) in the 1st dimension and Rxi-17Sil MS 1.50 m x 0.15 mm x 0.15 µm, Restek Bellefonte, USA) in the 2nd dimension. A

non-moving quad jet dual-stage thermal modulator was used to couple the two columns. The MS signal was obtained with the Pegasus IV time-of-flight mass spectrometer (Leco Corporation, St. Joseph, MI).

The applied column flow was 1.2 mL/min. The oven temperature was programmed from 40 ºC (4 min holding time) to 250 ºC at a rate of 6 ºC/min. The final temperature of 250 ºC was reached and then maintained for 15 min. The temperature offsets of the second oven and modulator were set at +5 ºC and +15 ºC, respectively. Within the 7-second-modulation time, 1.4 seconds were used for a hot pulse. The transfer line was kept at 250 ºC. The TOFMS was operated in electron ionization mode at 70 eV. The ion source temperature was 230 ºC. The acquisition frequency was 200 Hz within a mass scan range from 40 m/z to 350 m/z. The detector voltage was -1341 V. A 7-min acquisition delay was applied for liquid injection.

GC×GC-MS data acquisition and untargeted processing were achieved with LECO ChromaTOF (Version 4.22). ChromaTOF performs the m/z alignment automatically before the signal processing, rounding the measured m/z value to the nearest integer. The baseline offset was 1 for the signal pre-processing. The expected peak width was 0.8 s. For the peak picking, the signal to noise ratio was 100. And minimal 5 ion fragments were required.[17] Peak annotation was achieved by comparing the constructed mass spectrum (MS) to the reference spectrum in the database. Used MS databases were NIST/EPA/NIH 11, Wiley 8 and the FFNSC 2. The MS similarity threshold for the peak annotation was 700. Inter-measurement alignment was perform by Statistical Compare package (ChromaTOF build-in) to improve the peak annotation and quantification.

3.2.4 Saturated peak subtraction

Unsaturated and saturated chromatograms were obtained from the analysis of diluted pooled white wine and diluted pooled white wine plus standard solution (10 µL 200 mg/L), respectively. The saturation subtracted chromatogram is the output of the proposed target method applied to the saturated chromatogram.

3.3 Results and discussion

3.3.1 Detector and column saturation

Over-concentration may lead to column and detector saturation, consequently, to peak picking errors. An example of the possible issues is presented in Figure 4. The plot illustrates the effects of the detector and column saturation on the peak shape in the presence of a highly concentrated 2-phenylethyl acetate and methyl anthranilate respectively. In the case of detector saturation (Figure 4a), the ion channel at 104 (also noted as m/z = 104) shows a clear flat region close to the apex point. The presence of this flat area has two immediate consequences. First, the apex of this m/z is not in line with the ones of other unsaturated ion signals associated with the same neutral compound. Secondly, the presence of more than one local maximum on the ionic signal results in the "splitting" of the peak into two sub-peaks. Sub peak 1 contains the signal at m/z = 104 only. Due to the absence of the higher intensity peak from the compound spectrum, MS similarity-based annotation on this peak will most likely be incorrect. Sub-peak 2, instead, contains the signal from all aligned ion channels, including m/z = 104. The annotation here will be correct, but, since a part of the signal from m/z = 104 is assigned to sub-peak 1, if this ion is be used for the quantification, the peak area will be underestimated.

The effects of column saturation are presented in Figure 4c. In this case, the signal intensity is below the detector saturation limit (10^6), but the signals of all the ion channels are flattened at the top with noise from detector oscillation. Under these circumstances, the apex alignment result is based on random electronic noise and the annotation and quantification error is unpredictable.

The conventional solution implemented in commercial software to cope with saturation effects is to apply signal smoothing. Figure 4b shows the smoothed

signal of the saturated ionic channels (Figure 4a). After smoothing, the situation is apparently improved, as the apexes of all channels are now aligned. Nonetheless, a generalized smoothing of the ionic signals also presents severe drawbacks. The size of the smoothing window must be well adjusted to counterbalance the effects of saturation, but, at the same time, as small as possible. In our example, the required minimal smoothing window size is 7 data points, which is nearly one-fifth of the peak width. In the case of column saturation, because of the larger flatten region, a wider smoothing window is required. As an example, to align the signals from all the channels in Figure 4c, the size of a moving average window has to be larger than 19 points (Figure 4d) which is half of the peak width. Such a severe smoothing will result in the loss of small peaks, which can be detrimental for the coverage of the overall analysis. To counterbalance this effect, it is necessary to treat the signal of major and minor compounds differently.

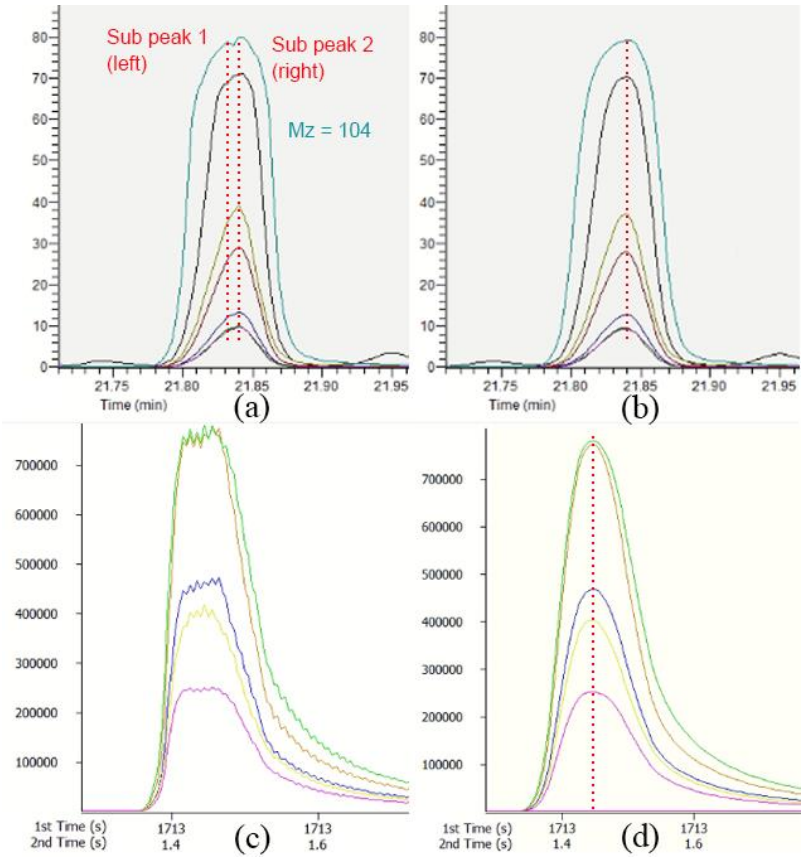


Figure 4. Chromatograms of 2-phenylethyl acetate by Xcalibur, Thermo Scientific™: (a) detector saturation; (b) Gaussian smoothed signal of (a). The

applied smoothing window was 7 points; Chromatograms of methyl anthranilate by ChromaToF, LECO: (c) column saturation, (d) Moving averaged signal of (c). The applied smoothing window was 19 points.

3.3.2 Targeted data processing

In a metabolomics study, the nature of the samples is clearly defined and the major compounds that have the potential to lead to saturation are almost invariably known. Their retention time (indices) and MS information can then be collected and stored in a matrix-specific library. This library will have a crucial role in the proposed two-stage approach. Column and detector saturations are easily determined by looking at the peak shape at different ion channels. Detector saturation can be detected by applying signal intensity thresholding. The value was 1800000 in our study. Column saturation, in reality, is difficult to be determined for every chromatogram. It's not possible to know whether the peak deformation is due to saturation or peaks overlap. It's necessary to know in advance which major compounds are likely to cause saturation, either from past experience or by analyzing the standards at the expected concentration.

Figure 5 illustrates the scheme of the proposed workflow. The workflow starts from the raw GC×GC data in netCDF format. These "raw" data are exported from ChromaTOF after automatic m/z alignment. The proposed workflow starts after the removal of background noise when the regions corresponding to "true" signals are detected in the 2D chromatographic plane (step 2). These regions of "interests" are then matched with the reference library to annotate the targeted compounds (steps 3, 4). Due to the presence of potential co-elution within the annotated signal region, it is necessary to detect how many peaks are actually present within each region of interest. (step 5). For each targeted compound, an unsaturated m/z that is exclusively present in the reference library is selected for peak quantification (steps 6, 7, 8). The signal of the

annotated peak is then subtracted from the chromatogram (step 9). The rest of the chromatogram still contains the signal of unannotated peaks. It will be analyzed by a downstream, untargeted data processing approach (step 10). The steps from 2 to 9 are performed by a set of bioinformatic scripts developed in R. The steps are presented in more detail below.

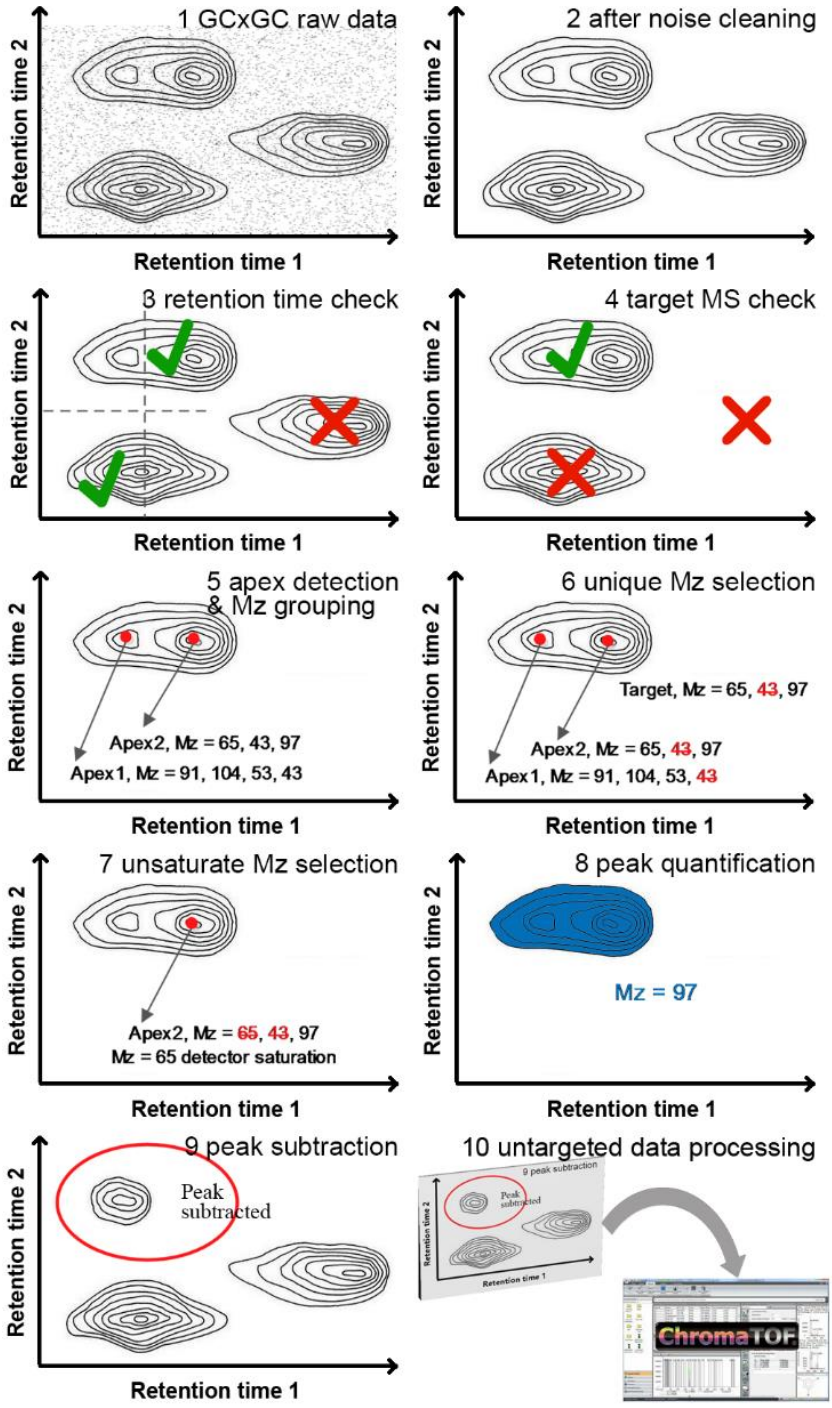


Figure 5. Target guided data processing flowchart

To construct the retention time and MS library for targeted compounds, pure standards have to be analyzed at the level just below saturation to obtain as much signal as possible. After the data processing, normalized MS of identified peaks can be added to the ChromaTOF library. Retention time is manually added on the compound's label. The ChromaTOF library can then be converted to a text file (SDF format) by lib2nist (NIST MS Search). This text file is ready to be processed by R.

Before starting the target data processing, the GC×GC-MS data must be exported from the GC×GC instrument control software (step 1). In this study, GC×GC data were acquired by ChromaTOF and then exported to netCDF format. During the export process, the MS data were centroid to integer mass-to-charge ratios to compensate for potential mass shifts.

In the second step, noise estimation and signal region detection were performed on the total ionic chromatogram (TIC). The background noise of GC×GC data was estimated by a local minimum method.[18] This method is rather robust, since it does not assume any statistical property of the background. As usual, it is necessary to set signal-to-noise thresholds to differentiate the signal coming from the peak and the one resulting from the noise. In our experiment, if the signal is higher than 1.7 times the local maximum electronic noise, it is considered as the true signal. The regions containing true signals are then labeled as peak regions. It is important to point out that due to co-elution, a detected peak region may contain more than one peak. However, this does not affect the annotation step.

After peak region detection, the position (1st and 2nd GC dimension) of the peak regions is known. Based on the retention times (indices) stored in the library, the potential peak areas containing each targeted compound can be identified (step 3). Retention time (RT) tolerance depends on the shifts of experimental chromatograms. Our tolerance window is a rectangle on the GC×GC plane. The edge length is 30 seconds in the 1st dimension and 2 seconds in the 2nd dimension of the chromatogram, but these tolerances have to be matched with the characteristics of the analytical method. As long as a peak region partially overlaps the tolerance window, the region of interest is “flagged” as a potential peak region.

Each potential peak region of the targeted compounds was then annotated according to the MS stored in the library (step 4). Annotation criteria will be explained in the section of Annotation criteria.

At the end of step 4, a region will be flagged as “annotated” if it contains at least one peak of the matching compounds. As already mentioned, co-elution is always possible. To check the purity of the annotated peak regions, a 2D apex detection algorithm was applied. The algorithm looks for local maxima on each m/z (step 5).[10] The size of the sliding window should be large enough to ignore the effect of detector oscillation. We used 3 times modulation time in 1st dimension and 21 times data acquisition time in 2nd dimension. When severe column saturation occurs, signal smoothing could also be applied to the second chromatographic dimension. Whether to use a smoothing process or not should be assessed by pre-analyzing the standards at the expected concentration. Our study applied moving average smoothing with 101 data points and 201 data points windows on the 2nd dimension for the 200 mg/L and 2000 mg/L standard solutions, respectively. Since false apexes may result from detector oscillations in each m/z, it is possible that a false peak be detected by the alignment of false apexes if the alignment involves hundreds of ion channels. However, this

phenomenon should only affect a small number of ion channels. In our experiment, if a location had aligned apexes from more than 15 ion channels, it was identified as the apex position of a true peak. The number of matching channels must be larger than the m/z number of aligned false apexes, usually 2 to 3, and smaller than the number of m/z collected in the MS library of targeted compounds. At this point, if an annotated peak region contains only one true peak, it is flagged as a pure peak region. Otherwise, it is a region of mixed peaks. For a pure peak region, the unique m/z selection (step 6) is skipped. Our interpretation continues with the region of mixed peaks.

For a region consisting of more than one peak, it is necessary to identify unique ion channels of the annotated compound (step 6) in order to proceed to quantification and signal subtraction. In the previous step, the apexes were detected at each m/z. After the apex alignment, for each “true” peak, the aligned ion channels are known. Based on the MS of the target compound recorded in the library, m/z belonging to the target compound are checked one by one. If an m/z appears on only one peak, it can be judged as the unique m/z for the target compound. If unique m/z exists, the approach can be continued, even for isomers and isobaric compounds. If a co-eluting compound shares the same ion channels with targeted compounds, it is not possible to obtain a unique m/z. In this case, the signal in this peak region will be untouched by the targeted approach and sent downstream to the untargeted workflow. It is worth noting that if two co-eluting compounds are not distinguished on the MS and the peak shape is disrupted by the saturation of the detector or column, distinguishing these peaks is not possible.

The intensities of the unique ion channels are collected at the peak apex and are flagged for saturation. An m/z is considered saturated if its intensity is higher than the manually defined detector saturation level (step 7).

The most intense unsaturated unique m/z is then used for peak quantification (step 8). Peak quantification is achieved by summing the intensity values over the entire detection peak region for the selected m/z.[19]

After peak quantification, the MS signal of the targeted compound is subtracted from the chromatogram (step 9). The relative intensity of each m/z is available for targeted compounds. According to the intensity profile of unique m/z, the profiles of other ion channels can be estimated. Then, in the entire peak region, the estimated signal of the targeted compound is subtracted from the experimental chromatogram. The signal subtraction only occurs on the ion channels belonging to the targeted compound. The leftover signal comes from the untargeted compound(s). For saturated ion channels, the estimated intensities are greater than the experimental intensities recorded in the chromatogram. The subtracted residual is below zero and is replaced by zero.

The signal remaining in the chromatogram will be reloaded to ChromaTOF for untargeted processing (step 10).

3.3.3 Annotation criteria

Four criteria must be satisfied for a successful peak annotation: 1) the difference in retention time (indices) (RTI) should not exceed an experiment-specific threshold, 2) the similarity between the spectra should be high, 3) the ion channels with the highest intensity in the database should always be present in the sample, 4) the sequence of the ion channels with the highest intensity should be preserved in the sample. Retention time (indices) and MS similarity are widely used in untargeted approaches and well-known to every chromatographer. In routine analysis, the acceptable tolerance in the retention time dimension for each experimental setting is known and can be assessed by

checking the retention times obtained from past analyses. RTI, instead, can be extracted from MS databases. It is important to notice that the deviation of the indices value present in the public database is in general much larger than the deviation of the experimental retention time measured on a specific instrument. With RTI, more peak regions may be selected for further MS annotation. This will increase the uncertainty on the final result.

MS similarity is calculated by composite optimized dot product algorithm.[20] This algorithm uses small computation power and does not require programmatic access to the large MS database. The calculated MS similarity value is between 0 and 1000. Since a large peak is less likely to be covered by small peaks, we believe that the peak of the targeted compound is at least partially separated from co-elutes. The minimal MS similarity threshold used in our study was 995.This threshold can, however, be manually tuned by the user. As usual, a more liberal threshold will improve annotation efficacy at the price of potential false results.

In our approach, we decided to add two additional rules which take into account the relative intensity, the presence and the order of the most intense ion channels to these two well-established criteria. These rules increase the confidence in annotation when saturation and co-elution occur. To be comparable with common ChromaTOF procedures, the examined m/z number was 5 in our study.[17]

The first criterion relies on the relative intensity of the most intense ion channels in the library MS in the sample. Since the targeted peak is the major peak, its intense ions must be the most intense in the MS by at least some pixels in the GC×GC peak region. This criterion is robust to the interference caused by detector and column saturations. The second criterion requires that not only the

intense ion channels are present, but also that they are showing the same rank they have in the library spectrum. The rank maybe altered by the signal of co-eluted compounds. However, our targeted approach focuses on the major metabolites in the sample. It is unlikely that the MS of a major compound will be largely altered.

In addition, in terms of similarity, presence, and sequence of the mass spectrum, the comparison is based on the single MS. As long as one data point (or one pixel) in the peak region passes the check, positive feedback is returned for the whole region. One of the advantages of this annotation is the use of all mass spectra of a single peak region. It avoids the apex detection, ion channel grouping, deconvolution, and mass spectrum construction for the apex point. Therefore, the annotation results are not affected by errors in these processes.

3.3.4 Improve the annotation and quantification for targeted peaks

The first benchmark of the proposed pipeline was performed on a dilution series of standard compounds (2, 20, 200, 2000 mg/L). The results of the analysis were compared with the ones obtained by a full untargeted approach. It is well known that choosing a suitable peak width is important for an untargeted approach, so the untargeted pipeline was run with both a standard 0.8s and an optimal peak width. The optimal peak width can be estimated by looking at the chromatogram manually. As expected, when the column was saturated with 200 and 2000 mg/L standard solution, peaks were split into a few sub peaks by deconvolution (Figure 6). The number of subpeaks detected by the different processing pipelines for the different compounds injected at different concentrations is shown in Figure 6. Severe over-deconvolution may lead to incorrect quantification. In our study, with 2000 mg/L standard, the peak of phenethyl acetate is overspilt into several sub-peaks, m/z = 48 was chosen for

the quantification. This m/z does not exist in the library-recorded-MS of phenethyl acetate.

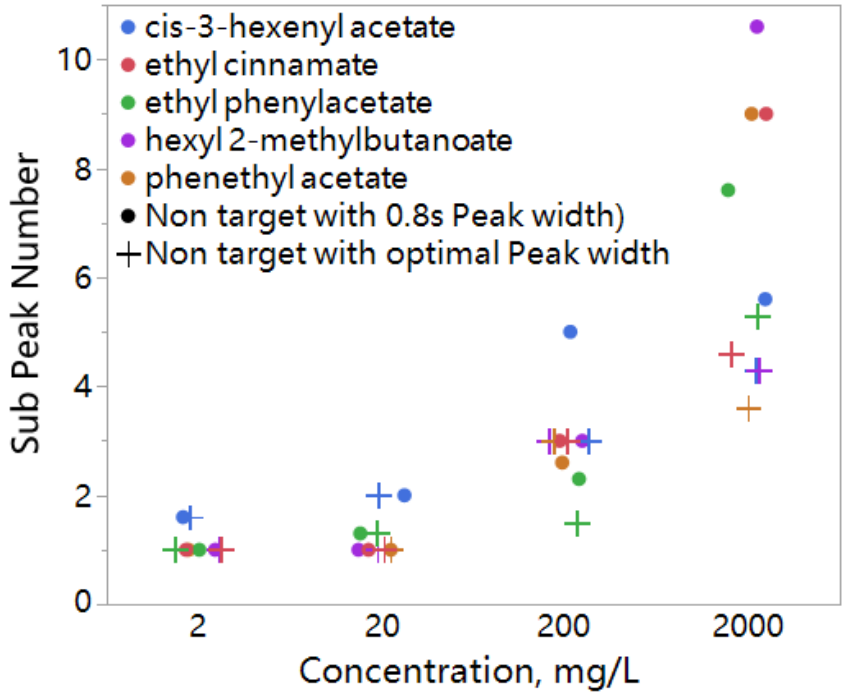


Figure 6. Compare the sub peak number for untargeted approaches with fixed and optimal peak width

When a suitable peak width is applied, over-deconvolution is significantly reduced, as shown by the low hanging cluster of crosses in Figure 6. The peak list reported by ChromaTOF usually consists of one major peak and several sub peaks for each compound. All peaks are quantified using ion channels, which are easily found in the library MS. A weighted linear regression (by GraphPad Prism) was applied to examine the linear relationship between peak area and concentration of the detected major peaks, and the results are shown in Figure 7. With the proposed approach, linearity (measured in terms of R^2) is generally improved, with the only exception of hexyl 2-methylbutanoate, where both targeted approach and peak width optimized untargeted approach provide high linearity. This means that the samples can be quantified more accurately in the concentration range tested. It is worth noting that in this case we are using pure

standard solutions. It is not difficult to pick up the correct peaks manually. However, for real samples, manual peak picking is still tricky and subjective.

Table 1. Compare peak area RSD% among untargeted data processing with 0.8s peak width (p.w.), 3s p.w., and target data processing for 2000 mg/L standard solution

Compounds	Untargeted 0.8s p.w.	Untargeted optimal p.w.	Targeted
cis-3-hexenyl acetate hexyl	3.0	23.2	0.9
2-methylbutanoate	54.3	1.7	1.5
phenethyl acetate	n.a.*	3.9	2.3
ethyl phenylacetate	38.5	22.7	3.9
ethyl cinnamate	26.2	2.3	1.5

* When column saturation occurs. For phenethyl acetate, peak quantification is not possible when untargeted approach and 0.8s peak width applied.

In addition, with the proposed targeted approach, peak splitting does not exist. Because the entire signal region is used to calculate the peak area, differences in quantification results only arise from background noise cleaning. This also improves the analytical reproducibility and the relative standard deviation across multiple injections (Table 1). Besides, the proposed targeted approach does not require any manual supervision. It reduces labor cost and human error during the analysis of a complex sample.

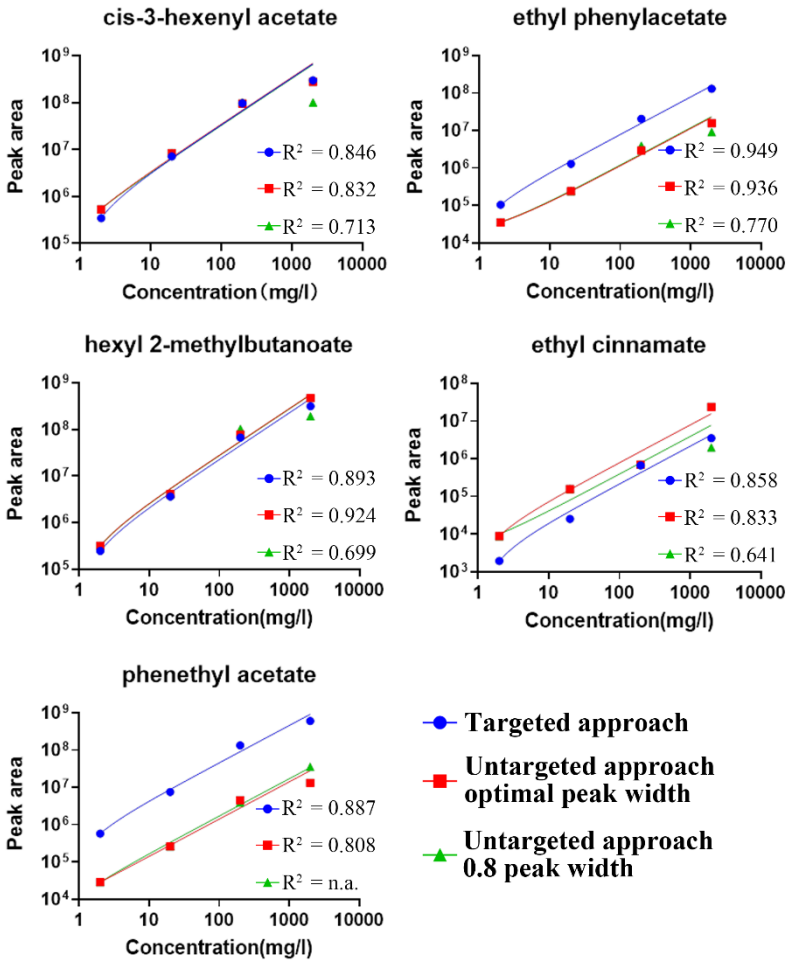


Figure 7. Compare the weighted linear regression for standards quantified with the proposed targeted approach, and untargeted approaches

3.3.5 Improve the peak annotation for untargeted co-eluting peaks

Targeted peak signals were subtracted from the original chromatogram according to the peak profile estimated with the experimental profile of unique m/z. The subtraction result is displayed in Figure 8 for the solution of standards (top panels) and for the wine sample. It can be seen clearly that peak signals are mostly removed in both standards and pooled wine chromatograms. There

are, however, some small artifacts which arise from two phenomena. Firstly, the intensity of the MS of the targeted compound in the library is not recorded with sufficient accuracy by ChromaTOF. Relative intensities for each m/z are, indeed, recorded as integers. This lack of accuracy propagates through the estimation of the signal that has to be subtracted from the raw data, leading to a mismatch between the estimated and the true profile of each peak. The other source of incomplete signal subtraction is the procedure applied during peak region detection. For convenience, the peak region was detected based on the TIC. By summing up the signal of each m/z, the electronic error is also enhanced. Detecting the peak region on every m/z then merging regions will provide more accurate peak region detection.

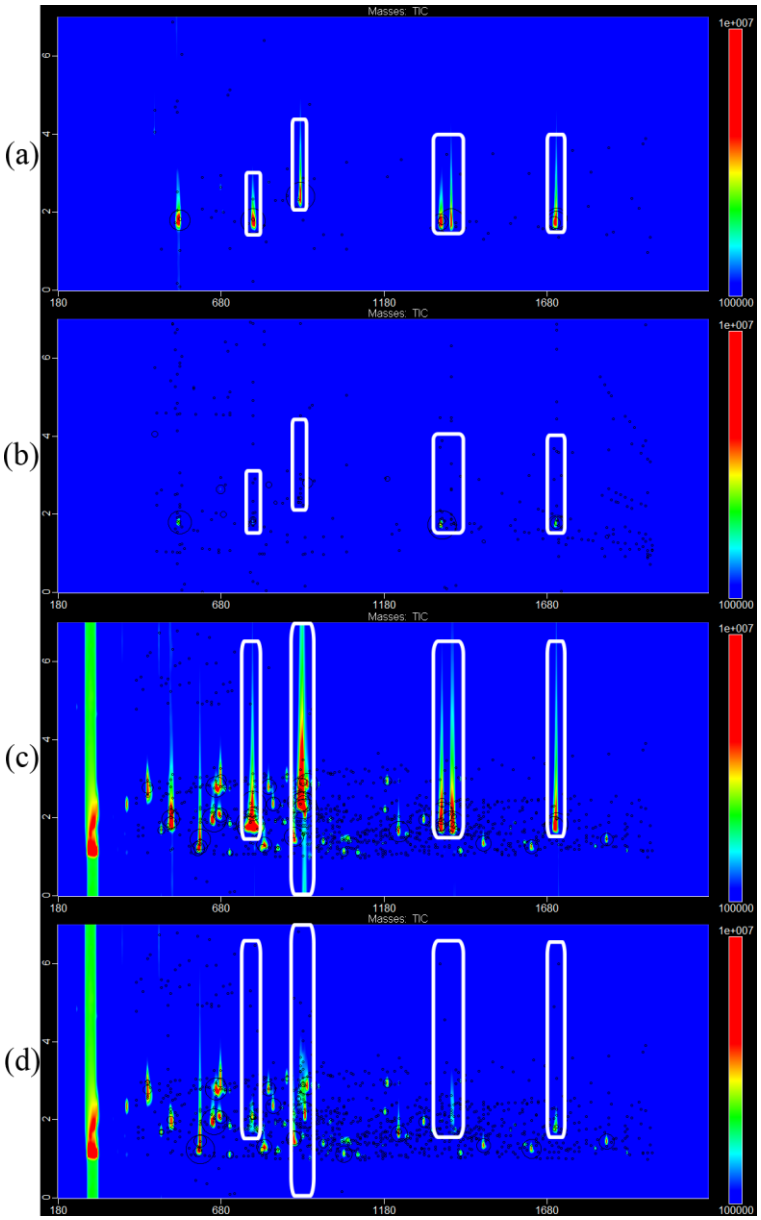


Figure 8. Peak signal (marked by white label) subtraction results. (a) chromatogram of 200 mg/L standard solution; (b) peak signal subtracted of (a); (c) chromatogram of 10 times diluted mixed white wine plus 10 μL of 200 mg/L standard solution; (d) peak signal subtracted of (c)

Predictably, artifacts resulting from incomplete signal subtraction may affect the peak picking of co-eluted compounds. To assess this effect, chromatograms of unsaturated, saturated, and subtracted saturated peaks were analyzed by the untargeted approach of ChromaTOF. The peak-picking performances of the

peaks co-eluted with saturated peaks were compared based on the MS similarity. The similarity results are listed in Table 3, mass spectra can be found in the supplementary, Figure S20 to S28. Under good separation conditions and in the absence of saturation phenomena, the constructed mass spectra of all peaks were highly similar to the mass spectra stored in the public database (similarity over 850). Subtraction of the signals of the target peaks from the saturated chromatograms did not eliminate the interference of the co-eluting compounds. The MS similarity of 6-methyl-5-hepten-2-one and linalool ether decreased to about 750. This may be attributed to the artifacts brought about by subtraction. However, this result also suggests that our proposed saturation subtraction approach has the potential to reduce the interference of large saturated peaks with the co-eluting peaks during peak picking. The mass spectra constructed after signal subtraction showed an improvement by almost 100 units in the similarity between 6-methyl-5-hepten-2-one and linalool ether, compared to the mass spectra constructed when the saturated peaks were co-eluting. This improvement is significant because the mass spectrum similarity threshold used for annotation in most recommended untargeted approaches is 700. In the case of benzene, 1-vinyl-4-ethyl, the saturated standard peaks did not interfere with the MS construction. This indicates that the algorithm developed by Leco successfully deconvoluted the signal.

Table 3. Compared MS construction results (similarity) among unsaturated, saturated, and saturation subtracted chromatogram

Compounds	Unsaturated	Saturated	Saturation subtracted
6-methyl-5-hepten-2-one	859	665	744
benzene, 1-ethenyl-4-ethyl	864	892	875
linalool ethyl ether	874	610	762

3.4 Conclusion

Metabolite concentrations differ from one another by orders of magnitude. This poses difficulties during metabolite profiling. Errors can occur during peak-picking and MS construction when the detector and/or column are saturated with major metabolites. In metabolomics studies, the nature of the sample is well defined and the major compounds that have the potential to cause saturation are almost invariably known. Data processing results can be improved by a two-stage data processing strategy that will incorporate a targeted data processing and cleaning approach upstream of the “standard” untargeted analysis. Our experiments show a significant improvement in annotation and quantification results for targeted compounds causing instrumental saturation. After subtracting the saturate signal of targeted compounds, the MS construction was improved for co-eluted compounds. Incomplete signal subtraction may occur. It leads to the detection of false positive peaks or to interferences with the construction of mass spectra of co-diluted peaks. High-resolution MS libraries and more accurate peak area detection methods should be tested for further improvement.

References

(1) Bean, H. D.; Rees, C. A.; Hill, J. E. Comparative Analysis of the Volatile Metabolomes of Pseudomonas Aeruginosa Clinical Isolates. *J. Breath Res.* **2016**, *10* (4), 047102. https://doi.org/10.1088/1752-7155/10/4/047102.

(2) Pinu, F. Grape and Wine Metabolomics to Develop New Insights Using Untargeted and Targeted Approaches. *Fermentation* **2018**, *4* (4), 92. https://doi.org/10.3390/fermentation4040092.

(3) Goldansaz, S. A.; Guo, A. C.; Sajed, T.; Steele, M. A.; Plastow, G. S.; Wishart, D. S. Livestock Metabolomics and the Livestock Metabolome: A

Systematic Review. *PLOS ONE* **2017**, *12* (5), e0177675. https://doi.org/10.1371/journal.pone.0177675.

(4) Vemuri, R.; Gundamaraju, R.; Shastri, M. D.; Shukla, S. D.; Kalpurath, K.; Ball, M.; Tristram, S.; Shankar, E. M.; Ahuja, K.; Eri, R. Gut Microbial Changes, Interactions, and Their Implications on Human Lifecycle: An Ageing Perspective. *BioMed Res. Int.* **2018**, *2018*, 1–13. https://doi.org/10.1155/2018/4178607.

(5) Qiu, Y.; Ree, D. Gas Chromatography in Metabolomics Study. In *Advances in Gas Chromatography*; Guo, X., Ed.; InTech, 2014. https://doi.org/10.5772/57397.

(6) Tsugawa, H.; Tsujimoto, Y.; Arita, M.; Bamba, T.; Fukusaki, E. GC/MS Based Metabolomics: Development of a Data Mining System for Metabolite Identification by Using Soft Independent Modeling of Class Analogy (SIMCA). *BMC Bioinformatics* **2011**, *12* (1), 131. https://doi.org/10.1186/1471-2105-12-131.

(7) Ong, R. C. Y.; Marriott, P. J. A Review of Basic Concepts in Comprehensive Two-Dimensional Gas Chromatography. *J. Chromatogr. Sci.* **2002**, *40* (5), 276–291. https://doi.org/10.1093/chromsci/40.5.276.

(8) Welthagen, W.; Shellie, R. A.; Spranger, J.; Ristow, M.; Zimmermann, R.; Fiehn, O. Comprehensive Two-Dimensional Gas Chromatography–Time-of-Flight Mass Spectrometry (GC × GC-TOF) for High Resolution Metabolomics: Biomarker Discovery on Spleen Tissue Extracts of Obese NZO Compared to Lean C57BL/6 Mice. *Metabolomics* **2005**, *1* (1), 65–73. https://doi.org/10.1007/s11306-005-1108-2.

(9) Chesler, S. N.; Cram, S. P. Iterative Curve Fitting of Chromatographic Peaks. *Anal. Chem.* **1973**, *45* (8), 1354–1359. https://doi.org/10.1021/ac60330a031.

(10) Min Chen; Reichenbach, S. E.; Jiazheng Shi. Automated Unmixing of Comprehensive Two-Dimensional Chemical Separations with Mass Spectrometry. In *2005 IEEE International Conference on Electro Information*

Technology; IEEE: Lincoln, NE, USA, 2005; pp 1–6. https://doi.org/10.1109/EIT.2005.1627049.

(11) Zheng, Y.; Fan, R.; Qiu, C.; Liu, Z.; Tian, D. An Improved Algorithm for Peak Detection in Mass Spectra Based on Continuous Wavelet Transform. *Int. J. Mass Spectrom.* **2016**, *409*, 53–58. https://doi.org/10.1016/j.ijms.2016.09.020.

(12) Reichenbach, S. E.; Tao, Q.; Cordero, C.; Bicchi, C. A Data-Challenge Case Study of Analyte Detection and Identification with Comprehensive Two-Dimensional Gas Chromatography with Mass Spectrometry (GC×GC-MS). *Separations* **2019**, *6* (3), 38. https://doi.org/10.3390/separations6030038.

(13) Rowan, D. D. Volatile Metabolites. *Metabolites* **2011**, *1* (1), 41–63. https://doi.org/10.3390/metabo1010041.

(14) Wilde, M. J.; Zhao, B.; Cordell, R. L.; Ibrahim, W.; Singapuri, A.; Greening, N. J.; Brightling, C. E.; Siddiqui, S.; Monks, P. S.; Free, R. C. Automating and Extending Comprehensive Two-Dimensional Gas Chromatography Data Processing by Interfacing Open-Source and Commercial Software. *Anal. Chem.* **2020**, *92* (20), 13953–13960. https://doi.org/10.1021/acs.analchem.0c02844.

(15) Carlin, S.; Vrhovsek, U.; Franceschi, P.; Lotti, C.; Bontempo, L.; Camin, F.; Toubiana, D.; Zottele, F.; Toller, G.; Fait, A.; Mattivi, F. Regional Features of Northern Italian Sparkling Wines, Identified Using Solid-Phase Micro Extraction and Comprehensive Two-Dimensional Gas Chromatography Coupled with Time-of-Flight Mass Spectrometry. *Food Chem.* **2016**, *208*, 68–80. https://doi.org/10.1016/j.foodchem.2016.03.112.

(16) Stefanuto, P.-H.; Perrault, K. A.; Dubois, L. M.; L'Homme, B.; Allen, C.; Loughnane, C.; Ochiai, N.; Focant, J.-F. Advanced Method Optimization for Volatile Aroma Profiling of Beer Using Two-Dimensional Gas Chromatography Time-of-Flight Mass Spectrometry. *J. Chromatogr. A* **2017**, *1507*, 45–52. https://doi.org/10.1016/j.chroma.2017.05.064.

(17) Reichenbach, S. E.; Ni, M.; Zhang, D.; Ledford, E. B. Image Background Removal in Comprehensive Two-Dimensional Gas Chromatography. *J. Chromatogr. A* **2003**, *985* (1–2), 47–56. https://doi.org/10.1016/S0021-9673(02)01498-X.

(18) Reichenbach, S. E. Quantification in Comprehensive Two-Dimensional Liquid Chromatography. *Anal. Chem.* **2009**, *81* (12), 5099–5101. https://doi.org/10.1021/ac900047z.

(19) Stein, S. E.; Scott, D. R. Optimization and Testing of Mass Spectral Library Search Algorithms for Compound Identification. *J. Am. Soc. Mass Spectrom.* **1994**, *5* (9), 859–866. https://doi.org/10.1016/1044-0305(94)87009-8.

Supplementary

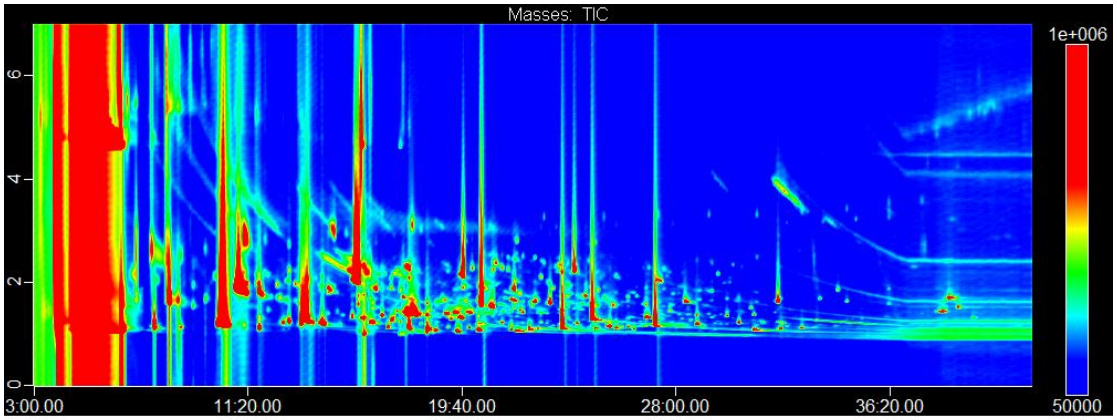


Figure S 6 0.5 ml pooled white wine is sampled by SPME and analyzed with VF-Wax column (100% polyethylene glycol; 30m × 0.25mm × 0.25µm) followed by Rxi-17Sil MS (1.50m × 0.15mm × 0.15µm)

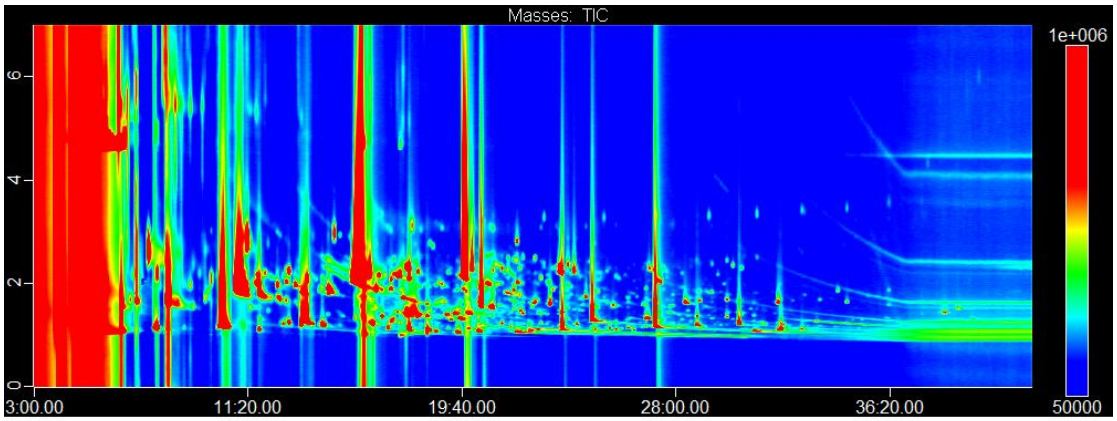


Figure S 7 1 ml pooled white wine is sampled by SPME and analyzed with VF-Wax column (100% polyethylene glycol; 30m × 0.25mm × 0.25µm) followed by Rxi-17Sil MS (1.50m × 0.15mm × 0.15µm)

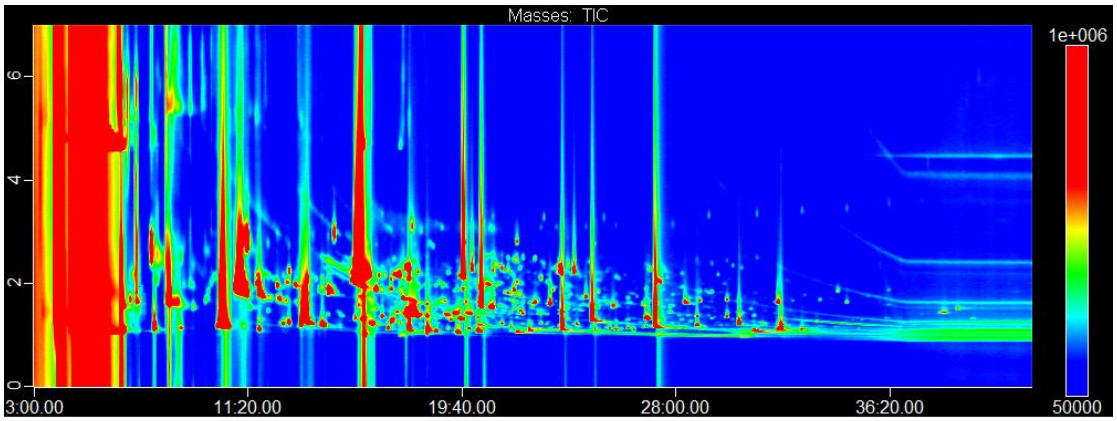


Figure S 8 2 ml pooled white wine is sampled by SPME and analyzed with VF-Wax column (100% polyethylene glycol; 30m × 0.25mm × 0.25µm) followed by Rxi-17Sil MS (1.50m × 0.15mm × 0.15µm)

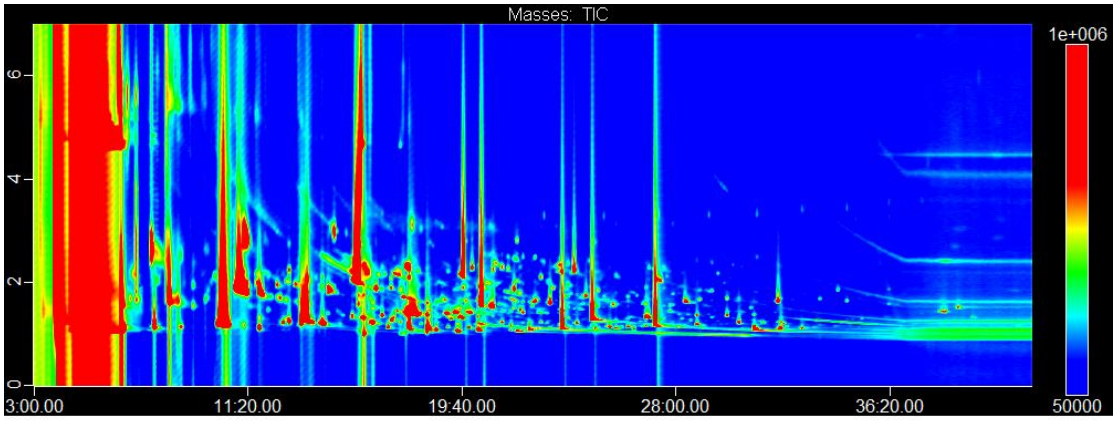


Figure S 9 4 ml pooled white wine is sampled by SPME and analyzed with VF-Wax column (100% polyethylene glycol; 30m × 0.25mm × 0.25µm) followed by Rxi-17Sil MS (1.50m × 0.15mm × 0.15µm)

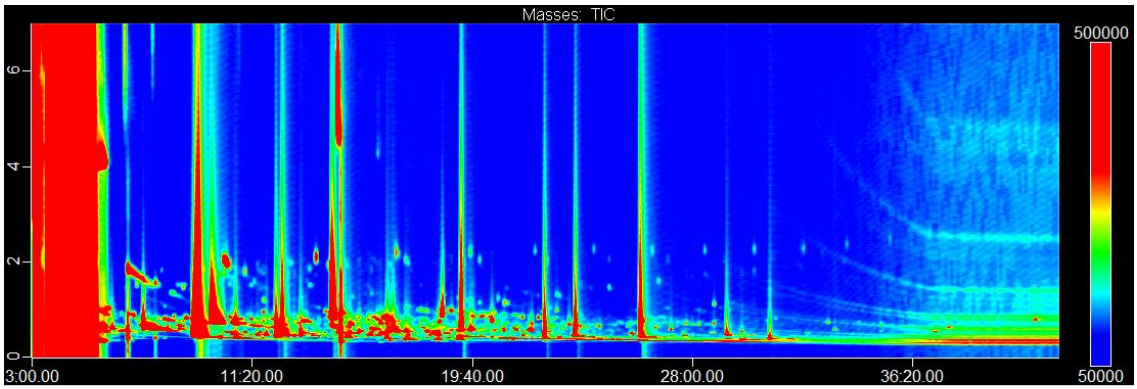


Figure S 10 0.5 ml pooled white wine is sampled by SPME and analyzed with VF-Wax column (100% polyethylene glycol; 30m × 0.25mm × 0.25µm) followed by DB-1701 (0.30m × 0.25mm × 0.25µm)

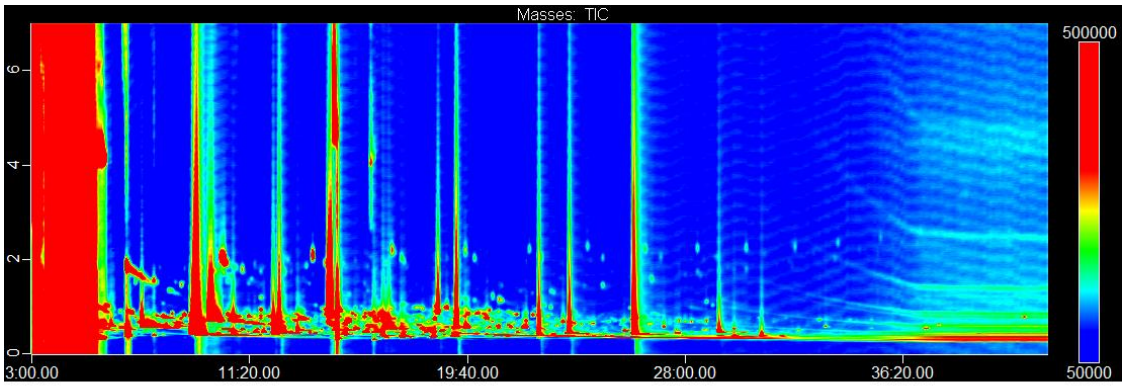


Figure S 11 1 ml pooled white wine is sampled by SPME and analyzed with VF-Wax column (100% polyethylene glycol; 30m × 0.25mm × 0.25µm) followed by DB-1701 (0.30m × 0.25mm × 0.25µm)

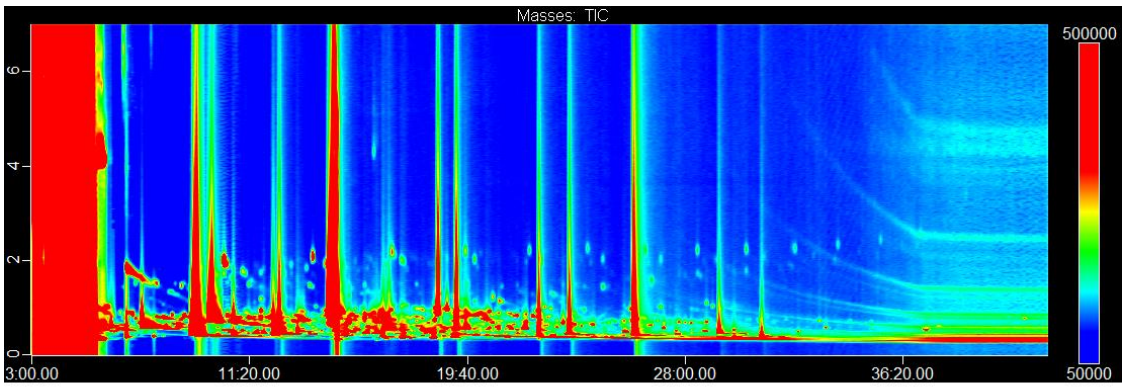


Figure S 12 2 ml pooled white wine is sampled by SPME and analyzed with VF-Wax column (100% polyethylene glycol; 30m × 0.25mm × 0.25µm) followed by DB-1701 (0.30m × 0.25mm × 0.25µm)

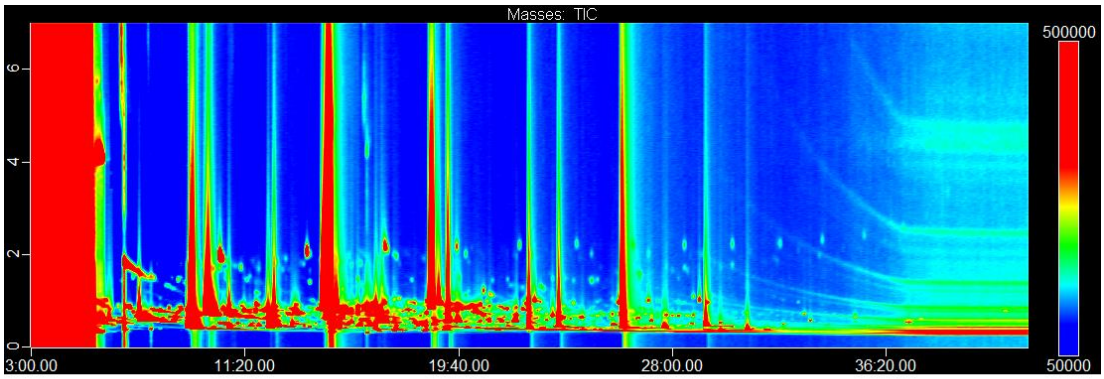


Figure S 13 4 ml pooled white wine is sampled by SPME and analyzed with VF-Wax column (100% polyethylene glycol; 30m × 0.25mm × 0.25µm) followed by DB-1701 (0.30m × 0.25mm × 0.25µm)

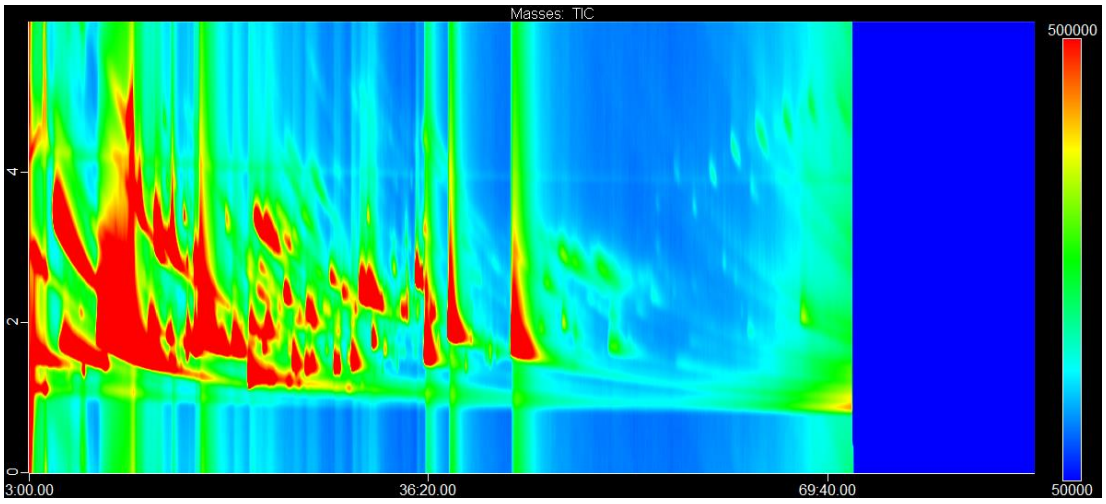


Figure S 14 cider VOC analysis with DHS

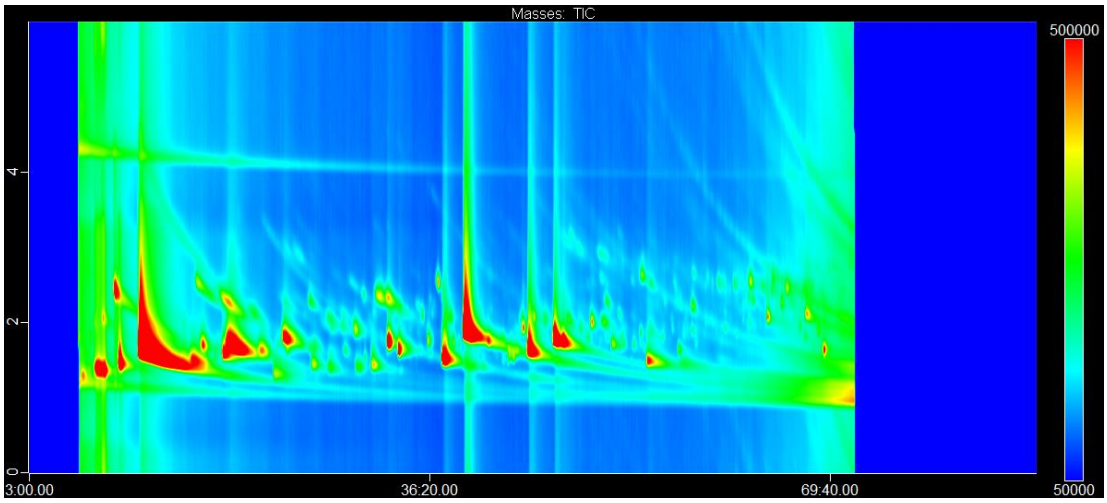


Figure S 15 cider VOC analysis with DLLME

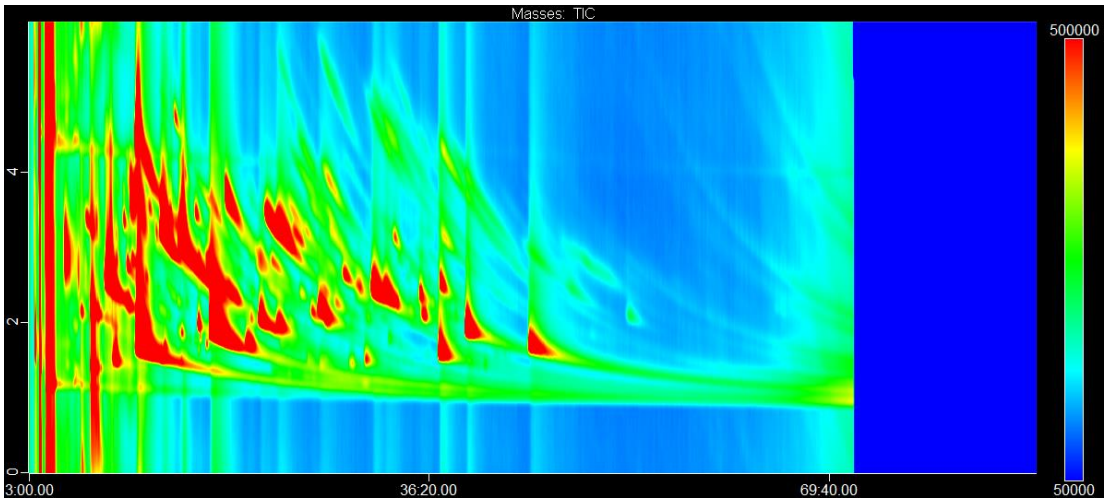


Figure S 16 cider VOC analysis with SPME

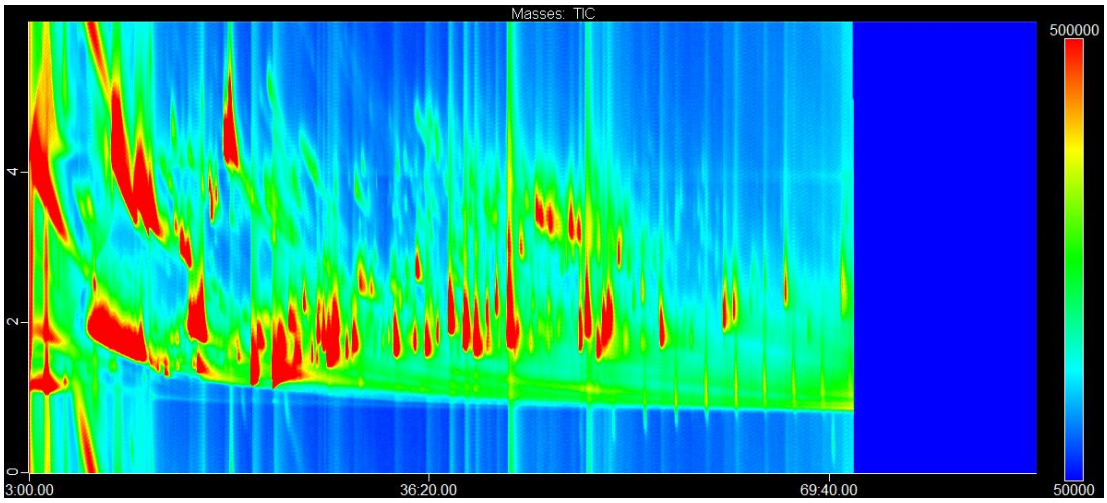


Figure S 17cider VOC analysis with mSBSE

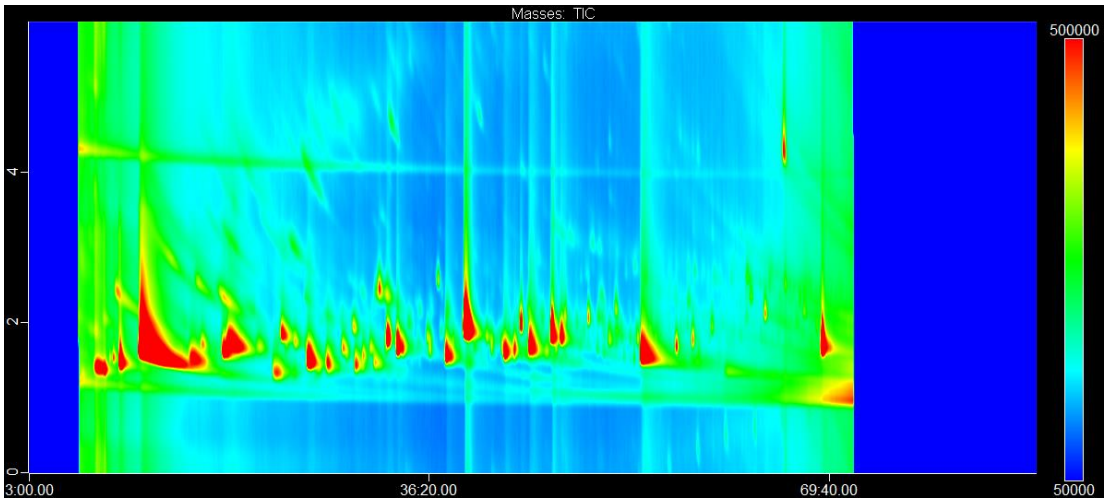


Figure S 18 cider VOC analysis with SPE

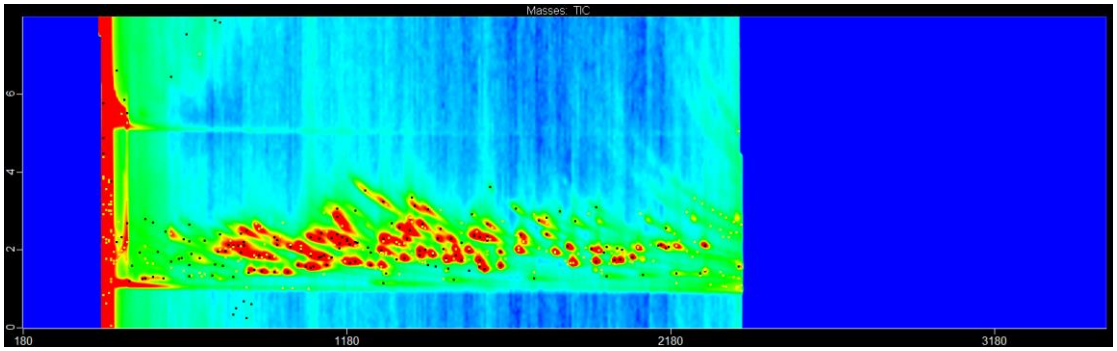


Figure S 19 common beverage VOC standards analyzed under GC×GC condition: column flow

0.8ml/min, temp program 5°C/min, 2nd oven temp offset 1°C, modulation temp offset 6°C, modulation time 8s, hot pulse time 2.8s

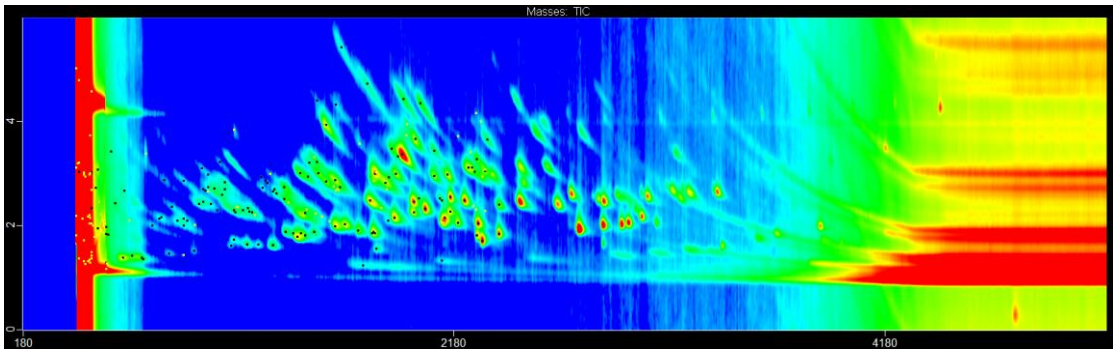


Figure S 20 common beverage VOC standards analyzed under GC×GC condition: column flow 0.8ml/min, temp program 3°C/min, 2nd oven temp offset 1°C, modulation temp offset 10°C, modulation time 6s, hot pulse time 2.1s

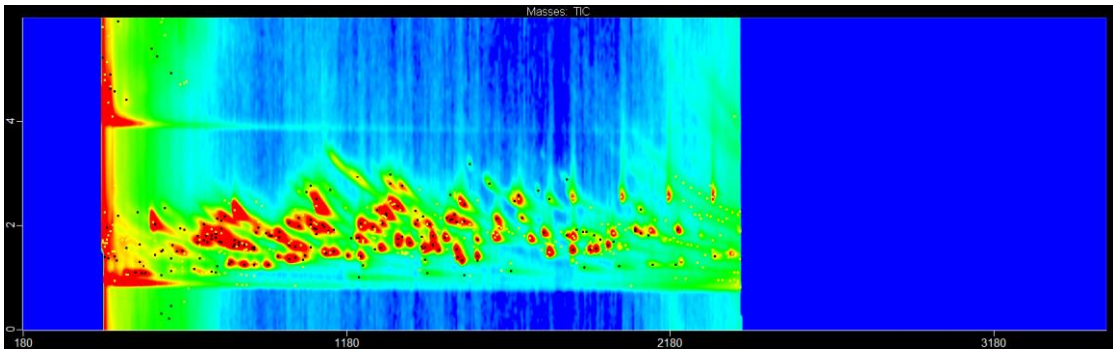


Figure S 21 common beverage VOC standards analyzed under GC×GC condition: column flow 1.2ml/min, temp program 5°C/min, 2nd oven temp offset 5°C, modulation temp offset 8°C, modulation time 6s, hot pulse time 0.9s

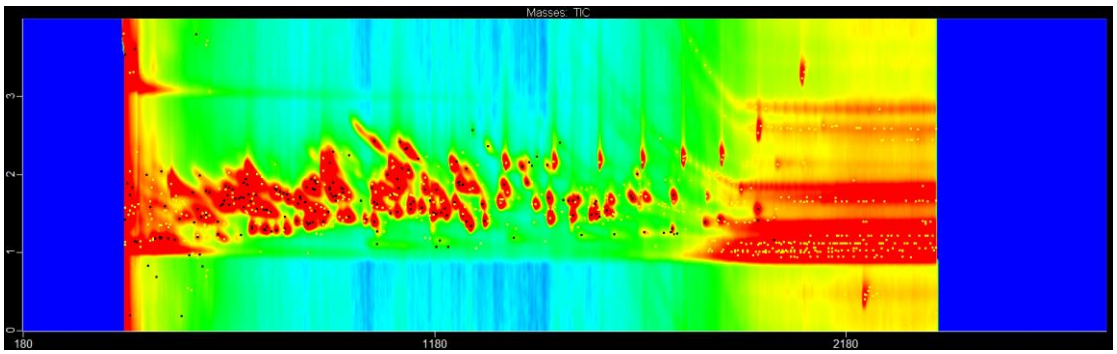


Figure S 22 common beverage VOC standards analyzed under GC×GC condition: column flow 0.8ml/min, temp program 7°C/min, 2nd oven temp offset 5°C, modulation temp offset 10°C, modulation time 4s, hot pulse time 1.4s

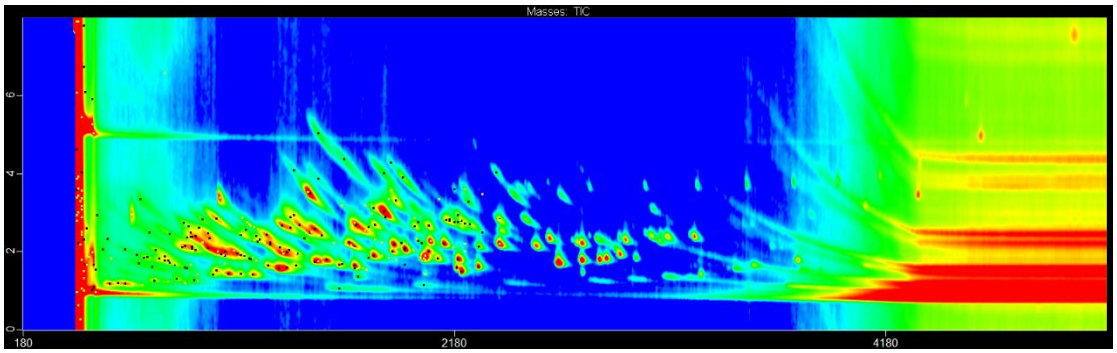


Figure S 23 common beverage VOC standards analyzed under GC×GC condition: column flow 1.2ml/min, temp program 3°C/min, 2nd oven temp offset 3°C, modulation temp offset 8°C, modulation time 8s, hot pulse time 2.8s

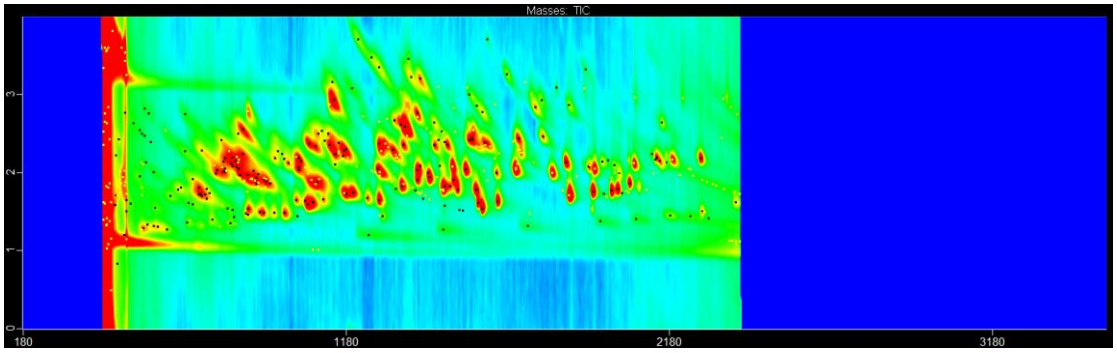


Figure S 24 common beverage VOC standards analyzed under GC×GC condition: column flow 0.8ml/min, temp program 5°C/min, 2nd oven temp offset 1°C, modulation temp offset 8°C, modulation time 4s, hot pulse time 0.6s

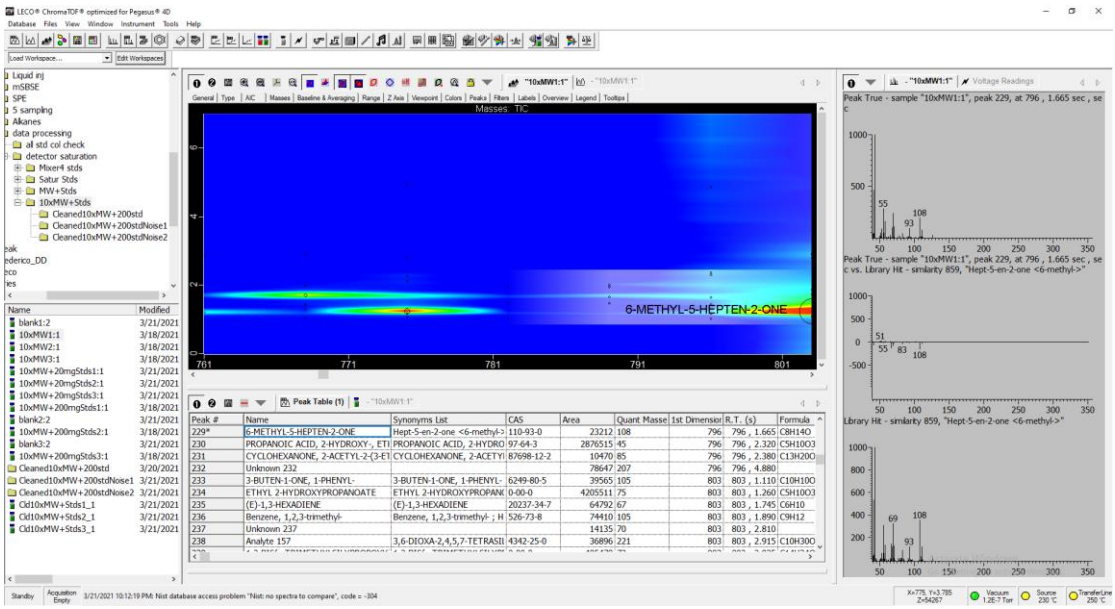


Figure S25. the mass spectrum of 6-methyl-5-hepten-2-one under unsaturated condition

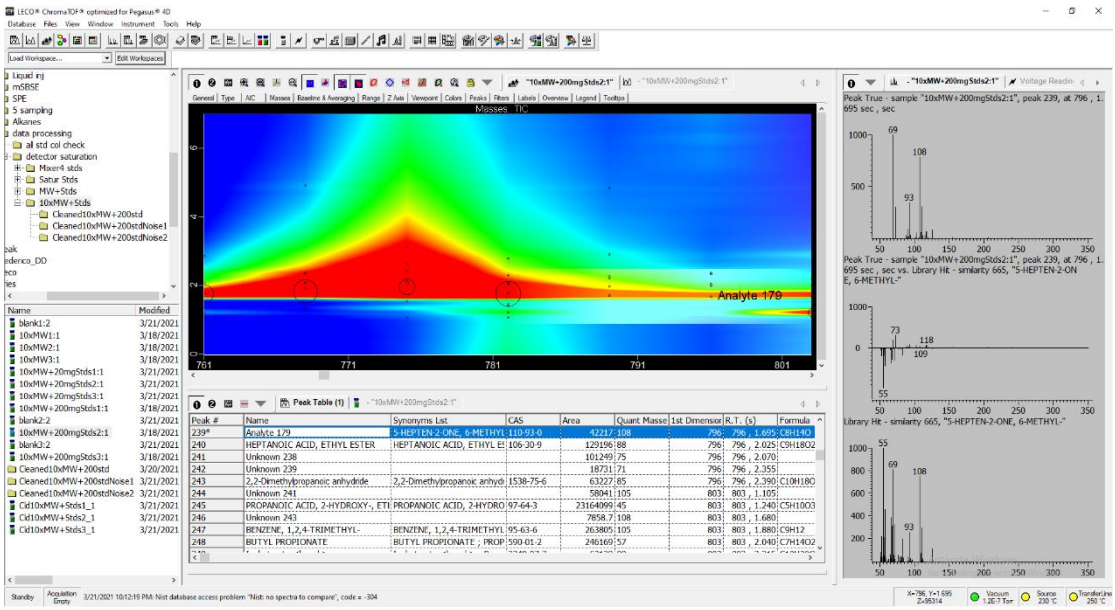


Figure S26. the mass spectrum of 6-methyl-5-hepten-2-one affected by saturation

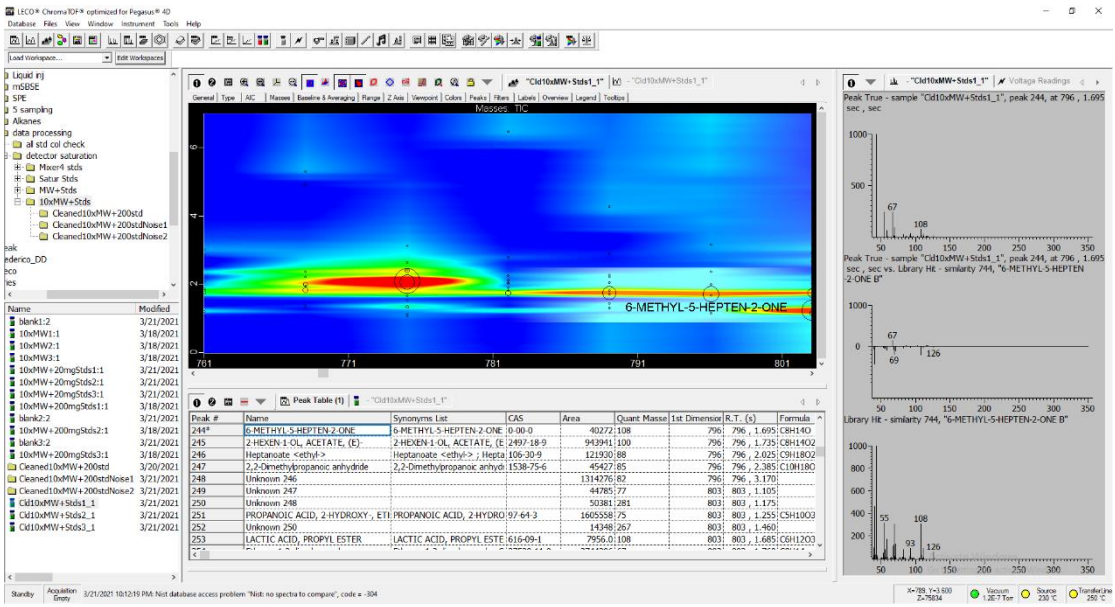


Figure S27. the mass spectrum of 6-methyl-5-hepten-2-one after saturation correction

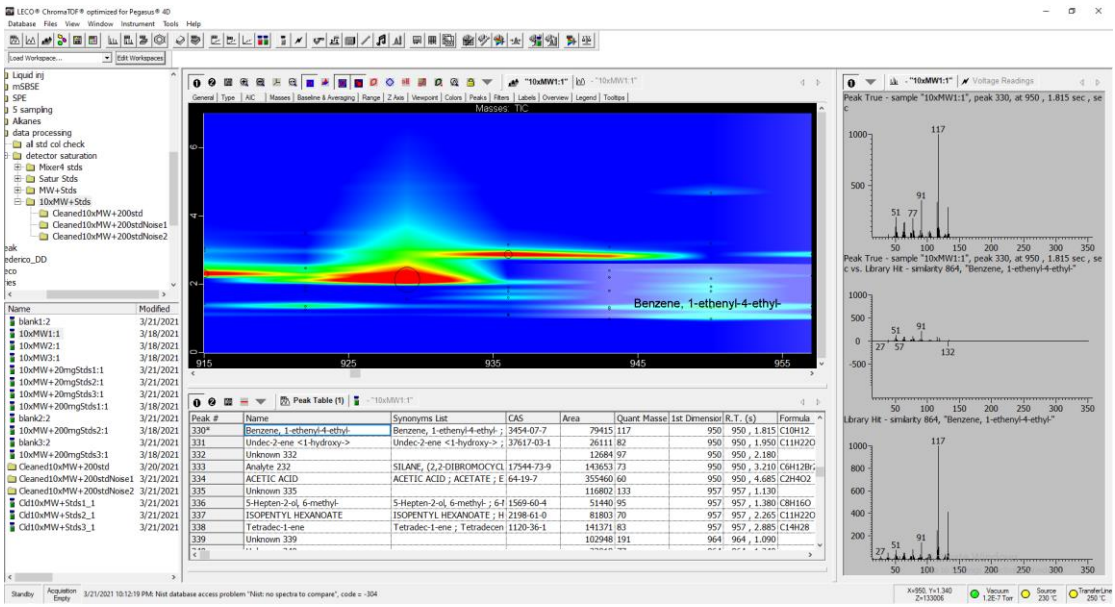


Figure S28. the mass spectrum of benzene, 1-ethenyl-4-ethyl- under unsaturated condition

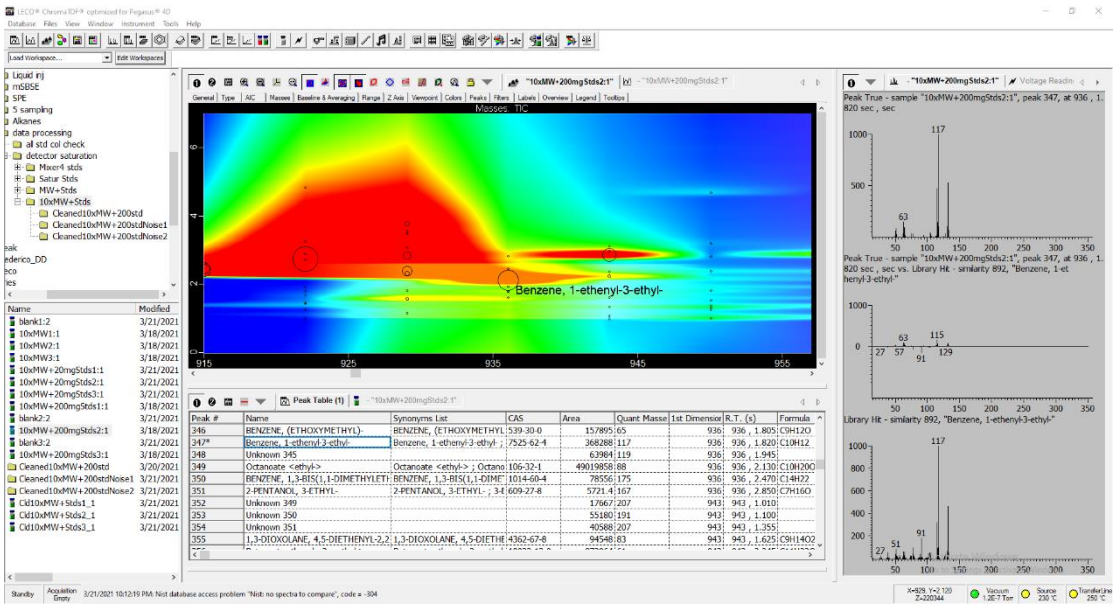


Figure S29. the mass spectrum of benzene, 1-ethenyl-4-ethyl- affected by saturation

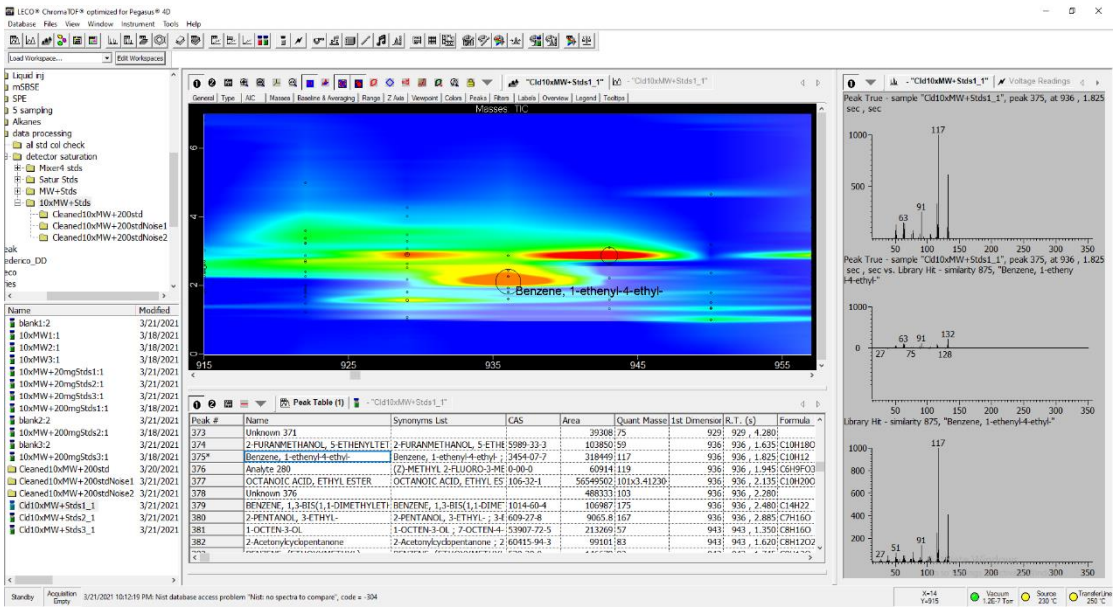


Figure S30. the mass spectrum of benzene, 1-ethenyl-4-ethyl- after saturation correction

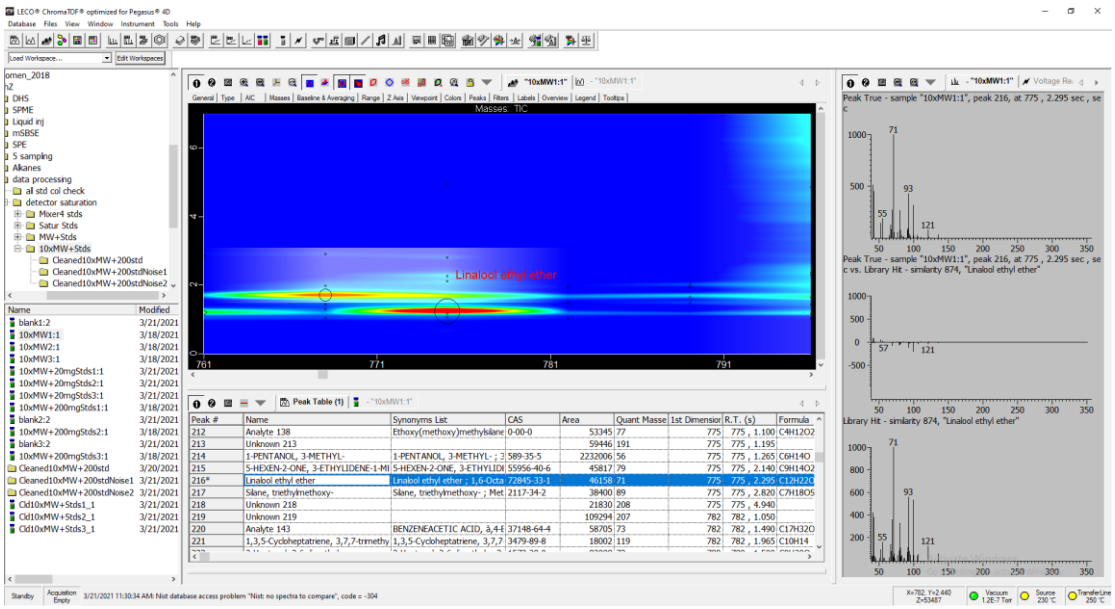


Figure S31. the mass spectrum of linalool ethyl ether under unsaturated condition

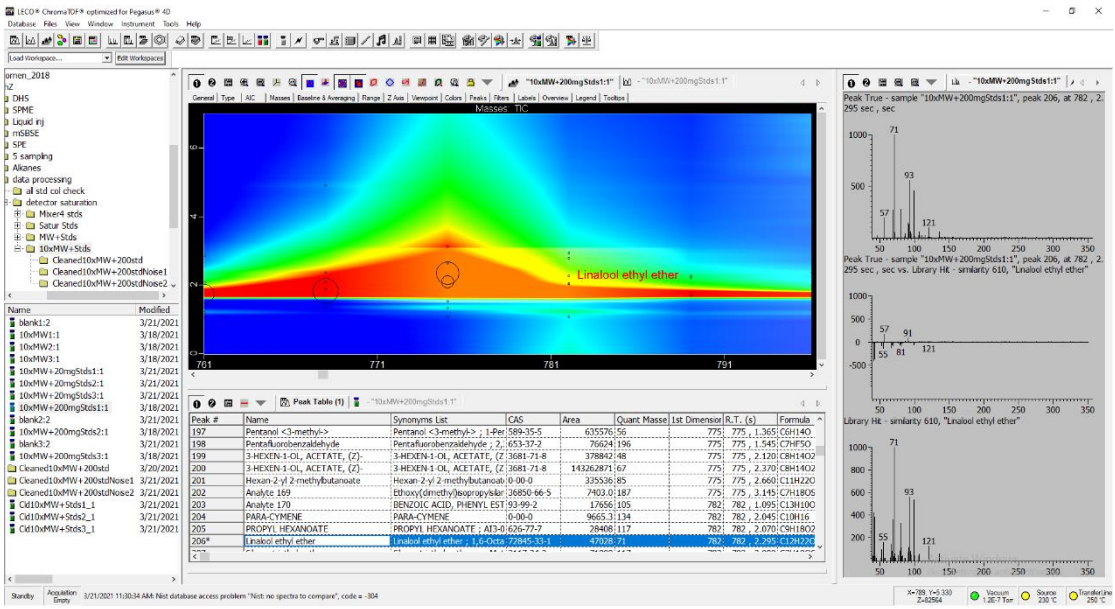


Figure S32. the mass spectrum of linalool ethyl ether affected by saturation

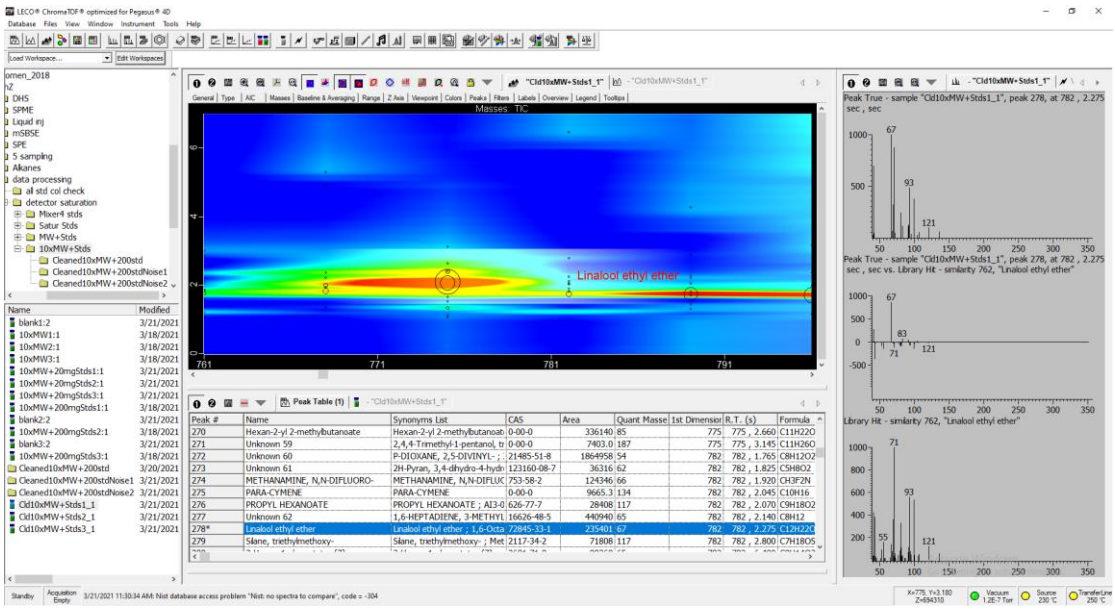


Figure S33. the mass spectrum of linalool ethyl ether after saturation correction

4 Conclusion and discussion

GC×GC is an efficient tool for the analysis of volatile compound. However, improvements are still required on VOC extraction, GC×GC setup and data processing.

Different sample preparation techniques and GC×GC setup were compared based on the literature study and experimental results. Each VOC extraction technology has its own drawbacks and needs new developments. For DHS, the sensitivity can be adjusted by optimizing the of parameters involved. However, this flexibility requires more complex instrumentation, more maintenance, and complex optimization. Find a universal parameter setup for beverage VOC profiling is not simple. The theory of DLLME requires cross disciplines knowledge. Mechanisms involved remain unclear. Many new methods belongs to DLLME have progressed, such as air-assisted liquid-liquid microextraction (AALLME) and switchable hydrophilic solvents. These methods are becoming increasingly popular and new developments, applications and improvements are expected in the future. Sample extraction efficiency limits the application of SPME. Its mass transfer efficiency can be improved by higher extraction temperature, larger contact surface and assistive technologies likes vacuum and ultrasound. With the proper setup, the extraction processing can be faster, milder, more selective and more sensitive. SBSE has only two sorbent phases are commercially available, PDMS and ethylene glycol/silicone phase. Back extraction is an interesting application of SBSE. It offers possibility of perform the analysis with HPLC, more choices on sorbent phases, and downstream odor and tasty test. SPE retains the molecule in varies mechanisms. A wide range of organic analytes, from non-polar to very polar analytes, can be absorbed with high extraction efficiency. The major drawbacks of using SPE for beverage VOC analysis is the cumbersome extraction procedure. Automation is required. Molecular imprinting-based solid-phase extraction has a bright

future. It increasing the selectivity benefits the quantification of target compounds.

There wasn’t an ideal sample preparation technique to recover all the VOCs from the beverage sample. Furthermore, the VOCs recovered by different techniques were very different. The discussion of the pros and cons of the different techniques in our study can serve as a guide for the further development and improvement of these techniques. Combining the results from different sample preparation techniques is necessary to achieve a higher coverage of global VOC profiling. For the known fermentative aromatic compounds, the best coverage can be reached by using SPME together with SPE for beer, and VALLME for wine and cider.

A fine GC×GC method development involves modulator selection, column combination and parameter optimization. Thermal modulator provides high detection sensitivity and allow exceptional trace analysis. Since the analytes coverage is the most important factor of in beverage VOC profiling, thermal modulation is a better choice. Columns are the key components of chromatographic systems. Better column orthogonality doesn’t mean enhanced separation resolution. Experimental evidence shows that semipolar columns in the second dimension is the right choice. In fermented beverages, there are more polar compounds than non-polar compounds. The most suitable column combination is polar-semipolar. Narrower diameter column in the second dimension was preferred for higher column efficiency, and short elute time. However, due to the column overload the data processing problems cause by it, use same column diameter is recommended.

GC×GC parameters such as, column-flow, 1^{st} and 2^{nd} oven temperature program, and the modulation setup must be optimized. These parameters

interact with each other therefore statistical prediction model is required. Mentioned parameters can be efficiently optimized by response surface model. Good global orthogonality is not equal to the separation suits to the typical need. In the case of non-target profiling for complex metabolic samples, direct measuring distances among the targeted peaks reflects the most real. The nearest neighbor distance was a suitable measurement for peak dispersion. When the column and detector were saturated by the major metabolites, which is unavoidable if the metabolic sample is measured at one dilution level, incorrect peak deconvolution and mass spectrum construction may happen. This limited the application of GC×GC-TOFMS to metabolites screening.

Occurred errors during peak-picking and MS construction were mainly caused by the detector and/or column saturation with major metabolites. In metabolomics studies, the nature of the sample is well defined and the major compounds that have the potential to cause saturation are almost invariably known. Data processing results can be improved by a two-stage data processing strategy that will incorporate a targeted data processing and cleaning approach upstream of the “standard” untargeted analysis. Our experiments show a significant improvement in annotation and quantification results for targeted compounds causing instrumental saturation. After subtracting the saturate signal of targeted compounds, the MS construction was improved for co-eluted compounds. Incomplete signal subtraction may occur. It leads to the detection of false positive peaks or to interferences with the construction of mass spectra of co-diluted peaks. High-resolution MS libraries and more accurate peak area detection methods should be tested for further improvement.