Volatile characterization

Optimizing GC-MS methods to reveal the chemistry and aroma of savoury food ingredients

Carmen Díez Simón

Propositions

1. Untargeted metabolomics techniques have great potential to deliver novel markers for the processing history of food materials.
   (this thesis)

2. We cannot fully rely on instrumentation to define sensory experience while human receptors in our nose remain more sensitive than a Mass Spectrometer.
   (this thesis)

3. As we try to reduce waste in the food industry, we should rethink how to diminish the waste created in laboratory/research studies.

4. Entrepreneurial courses must be integrated into all branches of research and education at schools and universities.

5. By shaving his hair, a Buddhist monk already has twenty less problems than a person with hair.

6. The over-stimulated and over-pressured mindset of people in today's world is our current mental pandemic.

Propositions belonging to the thesis, entitled

Volatile characterization. Optimizing GC-MS methods to reveal the chemistry and aroma of savoury food ingredients

Carmen Diez Simon
Wageningen, 10 September 2021

Volatile characterization

Optimizing GC-MS methods to reveal the chemistry and aroma of savoury food ingredients

Carmen Diez Simon

Thesis committee

Promotor
Prof. Dr. R.D. Hall
Professor of Plant Metabolomics, Laboratory of Plant Physiology
Wageningen University & Research

Co-promotor
Dr. R. Mumm
Researcher, Bioscience
Wageningen Plant Research

Other members
Prof. J-P Vincken, Wageningen University & Research
Prof. J. Hugenholtz, University of Amsterdam
Prof. J.G.M. Janssen, Wageningen University & Research / Unilever R&D
Dr. Y. Tikunov, Wageningen University & Research

This research was conducted under the auspices of the Graduate School VLAG (Advanced Studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences)

Volatile characterization

Optimizing GC-MS methods to reveal the chemistry and aroma of savoury food ingredients

Carmen Diez Simon

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 10 September 2021
at 1:30 p.m. in the Aula.

Carmen Diez Simon
Volatile characterization. Optimizing GC-MS methods to reveal the chemistry and aroma of savoury food ingredients, 210 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2021)
With references, with summary in English, Dutch and Spanish.

ISBN: 978-94-6395-917-9
DOI: 10.18174/550790

To my dear Alokito,
and the family that surrounded me throughout

Abstract

Volatiles as essential molecules for life on earth

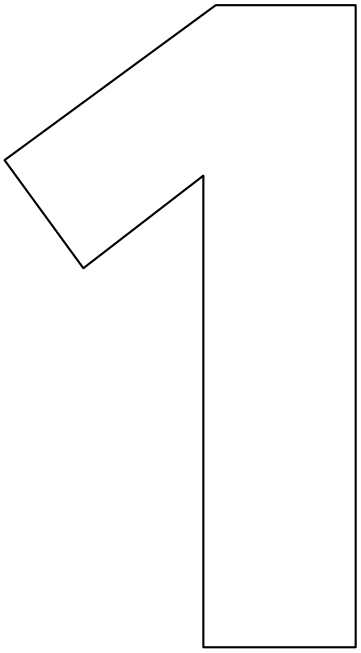
Volatiles are the responsible for all the pleasant and unpleasant smells on earth. From the delectable honey-crisp apple smell of a Riesling white wine, to the repulsive rotten odour of manure. They are small molecules that have the ability to travel through air and water, and thus, not only can they make food smell good or bad, but also they serve as signalling molecules both within and between organisms. For instance, plant volatiles attract pollinators and seed dispersers, and provide defence against pests and pathogens. For these reasons, measuring and determining the roles of volatiles is crucial in understanding living systems and reaction processes. In food, volatiles comprise thousands of molecules having highly diverse physicochemical properties, which makes the analysis of these compounds highly challenging. Therefore in this thesis we have focused on developing untargeted analytical techniques that improve and broaden the analysis of the vast array of volatiles present in a sample. More specifically, we investigated the application of metabolomics techniques to analyse the volatiles present in natural savoury food ingredients as these are superior flavourful constituents that are used to enhance the aroma and taste of the food we eat, such as soups and snacks. Moreover, work described in this thesis has also explored the intricate relationship between volatile compounds and the aroma characteristics of these food ingredients. The complexity of flavour perception added to the complex volatile profiles and reactions occurring during food processing is described in this thesis. As volatiles are known to contribute the most to the flavour of food, and metabolomics is now growing in application to processed food, this thesis sets the basis for future research into the exploitation of metabolomics technologies and on flavour perception in processed foods and their ingredients.

Table of contents

List of abbreviations

4-HDMF	2,5-Dimethyl-4-hydroxy-3(2H)-furanone
4-HEMF	2(or 5)-Ethyl-4-hydroxy-5(or 2)-methyl-3(2H)-furanone
AEDA	Aroma extraction dilution analysis
AH	Acid hydrolysed/hydrolysis
BER	Balanced error rate
CAR/PDMS	Carboxen/polydimethyl siloxane
CaSR	Calcium-sensing receptor
CIS	Cooling injection system
CV	Coefficient of variation
DHS	Dynamic headspace
DVB	Divinyl benzene
EG	Ethylene glycol
FAME	Fatty acid methyl ester
FD	Flavour dilution
FL	Flavour
GC-MS	Gas chromatography-mass spectrometry
GC-O(-MS)	Gas chromatography-olfactometry(-mass spectrometry)
HCA	Hierarchical cluster analysis
HLFSS	High-salt liquid-state fermentation soy sauce
(U)HPLC-DAD	(Ultra) High-performance liquid chromatography-diode-array detection
HSSE	Headspace sorptive extraction
HS-SPME	Headspace-solid-phase microextraction
HVPs	Hydrolysed vegetable proteins
LAB	Lactic acid bacteria
LC-MS	Liquid chromatography-mass spectrometry
LLE	Liquid-liquid extraction
LSFSS	Low-salt solid-state fermentation soy sauce
MMSE	Monolithic material sorptive extraction
MPS	Multi-purpose sampling robot
MS	Mass spectrometry
MSE	Mean squared error
MSG	Monosodium glutamate
m/z	Mass to charge ratio
NMR	Nuclear magnetic resonance
OD	Odour
ODP	Olfactory detection port
PA	Polyacrylate
PB	Process flavour blend

PCA	Principal components analysis
PDMS	Polydimethyl siloxane
PF	Process flavour
PLS	Partial least squares
PLS-DA	Partial least squares-discriminant analysis
PTR-MS	Proton transfer reaction-mass spectrometry
QC	Quality control sample
QDA	Quantitative descriptive analysis
RI	Retention index
(RP-)HPLC-MS	(Reversed-phase-)high-performance liquid chromatography-mass spectrometry
RSD	Relative standard deviation
R.T.	Retention time
SAFE	Solvent-assisted flavour evaporation
SBSE	Stir bar sorptive extraction
SBSE^2	Dual stir bar sorptive extraction
SD	Steam distillation
sd	Standard deviation
SDE	Simultaneous distillation and extraction
SPME	Solid-phase microextraction
SPTE	Solid-phase trapping solvent extraction
SVR	Support vector regression
SWATH-MS	Sequential window acquisition of all theoretical mass spectra
TAGs	Triacylglycerols
TDA	Taste dilution analysis
TDU	Thermal desorption unit
TF	Traditional fermentation
TIC	Total ion current
YB	Yeast extract blend
YE	Yeast extract
YP	Yeast derived products



General Introduction

Flavour is a combination of multiple stimuli created primarily by the senses of smell and taste, but also is driven by mouth-feel, pain (pungency), appearance, colour, temperature and texture (Dresow and Böhm 2009). The combination of all these factors tells us whether food is delicious, good, unpleasant, or downright disgusting. Smell is one of the first sensations that is perceived when about to eat food. However, the most immediate sensation when eating is taste. Nonetheless, if you hold your nose while you eat, you will notice that some foods are perceived quite differently. The same is often the case when you have a cold and you do not smell food. When we chew food, aromas are released which are moved through the retro nasal channel and activate odour receptors located in our nose (Guichard and Salles 2016). If this channel is blocked, such as when we have a cold or flu, aroma molecules cannot reach the sensory-specialized cells that are stimulated by odours, thus, food cannot be experienced in the same way. Without smell, foods tend to taste bland and flavourless, hence the importance of aroma compounds to our sensory pleasure (Spence 2015).

Food flavour is directly linked to the vast array of chemical constituents present in a sample. These chemical species are molecules that are naturally present in food, or can be formed by the action of enzymes and/or (thermal) degradation processes from their primary precursors (carbohydrates, proteins, lipids, nucleotides and vitamins) (Diez-Simon et al. 2019). For instance, the flavour of grapes is characterized by hundreds of molecules, however, when grapes are fermented into wine, these molecules will enable chemical transformations (thanks to microbial activity, temperature, etc.) that will result in distinct new flavour compounds and, in most cases, a larger number of molecules constituting wine flavour. Compounds contributing to the flavour of food products are mainly divided into two sub-categories: volatiles and non-volatiles. Volatiles are mostly associated with the smell/aroma of food, whereas non-volatile compounds are more closely related with taste sensations (sweet, salty, sour, bitter, and umami). Volatile molecules are present in the vapour phase already at ambient conditions and will interact with the smell receptors in our nose, while non-volatile compounds activate the different taste receptors in our mouth and tongue. Examples of non-volatile molecules are amino acids and small peptides (contributing to e.g. sweet and umami sensations) as well as fatty acids, organic acids, alkaloids, phenolics and many other metabolites. Aroma, which is mostly perceived through retro nasal olfaction, is the most important component of flavour (Kerth and Miller 2015; Shepherd 2005), and thus volatile composition plays an important role in flavour studies. Aroma compounds are often grouped into nine functional categories (See **Table 1**, sourced from Vilgis and Vierich 2020): (1) aliphatic hydrocarbons (e.g. 3-hexenal, 2,6-nonadienal) imparting green, waxy, fatty, fruity and mushroom-like aromas; (2) sulphur compounds (e.g. dimethylsulfide, methional) characterized by sulphurous, pungent, cabbage-like and onion-like odours; (3) acyclic terpenes (e.g. geraniol, linalool) having floral, lemon-like, slightly flowery, rose-like and lavender-like odours; (4) cyclic terpenes (e.g. limonene, pinene) imparting balsamic, camphor-like, minty, spicy, bitter, earthy and thyme-like aromas; (5) sesquiterpenes (e.g. caryophyllene, bisabolol) characterized by heavy floral, camphor-like, woody, trementine-like and earthy-vegetable odours; (6) aromatic compounds (e.g. vanillin, 4-ethylguaiacol) with aroma attributes such as almond-like, anise-like, thyme-like, vanilla and smoky; (7) phenylpropanoids (e.g. eugenol, coumarin) characterized by sweet, anise-like, nutmeg-like, spicy and cinnamon-like odours; (8) heterocyclic compounds (e.g. furfurals, pyrazines) imparting bread crust-like,

hay-like, caramel, woody, nutty and roasted odours; and, finally, (9) the trigeminal odourless compounds, such as capsaicin and quercetin, which impart irritating, warming, cooling, biting and demineralizing sensations.

Table 1. Functional categories for volatile compounds often used in food flavour (sourced from Vilgis and Vierich 2020).

	Group name	Odour characteristics	Example of volatiles
1	aliphatic hydrocarbons	green, waxy, fatty, fruity and mushroom-like	3-hexenal, 2,6-nonadienal, decanal
2	sulphur compounds	sulphurous, pungent, cabbage-like and onion-like	dimethylsulfide, dipropylsulfide, methional, allicin
3	acyclic terpenes	floral, lemon-like, slightly flowery, rose-like and lavender-like	geraniol, linalool, citronellol, nerol, myrcenol
4	cyclic terpenes	balsamic, camphor-like, minty, spicy, bitter, earthy and thyme-like	limonene, pinene, 1,8-cineole, 3-carene, borneol, safranal
5	sesquiterpenes	heavy floral, camphor-like, woody, trementine-like and earthy-vegetable	caryophyllene, bisabolol, bergamotene, camphene, selinene
6	aromatic compounds	almond-like, anise-like, thyme-like, vanilla and smoky	vanillin, 4-ethylguaiacol, thymol, benzaldehyde, cuminaldehyde, elemicin
7	phenylpropanoids	sweet, anise-like, nutmeg-like, spicy and cinnamon-like	eugenol, coumarin, anethol, estragol, myristicin, safrol
8	heterocyclic compounds	bread crust-like, hay-like, caramel, woody, nutty and roasted	furfurals, pyrazines, coumarin, sotolone
9	trigeminal odourless compounds	irritating, warming, cooling, biting and demineralizing sensations	capsaicin, quercetin, gallic acid, oxalic acid

In this thesis, I have focused on the profiling and characterization of these small molecules, more specifically volatiles and semi-volatiles, by using metabolomics technologies and sensory prediction models. Metabolomics aims to analyse the chemical constituents of biological material in a holistic way, while flavour is generally "measured" by sensory tests, such as evaluations performed by (trained, non-trained, or consumers-based) panellists who score the different perceptions and properties of flavour. The metabolomic and sensory data are then statistically evaluated in order to predict the sensory attributes from chemical patterns following a data-driven modelling approach. The outcome of such approaches is relevant for improving food formulation as well as understanding food processing reactions. Processing of the raw ingredients is an important way to improve palatability and to create specific flavours in a broad range of food products, from yogurt, to tinned vegetables, to biscuits, or to savoury snacks. In **Figure 1**, an integrated scheme exhibits the steps towards understanding flavour formation during food processing. This is essentially the basis for this thesis. In order to formulate a desired food product which follows certain processing steps, sensory and instrumental analyses are combined, and these help us evaluate the relationship between flavour, molecules and reaction pathways. Consequently, new re-formulations guided by the knowledge of the sensory and chemical information allow food scientists and industries to create

product diversity that suits consumer liking across the globe. Such approaches have been applied in this thesis to fermented savoury ingredients developed to boost the flavour experience of processed food.

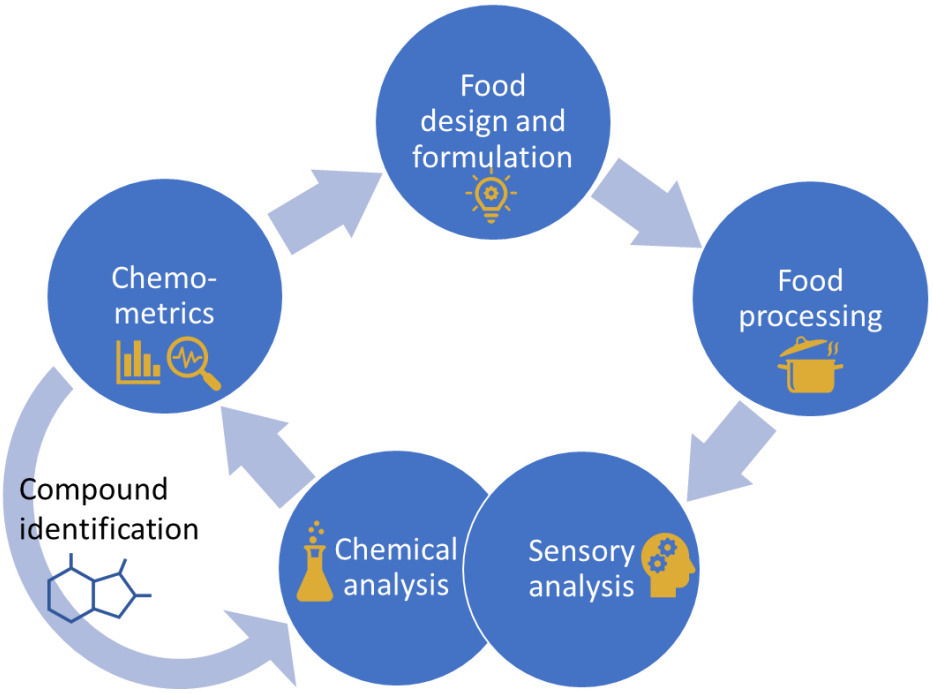


Figure 1. Schematic representation of the steps towards relating sensory and chemical data to understand flavour formation and processing strategies needed for the design of new, superior flavoured food.

Fermented savoury products that make food more palatable

Nowadays, food quality has become one of the most important traits in the food and flavour industry even though it is difficult to define in generic terms. The major challenges to the food and flavour industry are to provide safe, healthy, natural food products, which meet the high standards of consumers (meaning e.g. food with high, balanced flavour). Thus, one of the most used strategies by industry to create flavour is through the use of fermented ingredients. Although artificial flavourings are still widely used in industry, these generally do not meet the high-quality standards of food as mentioned above. Fermentation is one of the main processes which generates an enormous number of chemical compounds to enrich the human diet, and hence, creates new, specific or enhanced aroma, taste and texture experiences that cannot be found in fresh food (Steinkraus 2002). An example of a fermenting organism widely used in the flavour industry is yeast (*Saccharomyces cerevisiae*) and its derived fermented products. Yeast-derived products are generally used as flavouring agents, to simulate particularly meat-like, broth-like or cheese-like flavours, and to aromatize snacks, soups and cheese products (Sucan and Weerasinghe 2005). Yeast has the advantage of producing protein-rich and flavour-rich products which meet the industrial challenges mentioned above. Sometimes it is also used to create one specific aroma compound, which is expensive to make from the original source. Yeast can also be tweaked to make all kind of different products relatively easily and nowadays a large number of strains are available giving the flavour industry a large diversity of flavouring options. Moreover, yeast fermentation goes in line with more sustainable production approaches since it is

a very effective process and can be used to replace the flavour of certain animal-based products, the latter having higher environmental impact.

The main route for generating flavour compounds from the fermentation of yeast is the hydrolysis of proteins (peptides) and sugars (carbohydrates), which is initiated by a thermal treatment after the fermentation stops, upon which intracellular enzymes start hydrolysing, the so-called autolysis process (Běhalová and Beran 1979). When combined with high temperature processes, Maillard reactions in which non-enzymatic reactions between amines (amino acids, peptides, proteins) and reducing sugars are catalysed. These are known to create a vast range of flavour compounds which provide desirable attributes like roasted or baked aroma, but can also bring off-flavour notes. By developing production processes that can trigger the formation of tasty compounds while avoiding the formation of the less pleasant molecules, the food industry can continue to meet consumer preferences, and offer more sustainable alternatives of processed food products (**Figure 2**). This whole topic has been extensively covered in **Chapter 2** of this thesis which reviews the published literature on the different classes of aroma-related volatiles provided by yeast derived flavour products (e.g. yeast extracts and process flavours) that characterize processed food, and brings chemistry and technology aspects together.

Another example of a complementary but contrasting savoury food product which is also covered in this thesis, is soy sauce (**Figure 2**). Soy sauce is a highly valued and widely used fermented condiment that gives food an extra boost of flavour attributes such as salty, umami and smoky characters. Soy sauce has been widely consumed all around the world since ancient times, and is currently exploited in the flavour industry as an ingredient to enhance the umami taste and overall aroma of specific food products, such as bouillons, snacks and ready-to-eat soups. In **Chapter 3**, an overview of the available literature on the chemical and sensory characteristics of soy sauce is reported in a second comprehensive review paper. The description of the main processes involved, which go from raw ingredients to obtaining the diversity of final soy sauce types, is explained. Soy sauce types are not only described from a chemical point of view, but also from a cultural perspective. Quality differences exist between soy sauces found on the global market, and it is not yet well known what the fermentation products are that characterize a 'good' or a 'bad' soy sauce. However, some studies are helping to advance our knowledge of flavour formation and metabolite profiling of its aroma and taste compounds (Kong et al. 2018; Lee et al. 2019; Lee et al. 2006; Lioe 2007). The origin of these compounds relates to both the raw ingredients chosen (soybeans, wheat and brine) and the starter cultures used – but also to the specific parameters applied during production, which are often culturally linked. Soy sauce taste is dominated by umami and salty sensations. Free amino acids, nucleotides, and small peptides are among the most important taste-active compounds. Soy sauce aroma is characterized by caramel-like, floral, smoky, malty, and cooked potato-like odours (Devanthi and Gkatzionis 2019). Aroma-active volatiles are chemically diverse including acids, alcohols, aldehydes, esters, furan(one)s, pyrazines, and S-compounds.

There are two main challenges that need to be tackled in flavour studies. One being the limitation in methodological/analytical technologies intending to analyse the whole chemical profile of a sample (the

metabolome), and the other being the complex relationship between flavour sensory and chemical composition. These topics are introduced and discussed in the following sections.

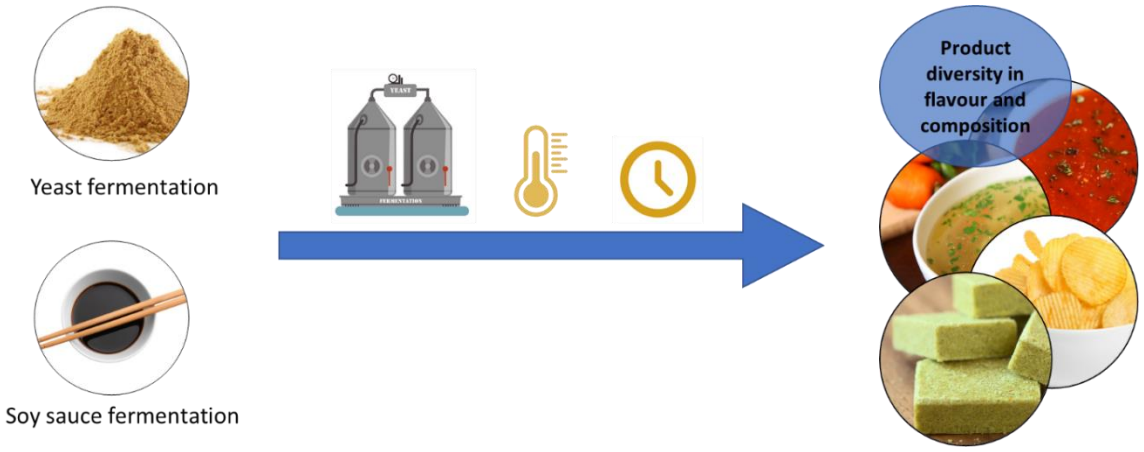


Figure 2. Representation of the production flow of fermented savoury food products. Products such as yeast extracts and soy sauce are used to create a high range of product diversity in flavour as well as in the final metabolite composition. Different conditions (fermentation, temperature, time and addition of other base-ingredients) can generate products with distinct flavour qualities.

Metabolomics: state-of-the-art technologies in flavour science

Metabolomics is the field of research that develops methodologies to analyse the vast array of molecules present in a biological sample, and is often used to compare the metabolite composition of different groups of samples (Zhang et al. 2012). Metabolomics mostly relies on advanced analytical techniques using e.g., mass spectrometry (MS) and spectroscopy (NMR) as detection methods. It has been applied to many fields of expertise, such as medical and pharmaceutical applications, environmental, microbial, plant and food sciences, etc. The broad scope of applications of metabolomics in food science and nutrition cover topics related to for example, food quality, food processing (pre-harvest, post-harvest and storage/packaging), toxicology, and microbiology and this has recently been well reviewed (Cevallos-Cevallos et al. 2009; Cevallos-Cevallos and Reyes-De-Corcuera 2012; Hall et al. 2008; Scalbert et al. 2009; Wishart 2008). In this thesis, I have expanded the knowledge and applications of state-of-the-art metabolomics technologies to the flavour quality of fermented savoury products.

Metabolomics is still advancing in terms of metabolite coverage by applying untargeted methodologies, which means taking into consideration all (or most) types of small molecules in one analysis, without targeting only the known/identified metabolites. It is clear that there is no single technique that can analyse the whole metabolome of a food sample, however, combining complementary metabolomics platforms with untargeted data mining methodologies allows us to explore a greater part of the metabolome of a food product and to study the relationships and interactions between metabolite classes. To achieve this, sophisticated sample preparation and extraction methods are key to ensure the accuracy and selectivity of the results needed for the comprehensive interpretation of the data (Carneiro et al. 2014). Techniques should ideally reach maximum extraction with minimal chemical alteration of the product (Carneiro et al.

2014). This thesis has developed and compared sample preparation and extraction methodologies that showed high reproducibility and high metabolite coverage using untargeted metabolomics techniques for the analysis of volatile compounds. Sorptive-based techniques have been chosen for this strategy, as these do not require altering the product matrix by using complex (and toxic) organic solvent extraction procedures, and are fast, cost-effective and not labour intensive (Nogueira 2015). Moreover, sorptive-based techniques have proven to work well for food flavour volatiles (Kataoka et al. 2000). These methods commonly use ab- and adsorptive polymer-based materials in which the volatiles are trapped either in (liquid) or above (headspace) a food matrix (**Figure 3**). Headspace extraction (e.g. SPME: Solid-phase micro extraction, DHS: Dynamic headspace extraction) or in-liquid extraction (e.g. SBSE: Stir bar sorptive extraction) are the most popular among all the techniques proposed in recent years (Nogueira 2015) and are mostly coupled to a Gas Chromatography – Mass Spectrometry (GC-MS) system in order to separate and identify those volatiles present. The extraction techniques which were tested in this thesis were: SPME, DHS, SBSE and headspace sorptive extraction (HSSE) as outlined in **Figure 3**. SBSE and HSSE are based on the trapping of volatiles onto an absorbent polymer coated on a magnetic stir bar which can be placed in the liquid sample or in the headspace (David et al. 2019). On the other hand, SPME uses a fused-silica fibre that is coated with one or more polymers to trap the volatiles. DHS is a dynamic headspace trapping technique, and its main difference to the aforementioned static techniques is that it traps volatiles on an adsorbent cartridge through flushing the sample with a continuous flow of gas, aiming to accumulate analytes more efficiently. All these sorptive-based techniques have their own advantages and disadvantages, which have therefore been further explored in this thesis. For instance, the adsorbent coatings have distinct volumes, configurations, and polymers, which make the comparison of the techniques interesting for certain food products. In this thesis, we compared these four extraction techniques, and several approaches were employed for selecting the most appropriate technique for savoury ingredients.

When developing metabolomics techniques, not only are the sample preparation and extraction steps important, but also particular challenges have to be considered when analysing aroma (volatile) compounds in food. These challenges are related to the complexity of food matrices (having complex mixtures of macromolecules and small molecules), the presence of aroma molecules at low concentrations (ppm, ppt), the high volatility of some compounds that risk being lost during analysis, and the instability of some components in dynamic equilibria with other constituents in the food matrix (Dresow and Böhm 2009). Processed food ingredients are known to contain highly complex mixes of both primary (e.g. sugars, lipids and peptides) and secondary metabolites, which can influence the detection of many aroma compounds. The release of aroma compounds from foods with headspace techniques is determined by the partition coefficient between the air phase and food matrix (and also between the hydrophobic and hydrophilic liquid phases). The nature of these interactions depends on the physicochemical properties of flavour compounds (Jeleń et al. 2012). For instance, fat composition, concentration, emulsion characteristics and temperature, can each significantly modify interactions between lipids and small molecules (Guichard 2002; Piraprez et al. 1998). In addition, studies have shown that carbohydrates can also influence the retention and release of volatile flavour compounds (Naknean and Meenune 2010). Proteins have also been shown to bind covalently to thiols and disulphides (Adams et al. 2001) removing,

in this way, potent odorants from the system. In this thesis, we have analysed complex processed food matrices such as tomato soups which contained savoury ingredients (yeast derived flavour products and high concentration of sugars, amino acids and fatty acids). Therefore, methodologies were developed in order to get the best picture of the volatiles present in the complex samples. The continuous development of more sensitive technology and the search for new (untargeted) methodologies is increasing our knowledge and the discovery of new taste and aroma compounds, and how they are formed during food processing.

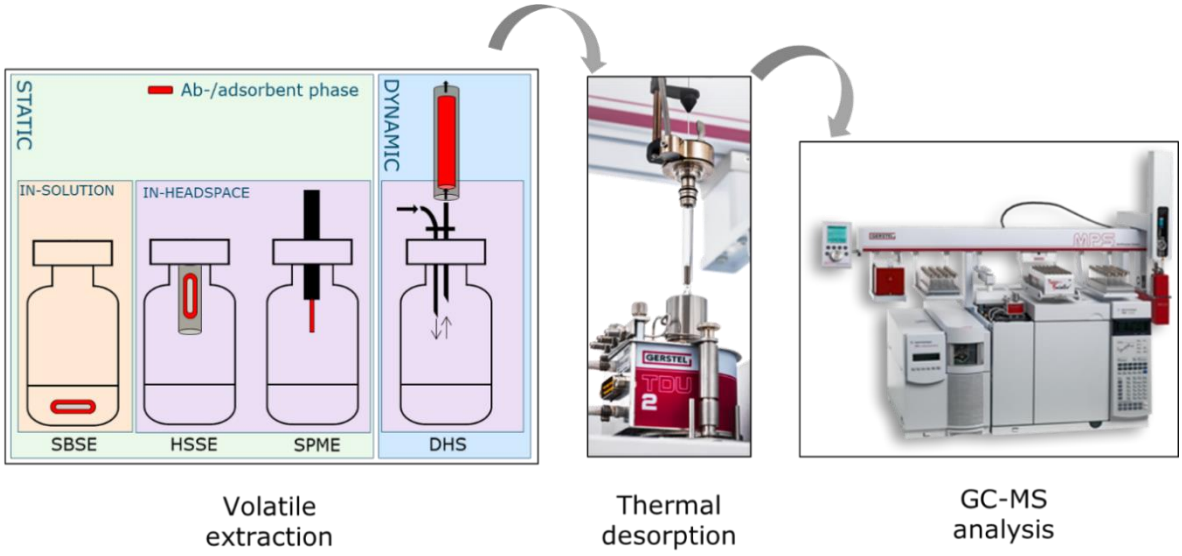


Figure 3. A schematic representation of the four volatile extraction techniques used in this thesis, which were coupled to a thermodesorption Gas Chromatography – Mass Spectrometry (GC-MS) system. The ab-/adsorbent coatings of the extraction techniques are highlighted in red. We differentiate between static in-solution, static in-headspace and dynamic in-headspace approaches. SBSE: Stir-bar sorptive extraction; HSSE: Headspace sorptive extraction; SPME: Solid-phase microextraction; DHS: Dynamic headspace system. The two images on the right side are sourced from Gerstel (Mülheim, Germany).

Chemometrics to link chemical composition with production processes and sensory attributes

The basis for developing metabolomics methodologies in food flavour studies is to get a comprehensive insight into the small chemical molecules which, in combination, contribute to the unique sensory characteristics and the overall aroma and taste of food. These sensory characteristics are important when it comes to consumer preferences. Hence, not only the analysis and identification of the chemical compounds is relevant, but also the sensory evaluation of the product and thus the correlation between the chemical composition and the sensory descriptors. Finding this correlation becomes more challenging in complex food matrices, where many interactions could impact the sensory and chemical evaluation.

Metabolomics technologies are used in combination with chemometrics analyses to evaluate and identify biochemical patterns within samples. The complex metabolomics datasets are commonly analysed using (un)supervised chemometric methods. The most common unsupervised methods used in food metabolomics are principal components analysis (PCA) and hierarchical cluster analysis (HCA). PCA is often

the first step in multivariate data analyses and helps us to get an overview of the general structure of the dataset and to visualize trends, groupings and outliers. In HCA, samples are typically grouped based on the similarity of their chemical profiles using dendrograms. HCA and PCA are useful methods that have been widely adopted to analyse volatile compounds from many food types, for instance rice varieties for quality traits (Mumm et al. 2016), or the aroma qualities of melon varieties (Allwood et al. 2014). However, PCA and HCA are not methods for classification approaches but rather are a step towards this. On the other hand, supervised methods, such as partial least squares discriminant analysis (PLS-DA), allow classification of data based on prior known information (by making classes beforehand), and is also widely used for metabolomics data. These chemometrics approaches help flavour scientists to evaluate the chemical data, and relate changes in the chemical composition to changes in the processing treatments as well as the flavour qualities. Partial least squares (PLS) analysis is the most commonly used statistical method in relating sensory properties with instrumental data (Seisonen et al. 2016). However, other methods such as elastic net and random forest are also used and have been exploited in this thesis (Shi et al. 2019).

Sensory evaluations are techniques which measure human perceptions of food (Lawless and Heymann 2010). Thus, sensory properties are determined to provide useful information to food scientists, product developers and related industries. Sensory panels can deliver important information about product quality and consumer preference which is essential for the development of new food concepts (Chin and Slupsky 2013). Nevertheless, challenges always remain when using intricate sensory data involving human responses, as humans are always associated with some degree of subjectiveness and other variation in their responses. However, advances in statistical methodologies are helping better understand the product-person interface (Lawless and Heymann 2010). Sensory evaluations are often combined with instrumental data in order to study both the sensory and chemical properties of a food product as well as their interconnection. Instrumental data includes any analytical measurement of flavour compounds, such as those from metabolomics approaches, whether it is (GC, LC, PTR)-MS, or whether it is from a sensory-guided instrumental analysis, such as receptor-based assays, cell-based assays, olfactometry, and electronic nose and tongue approaches (Batenburg et al. 2016; Feng et al. 2017; Riedel et al. 2017; Zou et al. 2018). When sensory data are combined with instrumental data and evaluated statistically, many other challenges are faced. Some of these are related to data fusion from different platforms (e.g. GC-MS and LC-MS), the non-linear relationship between sensory and chemical data, the presence of masking/amplifying effects, the validation of the model, etc. (Chambers IV and Koppel 2013). Nonetheless, it still remains a trending topic in food science as it allows flavour scientists to understand better taste and aroma formation in certain food matrices. Such data-driven approaches which combine both metabolomics analysis and sensory evaluations have led to success in relating flavour with chemical compounds. Examples can be found for various food products such as wine (Malherbe et al. 2013), olive oil (Procida et al. 2016), apples (Corollaro et al. 2014) and dairy products (Croissant et al. 2011). Additionally, several review papers explain the interaction between instrumental and sensory data (Auvray and Spence 2008; Poinot et al. 2013; Ross 2009). As an example, hexanal is a straight-chain aldehyde and it is naturally present in many vegetables and fruits, but also highly present in processed foods like dairy and processed meat, where it is formed by fat oxidation reactions. Hexanal has been associated with attributes such as green/grassy in many food types, by several

researchers (Chambers IV and Koppel 2013). However, in some cases, others have also found correlations with off-flavours in black walnuts, linking it to oxidization or rancidity (Lee et al. 2011). These, and many other examples found in the literature, illustrate how complex flavour is, and how multiple chemical interactions could play roles in complex food matrices regarding the final sensory perception.

It is important to use large datasets of samples in order to make associations between (non)linear relationships, like volatiles and sensory. Therefore, the work reported in this thesis has developed and assessed both untargeted metabolomics analyses and sensory approaches applied to an important category of savoury products in order to make associations that will help in the innovation of products towards designing better tasting food. In savoury products such as yeast derivatives, these approaches have not yet been developed in depth. Furthermore, these data-driven approaches not only work for the specific study shown here, but can be also extrapolated to several other processed products and approaches whose aim is to relate sensory changes to chemical composition.

Objectives

1. To review both the chemical and sensory characteristics of savoury food ingredients, as well as the chemical reactions typical of food processing, which lead to the formation of volatile aroma compounds (Maillard reactions, lipid oxidation, lipid-Maillard interactions, etc.). In this way, the origin of our chemical and sensory knowledge can be place in context.

2. To optimize and compare high-throughput untargeted metabolomics technologies that result in a high coverage of analytes. For this, comparative approaches will be developed to select the most appropriate method for each sample-application purpose.

3. To relate the abundance of volatile compounds with certain processing steps (or ingredients) that will allow us to better understand how flavour is formed during the production of savoury food ingredients (yeast derivative products, soy sauce), and to link volatile composition with (desired and/or off-flavour) sensory attributes that will allow us to predict and/or re-formulate new flavour experiences.

Thesis outline

The major aim of this thesis was to zoom in on the metabolite profile of savoury food ingredients to better understand what are the constituents (such as yeast derivatives) and the production processes leading to superior and desired flavour qualities. Therefore, first I include two literature overview studies, which cover the background knowledge of both the savoury ingredients used in this thesis (yeast derivatives and soy sauces). Secondly, I compared high-throughput metabolomics technologies that allow an accurate and comprehensive detection and identification of volatile molecules present in a savoury product and, thirdly, I expanded this to a data-driven approach aiming to understand and predict aroma and taste qualities related to the ingredients and/or processing steps of a set of contrasting savoury food products. **Figure 4** illustrates the thesis outline which is further summarised below.

Chapter 2 of this thesis provides a comprehensive and up-to-date review of the major flavour compound classes described in processed savoury ingredients, including the formation of these compounds from their precursors. Special attention is given to the interconnections between Maillard reactions and the different amino acid, lipid, and carbohydrate degradation pathways. Furthermore, the chapter provides insights into advancing metabolomics applications that have not yet being exploited in depth for processed food ingredients.

As a continuation, in **Chapter 3,** a second literature study compiles the chemical compounds being present in the most common soy sauces, from different origins, and their potential sensory relevance. This review also presents a sensory wheel of taste and aroma attributes that characterize soy sauce flavour. Soy sauce is a condiment used worldwide and is perhaps one of the most complex fermented condiments containing many small molecules that boost the flavour of many dishes. Here also little has yet been done in implementing comprehensive metabolomics analyses to study the metabolite composition of soy sauces. With these two reviews, I aimed to compile not only the list of relevant chemical groups but also promote the potential of metabolomics to study flavour in savoury processed foods.

The next step was to develop and compare untargeted metabolomics techniques that can detect as many of these reported metabolites as possible, using a single approach. In **Chapter 4**, four volatile extraction techniques were optimized and compared in order to select which approach gave the most comprehensive dataset of volatile metabolites in yeast derived process flavours (PF). As a consequence of this, SBSE was selected for further use (SBSE-GC-MS). This methodology has also been further optimized and used to analyse a set of soy sauces originating from different countries that were characterized by different ratios of ingredients or production methods (**Chapter 5**). I also report on a pilot analysis combining untargeted metabolomics with olfactometry analyses in order to correlate compounds with their individual aromas.

This thesis culminates in the evaluation of two statistical approaches applied to study the robustness of analytical techniques (**Chapter 6**) and the quality of a prediction model that aimed to link chemical profiles and sensory characteristics (**Chapter 7**). Chapter 6 provides a methodology to select between potential metabolomics techniques for a given metabolite - sensory relationship study using the untargeted approaches developed in the previous chapter. In Chapter 7, data-driven approaches were used to correlate certain aroma attributes with the presence of certain metabolite groups to start to develop a comprehensive picture of the presence of volatile components, and their potential links to specific sensory properties as derived from a descriptive trained panel sensory evaluation. For this, savoury products were used which are described as complex matrices (chicken bouillon) containing different yeast based savoury ingredients (yeast extracts, YE; or process flavours, PF) that have been designed to create a variety of products and flavours for the processed food market.

In **Chapter 8** I integrate all the findings of this thesis, from an extensive literature study, to the use of metabolomics, to the development of data-driven approaches, in order to close the circle of food processing design and innovation. I discuss the importance of advancing untargeted metabolomics studies for selecting those processing traits which give the desired taste and aroma in savoury food products. Some of these new

methodologies are now already implemented at the company partners. Moreover, chemical markers have been proposed which are able to predict certain sensory attributes (e.g. roasted odour and chicken odour) by combining metabolomics with sensory evaluations in a chemometrics approach.

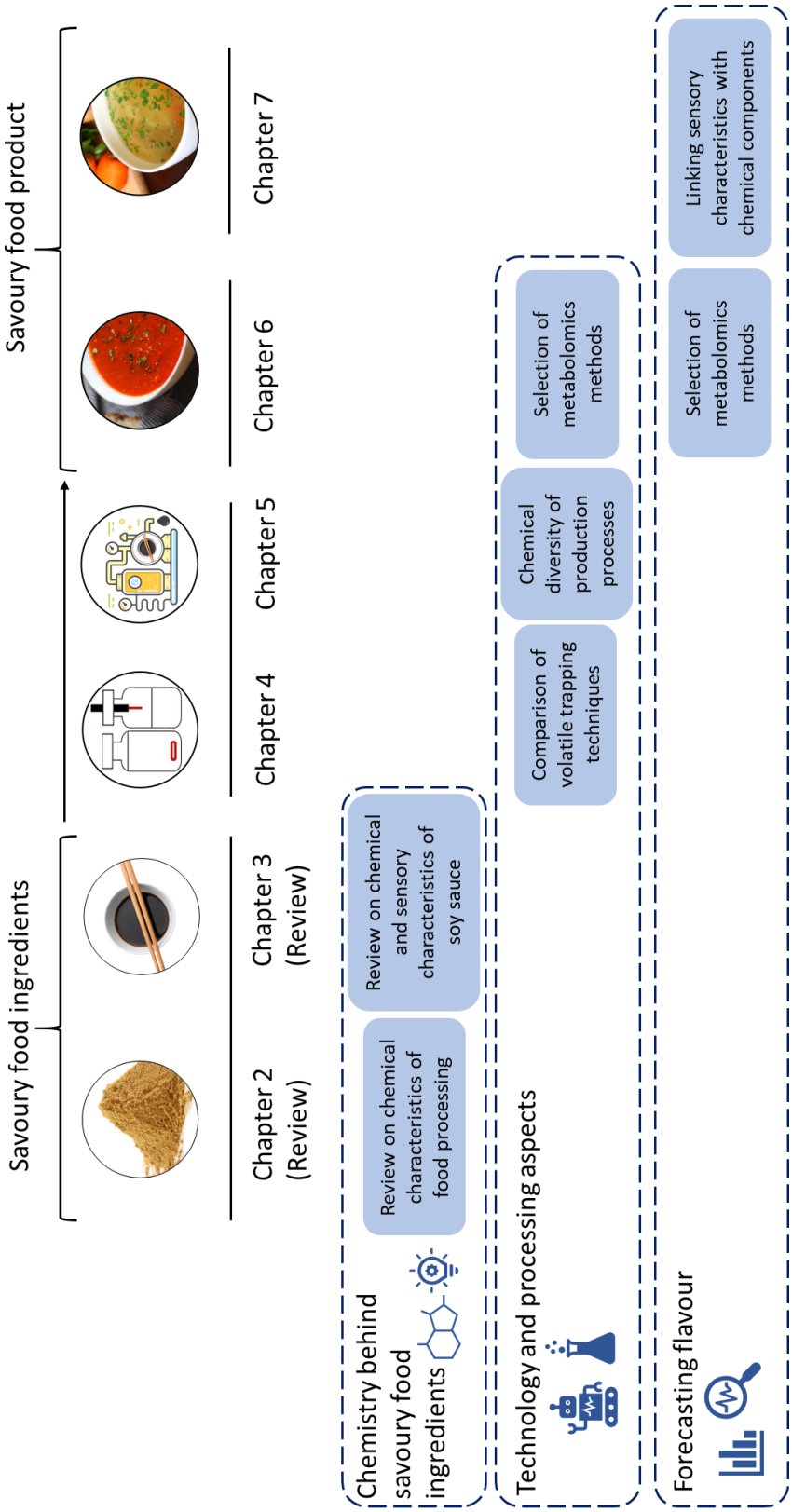


Figure 4. Outline of this thesis.

Mass Spectrometry-based metabolomics of volatiles as a new tool for understanding aroma and flavour chemistry in processed food products

Carmen Diez-Simon, Roland Mumm, Robert D. Hall

Metabolomics (2019)
DOI: 10.1007/s11306-019-1493-6

Abstract

Background: When foods are processed or cooked, many chemical reactions occur involving a wide range of metabolites including sugars, amino acids and lipids. These chemical processes often lead to the formation of volatile aroma compounds that can make food tastier or may introduce off-flavours. Metabolomics tools are only now being used to study the formation of these flavour compounds in order to understand better the beneficial and less beneficial aspects of food processing.

Aim of Review: To provide a critical overview of the diverse MS-based studies carried out in recent years in food metabolomics and to review some biochemical properties and flavour characteristics of the different groups of aroma-related metabolites. A description of volatiles from processed foods, and their relevant chemical and sensorial characteristics is provided. In addition, this review also summarizes the formation of the flavour compounds from their precursors, and the interconnections between Maillard reactions and the amino acid, lipid, and carbohydrate degradation pathways.

Key Scientific Concepts of this Review: This review provides new insights into processed ingredients and describes how metabolomics will help to enable us to produce, preserve, design and distribute higher-quality foods for health promotion and better flavour.

Keywords: mass spectrometry (MS); gas chromatography (GC); Maillard reaction; food processing; flavour chemistry; volatiles; process flavours

1. Introduction

Flavours are perceived by the aroma and taste receptors in the nose and the mouth, respectively. There are five base taste types (sweet, salty, sour, bitter and umami), which are supplemented by the specific perception of many different aroma compounds. The most important characteristic of flavour is the aroma, and the contribution of the odorous volatile substances. Flavour chemistry is the study of the chemical compounds considered to cause an aroma and/or taste. By performing taste panel studies, we are able to discover whether a food product tastes or smells good, bad or 'better than others'. However, it is still not well understood which (types of) compound(s) are imparting these characteristics (Sucan and Weerasinghe 2005) and how the interaction between different components including chemical composition, formulation, temperature, etc. influence the overall consumer perception. The chemical groups that influence the aroma characteristics of food are mainly volatile. However, non-volatile compounds can also play important roles in the sensory aroma profile, either as flavour precursors or directly, as flavourings. These are related to the sweetness, bitterness, sourness, saltiness and umami sensations and so, they make a strong contribution to the flavour of processed foods. Amino acid components, nucleotides, phenolic compounds, organic sugars and fatty acids are all examples of chemical groups that can have a role in determining the overall flavour of processed food. However, non-volatile compounds are beyond the scope of this review. Here, we focus on the volatile/aroma chemistry of process flavours (PF) through the use of metabolomic tools.

In this review, we give particular emphasis to the application of, and potential for, metabolomics approaches to advance our knowledge specifically on process flavours and flavourings as widely used in the food industry. We shall also focus on 'natural' ingredients as nowadays food companies are giving increasing importance to new food experiences which increasingly specifically involve only 'natural' flavours, colours, and preservatives (Attokaran 2017). According to the European Union, an ingredient is considered natural when it has been entirely derived from a source material that is vegetable, animal, or microbiological in nature, and at the same time, has been created through traditional food preparation processes (EU Regulation 1334/2008 Article 16 clause 2, 16 December 2008).

1.1 Food processing

Our modern lifestyle and the ever-growing global population have caused increased demands on the food processing industry. Food processing can be defined as the physical and/or chemical manipulation of raw food to generate products that can easily be prepared and served by the consumer. At the same time, processed foods can generally be more easily stored for longer periods, facilitating a broader availability to a global population. Preparation of food involves a great variety of processes; from mincing, to pasteurisation, to cooking, to fermentation, to packaging, etc. Thus, foodstuffs such as cheese, bread, breakfast cereals, tinned vegetables, savoury snacks, biscuits and milk are all processed foods.

Food quality can be influenced by the preparation process. This directly affects nutritional value and potential health benefits of food, as well as the sensory attributes (Tamanna and Mahmood 2015). Therefore, food processing can have both beneficial and detrimental effects. Beneficial effects include the

improvement of digestibility and bioavailability of nutrients, inactivation of food-borne pathogens, toxins or other detrimental constituents, prolongation of shelf-life and the improvement of the texture, taste and smell (van Boekel et al. 2010). All these changes increase consumer attractiveness. On the other hand, processing can also induce deleterious effects, such as loss of vitamins and other nutrients, the formation of toxic compounds or of compounds conferring negative effects on flavour perception, texture or colour. A well-known example for instance, is acrylamide which has been classified as a probable carcinogen in humans (Tareke et al. 2000). During preparation at high temperature, acrylamides can be formed in many types of foods via Maillard reactions (Mottram et al. 2002) including fried potato products (Vinci et al. 2012). By identifying which chemical species directly contribute to flavour perception, food manufacturers gain a better mechanistic understanding of how to produce more palatable and safer food through directing the formation of desirable flavour attributes and reducing the occurrence of undesirable ones.

Next to the basic raw materials used, other components are often added during processing for a wide range of reasons related to stability, appearance, flavour enhancement, aroma, etc. These additives can be constituted by single molecules such as so-called top notes as well as monosodium glutamate (MSG), sugar and organic acids or mixtures like protein hydrolysates or extracts. A specific group of ingredients are the so-called 'reaction flavours' or 'process flavours' (PF). These are complex mixtures which are often added to savoury products such as soups and sauces, but also to coffee, to modify taste and enhance specific sensory attributes in the final product (Sucan and Weerasinghe 2005). Process flavours have complex origins, related to spices, fruit (juice), vegetable (juice), yeast, herbs, bark, buds, dried roots, leaves or any other edible portions of a plant, or fermentation products. Processed foods therefore tend to have more complex biochemical profiles as compared to the fresh materials.

1.2 Process flavours (PF)

For many decades, the food industry has tried to make plant-based food taste 'meaty' in support of the growing vegetarian and vegan communities. Cooked meat flavour has been one of the main focus points in processed food flavours (Kerth and Miller 2015). The most important step in creating meat flavour is the Maillard reaction. Maillard reactions are a complex group of chemical reactions that occur between amino acids and sugars. They trigger a great number of reactions that lead to the formation of flavour compounds, also characteristic of brown colour formation. The first commercial use of Maillard reactions to produce process flavourings took place in the 1960s at the Unilever subsidiary company Food Industries Ltd. A number of landmark patents were filed (May and Akroyd 1960; Morton et al. 1960), which protected the reactions of the key precursors of meat flavour. Prior to this, a savoury character was usually generated through the use of hydrolysed vegetable proteins (HVPs), spice blends or actual meat extracts. However, these materials did not infer the desired meaty flavour (Parker 2015a) and were not appropriate for a vegan diet. Nowadays, alternative strategies to produce meat-type flavours are being developed by using natural yeast-based ingredients.

Yeast (*Saccharomyces cerevisiae*) has become a regular ingredient in the food industry. Some of the most outstanding ingredients now used for natural flavouring in PF are yeast extract-based products and yeast

autolysates (In et al. 2005), particularly when a meaty aroma is required (Lin et al. 2014). These enhance the food flavour by imparting cheesy, meaty or savoury notes, but can also be used as texturizers, stabilizers and thickeners (Sucan and Weerasinghe 2005). Yeast extracts are also known for their nutritional benefits as they have a relatively high content of protein, vitamins (B1, B2 and nicotinic acid) and minerals (Sucan and Weerasinghe 2005). In addition, they can be used for salt reduction in processed food without compromising the 'saltiness sensation' of the food product (Batenburg and van der Velden 2011). By varying the autolytic conditions used in yeast processing, such as temperature and time, different meaty and savoury notes can be obtained to suit different industrial purposes (Ames 1994; Münch et al. 1997; Mahadevan and Farmer 2006). Components can be either already present in the starting materials or they are formed as a result of the processing strategy used (**Figure 1**). Cooked foods develop characteristic flavours and colours, which are formed through complex series of reactions mainly related to Maillard reactions, lipid oxidation, and thermal degradation (Parker 2015a). These reactions and their importance to flavour development are explained in detail throughout this review.

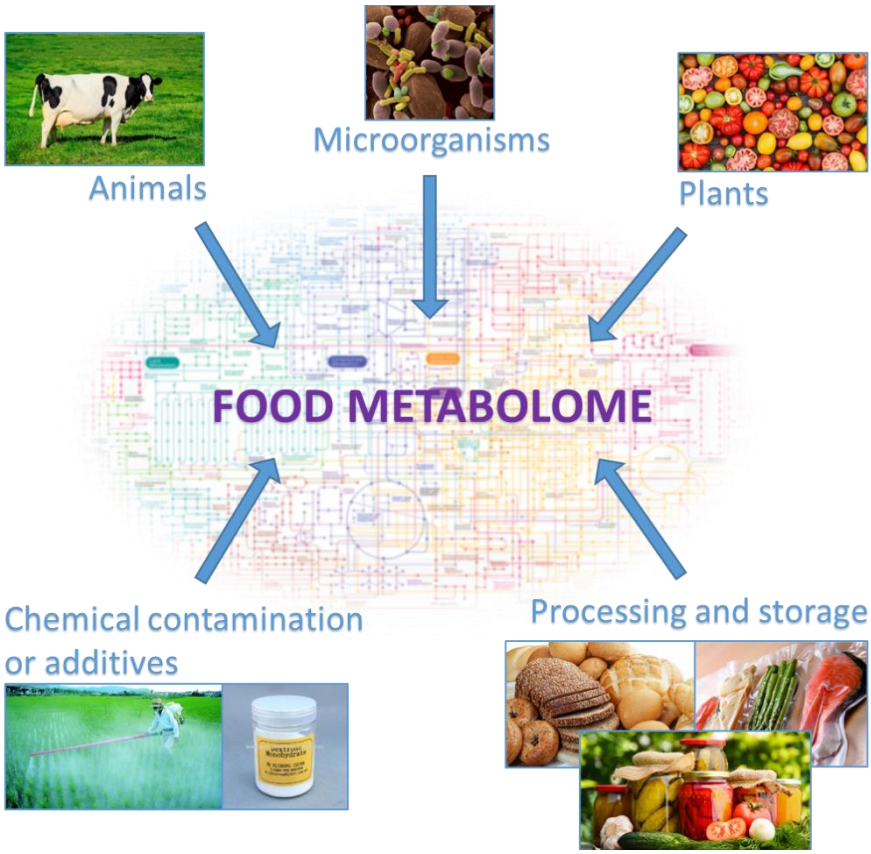


Figure 1. In terms of metabolites, food is a very complex material. Plant, animal and microbial materials cannot only be directly consumed but also often after highly-influential processing steps. Here the main sources of food metabolites are given which together constitute the food metabolome (modified from Johanningsmeier et al. 2016)

2. Metabolomics in food processing

Metabolomics is defined as the comprehensive characterization of all the small molecules present in a biological sample and is used to compare accurately the metabolite profiles between groups of samples

(Zhang et al. 2012). Small-molecule metabolites play a central role in food quality as they are often the coloured, fragrant or bioactive compounds contributing directly to nutritional value and to both positive sensory attributes as well as negative ones such as the so-called, off-flavours. The general application of metabolomics in food science and nutrition has been reviewed (Wishart 2008; Cevallos-Cevallos et al. 2009; Scalbert et al. 2009) and the importance of improvements in food analytical chemistry, such as high-resolution mass spectrometry and advanced statistical techniques to process the large data sets, have already been emphasised (Rubert et al. 2015). However, many limitations are still evident especially relating to the high complexity of the processed food matrices and the importance of low abundancy compounds with low aroma thresholds.

More than 20,000 compounds are known from food (www.Foodb.ca). To understand the relationship between food quality and processing, a complete analysis of the metabolites present in a food sample is needed (Thissen et al. 2011). State-of-the-art metabolomics comprises analytical platforms such as Gas Chromatography and Liquid Chromatography-Mass spectrometry-based techniques (GC-MS and LC-MS) and Nuclear Magnetic Resonance (NMR) spectroscopy. Each approach has its own advantages and disadvantages (which are beyond the scope of this review: see Johanningsmeier et al. 2016). Each can routinely be used to obtain metabolomic data sets due to their versatility, dynamic range, sensitivity, unique accessibility, etc. (Marshall and Powers 2017). Metabolomics has improved our capacity to analyse the overall metabolome as well as helping to perform pathway analysis and metabolite identification (Zhang et al. 2012; Van Duynhoven and Jacobs 2016; Marshall and Powers 2017). Here, we focus primarily on Mass spectrometry (MS) – based techniques as these are the most widely used in food science.

Significant compositional changes occur during pre-harvest, post-harvest, and processing of foods. Metabolomics approaches have been widely exploited at each of these stages to help advance our knowledge of firstly, which components are present in which parts of our food materials and how they change or appear in time (Kim et al. 2016) and secondly, how specific processing strategies influence final composition (Tamanna and Mahmood 2015). Van Boekel et al. (2010) and Tamanna and Mahmood (2015) have discussed general aspects of food processing on nutritional components. However, we still need more knowledge of the other quality aspects not directly linked to nutrition.

Studying how processing affects sensorial properties of food ingredients used as flavourings is of specific importance. Plants such as onion (*Allium cepa*) are used for enhancing the flavour of many processed flavours due to the high content and variety of sulphur containing compounds. Processing operations such as drying/dehydration, high hydrostatic pressure, ohmic heating, etc. can influence both the abundance and composition of metabolites in the final product. Colina-Coca et al. (2013) reported a new approach using dynamic headspace (DHS) GC-MS for the analysis of volatile compounds in onion to evaluate how processing operations affect the final aroma profile. Metabolomics has also been used for the analysis of polysulphides, which are the primary compounds determining garlic flavour, derived from the degradation of allicin when garlic is cooked/processed (Tocmo et al. 2017). To evaluate the effects of processing on polysulphides, GC-MS and LC-MS analyses of garlic extracts revealed that shorter boiling times enrich linear polysulphides,

especially trisulphides and allyl disulphides (next to garlic aroma, this confers meaty nuances to the taste of food already at 2 ppm; Burdock 2009).

There is currently limited metabolomics literature on flavours resulting from the processing of yeast derived products. Yeast extracts were studied for the first time by Izzo and Ho (1991) and Ames and Elmore (1992), and later in more detail (Ames 1994; Münch et al. 1997). Sucan and Weerasinghe (2005) provide an overview of process and reaction flavours. Recently, Zhang et al. (2017) reported the analysis of 'yeasty' off-flavour volatiles from yeast extracts used as food ingredients. Volatiles have been the most studied flavour compounds although metabolomics has yet been poorly exploited for yeast product analyses.

2.1 Two approaches: Untargeted and targeted metabolomics

Metabolomics analyses have already proved valuable to the food industry for the analysis of the aroma of fresh (e.g. tomato (Thissen et al. 2011), melon (Allwood et al. 2014); and processed, wine (Cozzolino 2016), or vegetable puree (Lopez-Sanchez et al. 2015) materials. Metabolomics analysis can follow targeted or untargeted approaches (Patti et al. 2012). In untargeted analysis, the aim is to detect as many components from the matrix as possible in an unbiased manner. An untargeted approach is chosen to evaluate the overall metabolite profile of the studied system, without anticipating which (classes of) compounds are responsible for differences in the metabolic profiles. Technically, this is achieved by metabolic fingerprinting or profiling approaches (Hall 2006). A typical objective in metabolomics studies is the detection of biomarkers. Biomarkers in the food context are compounds that indicate a certain state or the perturbation of a metabolic system either by their presence or abundance change. Often, a combination of untargeted analysis followed by one or several targeted analyses is needed in order to capture all the information (Esslinger et al. 2014).

Targeted analyses rely on *a priori* knowledge of the class of metabolites that are expected to contribute to the (sensory) properties of interest (Scalbert et al. 2009). However, food matrices are highly complex involving compounds with very different physical and chemical properties. Appropriate sample preparation methods and accounting for the influence of matrix effects are essential during data analysis and interpretation. Extraction of components has been traditionally done by using universal solvents (Patti et al. 2012). Solventless extraction of volatiles such as headspace techniques is a fast, sensitive and economical alternative (Kataoka et al. 2000). Metabolites can also be analysed without extraction by using e.g. NMR (Van Duynhoven and Jacobs 2016), direct infusion MS (Baker et al. 2012), PTR-MS (Biasioli et al. 2011), infrared (IR) spectrometry (Aernouts et al. 2011), RAMAN Spectroscopy (Goodacre et al. 2018) and MS imaging (Matros and Mock 2013; Kadam et al. 2016) as well as the emerging SWATH-MS technology (Stolle et al. 2018). There is no singular method that allows for accurate, sensitive and complete reporting of all chemical species in a food sample. However, with new analytical developments, comprehensiveness in coverage continues to increase (Lopez-Sanchez et al. 2015). LC-MS-based methodologies have been proposed to be best suited for the identification of novel bioactive compounds in plant foods because of the compatibility of LC separation with the diversity of metabolites present (Johanningsmeier et al. 2016). GC-MS-based methods have been broadly applied for the analysis of food volatiles and may also be applied to

the study of derivatised, non-volatile polar components such as mono- and disaccharides, sugar alcohols, organic acids, amino acids, and long-chain fatty acids. Headspace techniques are now regularly being used to study the volatile aroma composition of food products. Each technique has its own advantages and limitations and these are highlighted in more detail for specific flavour compound groups in Section 4 (See also **Table 1**). For a more detailed overview of the wide range of methodologies we can refer to the many recent chapters in (Antonio 2018).

2.2 The complexity of food matrices

The analysis of flavour compounds is challenging due to several factors including their high dynamic range, relevant presence at low concentrations (ppm, ppt), high range of polarity, extreme high volatility (high vapour pressures) and the instability of some flavour compounds in dynamic equilibria with other constituents of the food matrix (Dresow and Böhm 2009). Moreover, the complexity of the food matrix provides the greatest challenge. The physical and chemical interactions of all compounds present in a food matrix determine the 'overall sensory experience'. This complexity sometimes incurs limitations to the extraction/separation methods (Scalbert et al. 2009) but since perceived flavour is not determined by a single component, the identification of the overall sensory-relevant metabolite profile, through untargeted analysis is needed to allow us to recognise which metabolite(s) or metabolic pathway(s) correlate with the food processing strategy and final product quality.

Lipids, proteins, and carbohydrates in the food matrix can positively or negatively affect the flavour of food products. The release of aroma compounds from foods is determined by the partition coefficient between the air phase and food matrix (and here also between the hydrophobic and hydrophilic phases). The nature of these interactions depends on the physicochemical properties of flavour compounds and food components (Jeleń et al. 2012). For instance, fat composition, concentration, emulsion characteristics and temperature, can each significantly modify interactions between lipids and small molecules (Piraprez et al. 1998; Guichard 2002). The retention of different aroma compounds in lipid matrices is strongly influenced by their molecular weight and chemical structure (Piraprez et al. 1998). Lipid oxidation produces a variety of aldehydes that can participate in carbonyl-amine condensation and aldol condensations, potentially competing for reactive intermediates with the Strecker aldehydes. In addition, studies have shown that carbohydrates can also influence the retention and release of volatile flavour compounds (Naknean and Meenune 2010). Generally, mono- and disaccharides usually increase the vapour pressure, which causes an increase in volatility of flavour compounds relative to water. Polysaccharides, in contrast, incur a reduction in aroma release caused by an increase in viscosity and/or by molecular interactions with flavour compounds (Naknean and Meenune 2010). Proteins have also been shown to bind covalently to thiols and disulphides (Adams et al. 2001) removing, in this way, potent odorants from the system. In addition, the presence of polyphenols can reduce the emission (or amount) of pyrazines (García-Lomillo et al. 2016).

Studying the behaviour of these flavour-matrix interactions will enable us to understand better the dynamics of the formation and release of flavour compounds in food. Moreover, flavour compound-matrix interaction is not only a crucial step in flavour release research, but also is highly relevant for

metabolomics/analytical method development. In metabolomics studies, the complexity of food matrices has become a crucial aspect for method development, especially for untargeted analyses.

3. Flavour precursors and reaction pathways for aroma-active compound formation

Non-enzymatic reactions occurring in the formation of aroma compounds during food processing include Maillard reactions, caramelization, oxidative and thermal degradation of lipids, as well as degradation of sugars, proteins, ribonucleotides, pigments and vitamins. Once again, the interactions between degradation products can also result in additional chemical reactions. The Maillard reaction, lipid degradation and a combination of both are particularly important for aroma formation in PF.

3.1 Maillard reaction

The Maillard reaction, which occurs between amino compounds and reducing sugars, has been recognised for over 60 years as one of the most important routes to flavour and browning formation in cooked food (Shahidi et al. 2014). Louis-Camille Maillard first reported the reactivity of reducing sugars with peptides in 1912 (Maillard 1912). The Maillard reaction is a complex of hundreds of possible reactions. Even with the simplest sugars and amino acids, hundreds of different volatile and non-volatile compounds can be formed. This extremely complex reaction has been the subject of much research by food scientists seeking to discover new mechanisms of formation and to identify new compounds that provide the desired flavour and colour characteristics of heated foods (Jaeger et al. 2010; Tamanna and Mahmood 2015). The chemistry has been comprehensively reviewed by Ledl and Schleicher (1990) and more recently by Nursten (2005). Many authors also reviewed the importance of flavour formation through Maillard reactions in different processed foods (Manzocco et al. 2001; Newton et al. 2012).

Hodge (1953) divided the chemistry of the browning reaction into three stages, now generally adopted as the three stages of the Maillard reaction. These are: (i) the early stage (sugar-amine condensation, the Amadori rearrangement); (ii) the intermediate stage (sugar breakdown and dehydration, Strecker degradation) and (iii) the final stage (aldol condensation, aldehyde-amine condensation and formation of heterocyclic nitrogen compounds). **Figure 2** shows a schematic overview based on Hodge (1953) and van Boekel (2006). Two reactions occur during the early stage, the condensation of an aldose sugar and an amino compound (N-glycosylamine formation), and the rearrangement reaction leading to the Amadori compound (the N-substituted 1-amino-2-deoxy-2-ketose) or the Heyns compound if the reducing sugar is a ketose. While the first reaction is reversible, the second is not and is thought to be acid-catalysed. There is no formation of any aroma or colour at this stage. In the intermediate stage, the relatively stable Amadori/Heyns compound can react via two enolisation routes depending on the pH conditions. Sugar fragmentation and release of the amino group occurs, and Strecker aldehydes, among others, are formed (Rizzi 2008; **Figure 2**). These corresponding Strecker aldehydes contain one less carbon atom than the original amino acid (Whitfield 1992) and can be colourless or yellow. They are considered important contributors to the aroma of food products. Many patents have been granted which involve Strecker degradation to produce flavouring materials of foodstuffs such as maple syrup, chocolate, coffee, tea, honey,

mushrooms and bread (Morton et al. 1960). The final stage of the Maillard reaction is comprised of many complex and interconnected reactions of dehydration, fragmentation, cyclization, condensation and polymerisation, in which amino groups again participate. Strecker aldehydes formed from the previous stages can react with each other by an aldol condensation, or they can react with amines at high temperatures to give 'polymeric', high molecular mass, coloured products of generally unknown structure, called melanoidins. Heterocyclic ring systems, such as pyridines, pyrazines, pyrroles, and imidazoles, have also been shown to be present in food materials after these reactions (Nursten 2005). However, little is known about their mechanisms of formation.

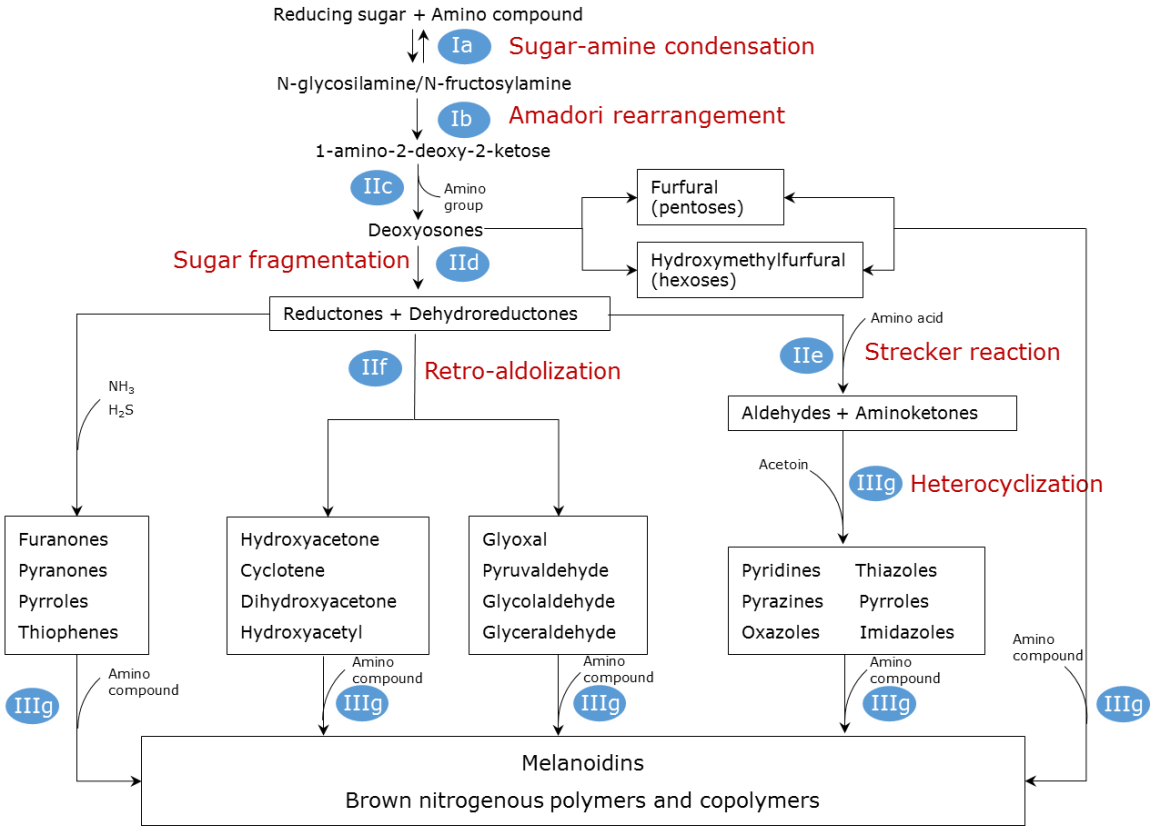


Figure 2. One of the most important sources of typical food metabolites which are of great influence to food flavour and quality arise through usually heat-induced chemical reactions generally grouped under the term 'Maillard reactions'. Here we present a schematic overview of the Maillard reaction, based on Hodge (1953) and Van Boekel (2006), that shows the main-end products contributing to flavour. (I) Early stage (a: Sugar-amine condensation; b: Amadori rearrangement); (II) Intermediate stage (c: Sugar dehydration; d: Sugar fragmentation; e: Strecker degradation/amino acid degradation); (III) Final stage (f: Aldol condensation; g: Aldehyde-amine condensation and formation of melanoidins).

The control of Maillard products during food processing is essential to prevent the formation of undesired products (like carcinogenic compounds and off-flavours) and to facilitate the production of savoury compounds. For this, we first need to learn more about the chemical reactions and mechanisms during processing conditions. Metabolomics now represents a new approach to help correlate pathway analysis, non-enzymatic conversions and food processing steps (Klevorn and Dean 2018).

3.2 Lipid oxidation

Importantly, lipids influence the aroma and flavour of other components, are precursors of odour and flavour compounds and many even have odours and flavours themselves (Forss 1973). Lipids are generally associated with more negative qualities of food flavour, as they are responsible for rancidity in oils or lipid-containing foods. However, they can also play a positive role as flavour enhancers depending on product characteristics. For example, short chain fatty acids are mainly responsible for rancid flavours in milk whereas the same acids are essential flavour constituents in cheese. Moreover, lipids also play an important role in food texture (i.e. mouth feel) and thus affect consumer attractiveness. During food processing, non-enzymatic (auto)oxidation of lipids may occur, and degradation pathways will lead to the formation of a great number of secondary aroma-related metabolites that can affect flavour.

Understanding the mechanisms underlying thermal lipid processing needs more attention. Most lipids are hydrophobic, non-polar compounds. Phospholipids and triglycerides are disassembled during heating releasing short-chain fatty acids with reduced saturation. At elevated temperatures, autoxidation of fatty acids occurs and hydroperoxides are produced. This process involves a free radical mechanism which can be divided into three stages: initiation, propagation and termination (**Figure 3A**). The initiation reaction is activated by direct thermal dissociation, metal catalysis or exposure to light, forming hydroperoxides (Frankel 1980) which are then decomposed via many routes leading to a broad variety of volatile and non-volatile secondary products. A better understanding of hydroperoxide formation and rearrangement is needed. Advanced metabolomics techniques, such as GC-MS and LC-MS are now being used for the characterization of lipid oxidation products (Xia and Budge 2017). GC-MS analysis provides information regarding the structures of individual oxygenated fatty acids, typically as methyl esters, isolated from oxygenated triacylglycerols (TAGs), while LC-MS techniques allow analysis of intact oxygenated TAGs and yields information on the position of the oxygenated acyl chain on the glycerol backbone (Xia and Budge 2017).

Hydroperoxide decomposition forms an alkoxy radical which mostly transforms into aldehydes, ketones, alcohols and furans. Which products are formed depends on the fatty acids present, the hydroperoxide isomers formed, and the stability of the decomposition products. However, the formation of hydroperoxides is not the only oxidation mechanism involved. According to Schaich (2012), alternative pathways to the hydroperoxide can occur from competing reaction cycles to form peroxides. These peroxides can then either re-enter the traditional propagation stage or undergo alternate reaction pathways thus increasing the complexity of both the kinetics and the product mixture. Most of the research related to food science has been focused on the autoxidation of the most common relevant acids; oleic acid, linoleic acid and linolenic acid. Frankel (1980) and Ho and Chen (1994) list the expected decomposition products from linoleate and linolenate hydroperoxides (**Figure 3B**). The rate of autoxidation increases with the degree of unsaturation. Different mixtures of hydroperoxides and their derivatives formed from omega-3 lipids were characterised using SPME-GC-MS of oxidized and non-oxidized flaxseed oils (Nieva-Echevarría et al. 2017). An overview of the lipid degradation pathways that lead to flavour formation is shown in **Figure 3A**. All of these products

are of broad importance in (off)flavour determination – for example, 2,4-Decadienal is known to be one of the most important flavour contributors to deep-fat fried foods.

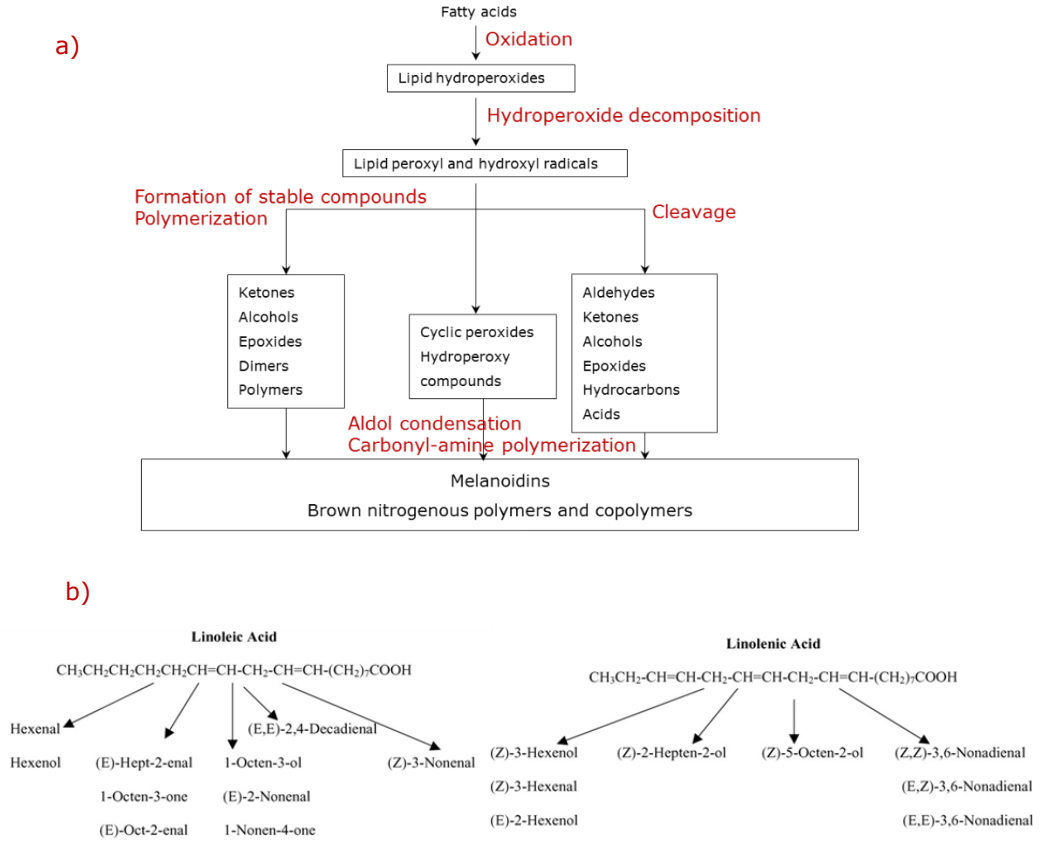


Figure 3. Lipids and their breakdown products are often hugely influential to the overall quality and flavour of food. These compounds can have both a positive or a negative influence on (off-) flavour. Here we present an overview of the lipid degradation pathway: **a)** provides an overall picture of the different fatty acid degradation routes and **b)** illustrates the complexity using two specific examples of important common fatty acids and their diversity in breakdown products of sensory relevance

3.3 Lipid-Maillard interactions

Both lipid degradation and Maillard reactions lead to the formation of a great number of compounds with a similar range of physicochemical properties. It is expected, that these compounds (both intermediate and end products) could further interact to create new volatile compounds and/or block some of the products found in one reaction process by the presence of products from another (Kerth and Miller 2015).

Lipid-Maillard interaction products have been mostly identified in cooked meat, French-fries, peanuts and beverages such as coffee, tea and cocoa. From a review on this topic the largest number of these compounds was found in French-fried potatoes (Whitfield 1992). These were mostly characterized as heterocyclic compounds containing one or more atoms of nitrogen or sulphur, with the presence of long-chain alkyl groups with four or more carbon atoms. Examples of such volatile groups are pyridines, pyrazines, thiophenes, thiazoles and oxazoles with alkyl side chains. **Figure 4** shows a few common volatile products formed in this way.

Shahidi et al. (2014) investigated the reactions occurring during the thermal processing of meat. They described how saturated and unsaturated aldehydes, formed via autoxidation of lipids, interact with products from the initial and later stages of the Maillard reaction. Several thiazoles with four-to-eight carbon alkyl moieties at the 2 position have been reported in roast beef along with other alkyl thiazoles. Using GC-MS, longer alkyl chains were identified in the volatiles of heated beef (Elmore et al. 2002). In general, volatile compounds formed from Lipid-Maillard interactions possess weak odour intensities and high odour thresholds. These contribute to the overall aroma at a lower level compared to compounds generated in the primary reactions. The strongest effect on the flavour profile when Lipid-Maillard interactions are present is due to the variation on the Maillard reaction products as affected by lipids (indirect impact on the aroma profile). In particular, Shahidi et al. (2014) reported that phospholipids and their degradation products inhibit important reactions involved in the formation of heterocyclic aroma compounds in the Maillard reaction. Therefore, in the case of cooked food, this interaction may help to maintain the concentrations of sulphur compounds at an optimum level. On the other hand, Elmore et al. (2002), using meat-like model systems, were able to detect the formation of alcohols and alkylfurans instead of the saturated and unsaturated aldehydes when polyunsaturated fatty acids were present. Possible pathways for the formation of those compounds have been proposed. However, many gaps in our knowledge are still present.

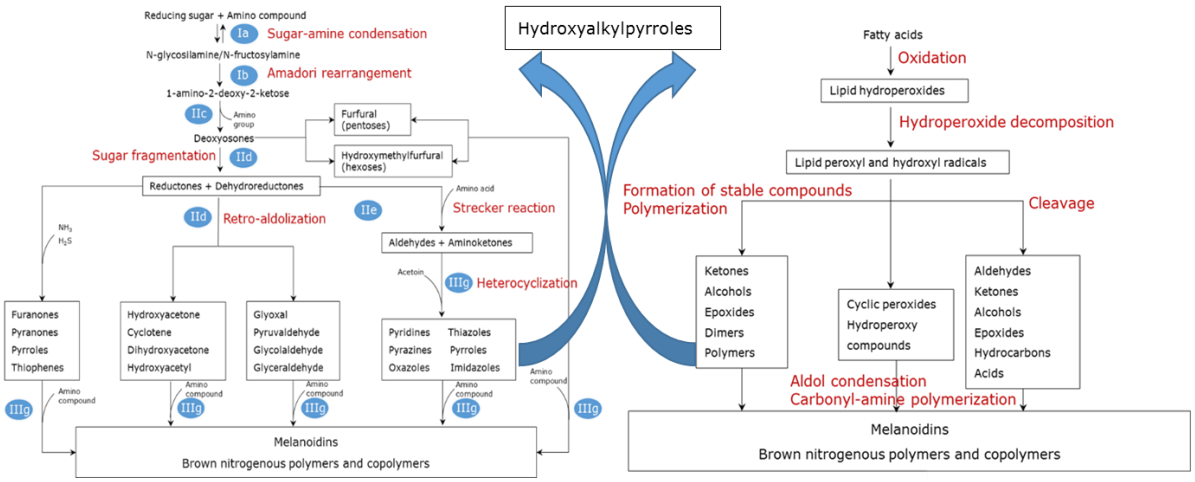


Figure 4. Foods generally have a more complex chemical composition that the fresh materials used for their production. Many new compounds are formed during the main processing steps involving non-enzymic reactions. Here as an example, we show an overview of the common volatile products formed from the interaction of Maillard reaction and lipid autoxidation (modified from Whitfield 1992).

3.4 Others

Apart from the great proportion of aroma compounds derived from sugars, amino acids and lipids, there are also other essential compounds found in processed food which can play an important role in flavour formation. Thiamine (Vitamin B1), ascorbic acid (Vitamin C) and carotenoids are examples of such compounds. During thermal processing, these compounds can undergo degradation processes and contribute to the formation of odour-active compounds. Thermal breakdown of these products produce

reactive intermediates that are common to the Maillard reaction. Therefore, thiamine degradation leads to the formation of S-containing heterocycles, thiols, sulphides and disulphides that contribute to a meaty flavour in cooked products (Khan et al. 2015). A schematic compilation of the degradation pathways that lead to aroma formation, from the primary precursors, is illustrated in **Figure 5**. Recently, Yu and Zhang (2010) described the volatiles generated when ascorbic acid and cysteine were heated in aqueous buffer at different pHs. Many of these were compounds that can contribute to meat flavour such as thiophenes, thiazoles, pyrazines and cyclic sulphur compounds. However, using dynamic headspace (DHS) extraction rather than SPME, Parker (2015) showed that 2-methyl-3-furanthiol and many related disulphides were formed in buffered model systems containing ascorbic acid and cysteine.

Another important reaction in processed food is caramelisation. This thermal degradation of sugars occurs in the absence of amino acids, and the products that are formed are similar to those of the Maillard reaction (**Figure 5**). However, when an amino group is present it acts as a catalyst resulting in a faster reaction and higher amounts of very reactive intermediate products (Van Boekel 2006).

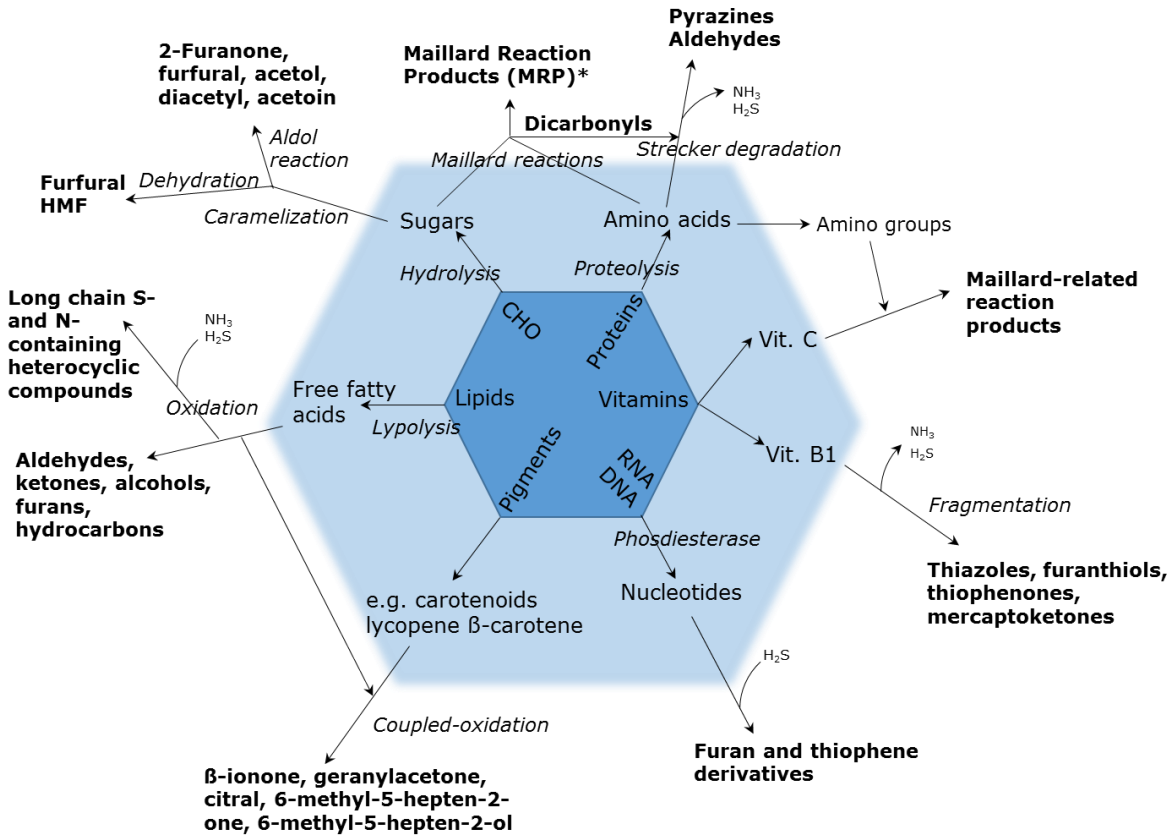


Figure 5. An overview of the chemical reactions relevant to flavour formation, modified from Sucan and Weerasinghe (2005). The typical complexity of the volatile components in food products is the result of both the modification of the original chemical constituents of the raw materials during processing as well as the subsequent interactions within and between the different chemical groups catalysed by enzymatic and thermo/chemical reactions.

4. Volatiles contributing to flavour

4.1 Aldehydes, alcohols and ketones

Aldehydes, alcohols and ketones are groups of chemical compounds that play a key role in the overall flavour of processed foods and e.g. are responsible for the characteristic aroma of fermented foods (Visessanguan et al. 2006; Gambacorta et al. 2009). Moreover, many aldehydes are also naturally present in fresh foods, such as fruits and vegetables, as well as in essential oils. Concentrations increase after thermal processing due to chemical reactions at high temperatures. They can also be formed by oxidation of unsaturated fatty acids, by normal fatty acid metabolism or by the conversion of amino acids (**Figures 2-4**). These molecules cover a broad range of different chemical and physical properties and hence, different analytical approaches are needed. Smaller aldehydes, such as formaldehyde and acetaldehyde, are more soluble in water. GC-MS and LC-MS approaches have been developed to identify short-chain aldehydes. One of the earliest detection techniques for fatty aldehydes was their analysis by GC as dimethylacetal derivatives (Berdyshev 2011). However, headspace extractions are now commonly used. On the other hand, LC-MS approaches have the advantage of the higher stability of hydroxy-aldehydes as compared with the higher temperatures used in GC-MS.

Soy sauce aroma is commonly used in food processing as a flavouring. Its main flavour attributes are based on Maillard and fermentation processes. Aldehydes, alcohols and ketones were the main volatile compounds found in soy sauce products (Fan et al. 2011; Zheng et al. 2013). Several methods have been used for the extraction of aldehydes, alcohols and ketones in soy sauce, such as liquid–liquid extraction (LLE) (Xu et al. 2007) and solid-phase microextraction (SPME) (Gao et al. 2010; Chen et al. 2015). More recently, stir bar sorptive extraction (SBSE) has been tested in Chinese soy sauce aroma type liquor for the analysis of volatile compounds (Fan et al. 2011). SBSE is an environmentally-friendly alternative characterized by its ease of use, high selectivity, high sensitivity, and reproducibility. Fan et al. (2011) also optimized the SBSE method for the analysis of the overall aroma volatile profile of soy sauce liquors, including esters, alcohols, aldehydes and ketones, aromatic compounds, furans, nitrogen and sulphide-containing compounds, acids, phenols and terpenes. **Table 1** shows a summary of some characteristics from the different chemical groups explained throughout Section 4.

4.2 Organic acids and derived esters

Volatile esters are commonly found in many food products. Aliphatic esters are very important constituents of many PF as they give very intense fruity notes. Esters are often characterized as contributing ‘banana flavour’ (isoamyl acetate) or ‘apple flavour’ (ethyl hexanoate, ethyl octanoate). Ethyl acetate is often characterized by an ‘adhesive’ flavour note. Ethyl butyrate, valerate, hexanoate and octanoate are all considered to have strong flavours.

Esters are formed from lipid metabolism (**Figure 3**). Lipids are converted to a large number of alcohols and acids that may then undergo esterification. Esters are also present in process PF, but at low concentrations and thus have little effect on the whole aroma. Lin et al. (2014) identified two esters (acetic acid ethenyl

ester and (S)-2-hydroxy propanoic acid ethyl ester) in yeast extract pastes using simultaneous distillation extraction, and detected by combined GC-O/MS.

Table 1. Examples of aroma/flavour related compounds analysed with different sample preparation techniques in combination with GC-MS during food processing. DHS, Dynamic headspace; LLE, Liquid-liquid extraction; SAFE, Solvent-assisted flavour evaporation; SBSE, Stir-bar sorptive extraction; SD, Steam distillation; SDE, Simultaneous distillation and extraction; SPME, Solid-phase microextraction; SPTE, Solid-phase trapping solvent extraction.

Volatile chemical group		Main sensory attributes	Formation pathways	Extraction technique	References
Aldehydes, alcohols and ketones		Fatty Herbal Fruity Nutty Earthy	-Fatty acid oxidation -Thermal degradation of carbohydrates -Thermal degradation of amino acids -Oxidation of carotenoids	SPME, SAFE	Xu et al. 2007; Gao et al. 2010; Chen et al. 2015
				LLE	Zheng et al. 2013
				SBSE	Fan et al. 2011
				Dimethylacetal derivatives	Berdyshev 2011
Organic acids and derived esters		Fruity Sweet-like	-Lipid metabolism	SDE	Lin et al. 2014
N- and O-containing heterocyclic compounds	Furan-derivatives	Caramel-like Fatty Nutty	-Thermal degradation of carbohydrates -Thermal degradation of amino acids -Maillard reaction -Oxidation of polyunsaturated fatty acids -Oxidation of carotenoids -Oxidation of Vitamin C	SPME	Feng et al. 2015; Seok et al. 2015
				AEDA	Kaneko et al. 2013; Poehlmann and Schieberle 2013
	Pyrazines	Roasted Nutty Cocoa Sweet	-Thermal degradation of amino acids -Maillard reaction	DHS, SDE	Lin et al. 2014
				SPME	Adams et al. 2008
S-containing compounds	Alkyl sulphides and polysulphides	Cooked meat aroma Alliaceous Off-flavour	-Maillard reaction -Thermal degradation of Vitamin B1	DHS	Colina-Coca et al. 2013
				SD, SDE, SPTE, SPME	Lee et al. 2003; Murray 2001
				LLE	Tocmo et al. 2017
	Thiophenes	Meaty aroma	-Maillard reaction -Thermal degradation of Vitamin B1 Thermal degradation of nucleotides	DHS, SDE	Lin et al. 2014; Mahadevan and Farmer 2006
	Thiazoles and thiazolines	Roasted	-Interactions between Maillard reaction products and lipid-derived aldehydes	DHS	Elmore et al. 1997; Shahidi et al. 2014

4.3 Nitrogen- and Oxygen-containing heterocyclic compounds

Heterocyclic compounds are found at relatively low levels in food. However, they do add a significant degree of complexity to food and, thus, they boost the overall flavour of a product and increase its desirability to consumers. Heterocyclic compounds are strongly related to roast meat flavour formation during heating and are important compounds in food processing and flavouring. They have been observed in reaction mixtures of amino acids and sugars but some can also be formed from lipids (Whitfield 1992). Furans, pyrroles, pyridines, pyrazines, oxazoles and oxazolines are typical volatile N- and O-containing heterocyclic compounds which are mainly derived from the reaction between cysteine and xylose (Cao et al. 2017). Apart from sulphur-containing compounds, heterocycles such as methylfuran, pyrazine, and furfural are recognized as potent meaty flavour compounds. Oxazoles, oxazolines, pyrroles and pyridines have also been identified in cooked meats (Devine & Dikeman 2014). However, their contribution to the overall aroma is not as significant as that of sulphur-containing compounds which possess closely related chemical structures, such as thiazoles and thiazolines (See next section).

4.3.1 Furan-derivatives

Furan-derivatives have a planar enol-carbonyl structure in a cyclic dicarbonyl compound. They originate from the early stages of the Maillard reaction and deliver caramel-like aromas (F. Shahidi et al. 2014). These compounds are also important intermediates in the formation of other N- and S-containing flavour volatiles as they can easily exchange oxygen with nitrogen and sulphur. Examples of furan-derivatives are illustrated in **Figure 4**. There are multiple pathways underlying furan and furan-derivative formation. According to Seok et al. (2015), furan is formed from (i) thermal degradation or rearrangement of carbohydrates alone, or in the presence of amino acids, (ii) thermal degradation of certain amino acids, (iii) oxidation of ascorbic acid under high temperatures, and/or (iv) oxidation of polyunsaturated fatty acids and carotenoids (See **Figures 2-4**).

Furan-derivatives are important flavouring compounds and their characterization and identification by analytical methods is of great relevance for designing and producing high-quality food products. An important food ingredient used as seasoning and widely consumed around the world is soy sauce. Over 300 volatile compounds have already been identified in soy sauce (Kaneko et al. 2012) with furan-derivatives being the most abundant and sensory-relevant compounds. Kaneko et al. (2013) investigated the volatiles of soy sauce and how these change on heating. Aroma extraction dilution analysis (AEDA) was used. This is the most frequently applied method for the screening of flavour-impact compounds when using GC-O. During AEDA, the original flavour extract is sequentially diluted and the diluted extracts are then evaluated by GC–O to provide flavour dilution (FD) factors. These dilutions are usually combined with extraction techniques using solvents, such as liquid–liquid extraction (LLE), simultaneous distillation/extraction (SDE), or solvent-assisted flavour evaporation (SAFE) (Curioni and Bosset 2002; Poehlmann and Schieberle 2013; Munafo et al. 2014) and recently with SPME (Feng, Su, et al. 2015). Several parameters, including fibre type, extraction (exposure) time/temperature, and the saturation of the aqueous phase in the headspace vial, affect optimal SPME conditions. Some studies focused on the selection of optimal fibers for the analysis of furan (Seok et al. 2015) and report the best option is a carboxen/polydimethyl-siloxane (CAR/PDMS) fiber, as this showed marked advantages such as selectivity.

4.3.2 Pyrazines

Pyrazines represent an important component of PF and constitute a major class of volatiles formed via the Maillard reaction, conferring a 'roasted flavour' character. Currently, pyrazines are included in the list of flavouring agents authorized by the European Union which may be incorporated into food products to imitate meat flavours (García-Lomillo et al. 2016). However, the majority of the pyrazines in food are naturally occurring and often have a very low odour threshold. The common structure of pyrazines is a six-membered aromatic ring with two nitrogens in *para* position. They differ in the substituents at one or more of the four ring carbon atoms. Alkyl and alkoxy pyrazines are the most predominant forms used in PF. A few examples are illustrated in **Figure 5**. Alkyl pyrazines can help enhance the flavour of cooked foods by adding a savoury taste to the overall product. They are characterized by a roasted, nutty, cocoa, sweet flavour. Alkoxy pyrazines are also commonly used to add a characteristic savoury, and sometimes, spicy aroma to a

number of different products. Müller and Rappert (2010) describe the occurrence, formation and biodegradation of pyrazines as used in flavourings. Pyrazines are routinely analysed by headspace GC-MS because of their volatility. However, organic extractions are also used (Lin et al. 2014). They compared the analysis of the volatile content of two different yeast extract pastes by using dynamic headspace and simultaneous distillation extraction (SDE). SDE extracted and detected more pyrazines than DHS and is hence to be preferred. Adams et al. (2008) used SPME-GC-MS for the analysis of different pyrazines formed from several model reactions and also discovered the formation of a novel pyrrole in the model reaction of alanine with 2-oxopropanal.

4.4 Sulphur-containing compounds

Sulphur-containing compounds are among the most important aroma volatiles in many processed foods (Landaud et al. 2008; McGorrin 2011). Both aliphatic and heterocyclic sulphur volatiles are present in food at low concentrations, but their low odour thresholds make them potent aroma compounds conferring sulphurous, onion-like and meaty aromas to foods. Sulphur compounds are volatile and many GC-MS techniques have been used for the detection of these compounds, mostly following solvent extraction. However headspace techniques have also been developed for the qualitative and quantitative analysis of these compounds as shown below.

4.4.1 Alkyl sulphides and polysulphides

Allium vegetables and their organosulphur compounds are possible cancer-preventative agents in humans as they appear to inhibit the formation of a range of different cancer types (Omar and Al-Wabel 2010). Alkyl sulphides and polysulphides are important compounds for the aroma of many flavouring agents. Polysulphides are formed from the thermal degradation of thiamine and components of the methionine and cysteine pathways via the Maillard reaction (**Figure 2**). They contribute to a cooked meat aroma but, at high concentrations they can give food an off-flavour. Many studies have identified these compounds in garlic and onion samples (Jones et al. 2004; Tocmo et al. 2017). They are created by the degradation of relatively stable, odourless, S-alk(en)yl cysteine sulphoxide flavour precursors and these polysulphides are known to be unstable, depending on the extraction technique. Lee et al. (2003) compared several sampling techniques for the determination of Korean cut garlic flavour components by GC-MS including steam distillation (SD), simultaneous distillation and solvent extraction (SDE), solid-phase trapping solvent extraction (SPTE), and headspace solid-phase microextraction (HS-SPME). Thermal degradation of components such as allyl methyl sulphide, dimethyl disulphide, and thiirane were observed for SDE and SD but not for SPTE or HS-SPME. HS-SPME had several advantages compared with SD, SDE, and SPTE such as rapidity, no apparent thermal degradation, is less labour intensive, and small sample size (Lee et al. 2003). Five different fibre coatings were also evaluated for HS-SPME of garlic flavour components. DVB/CAR/PDMS was the most efficient of the five types investigated. However, some researchers report limitations on the use of SPME if quantification is desired (Murray 2001).

Many vegetables, when cooked, are used as natural flavourings in a wide range of dishes. Garlic (and other species from the genus *Allium*) and ginger are routinely used in stir-fried Chinese dishes. The most important precursor of garlic flavour is allicin (allyl 2-propenethiosulfinate). Allicin is very unstable and readily degrades into other secondary sulphur compounds and a variety of sulphides, when garlic is crushed and then heated. The sulphur compounds thus formed contribute to the pungent flavour of cooked garlic. It was found that in a GC column, allicin will decompose into 3-vinyl-[4H]-1,2-dithiin, 2-vinyl-[4H]-1,3-dithiin, and a few trace compounds (Ho and Chen 1994). The major compounds of heated garlic oil have been identified and are likely to be sensory-relevant. On the other hand, Colina-Coca et al. (2013) reported the analysis of many alkyl sulphides, disulphides and trisulphides by dynamic headspace (DHS) and GC-MS in processed onion. They found differences in the aroma and volatile content of the samples according to the way the onion was processed (raw, high pressure processing). Tocmo et al. (2017) also reported the analysis and identification of alkyl polysulphides by organic extractions.

4.4.2 Thiophenes

Thiophenes are important Maillard products which confer a desirable 'meaty' aroma to food. Thiophenes with a thiol group at the 3-position possess a strong meaty-like aroma and have low odour threshold values. There are a number of possible routes to the formation of thiophenes, involving the reaction of a sulphur-containing amino acid (e.g., cysteine, cystine, methionine) or thiamine, with intermediary sugar degradation products coming from the Maillard reaction (See **Figure 2** and **4**). Lin et al. (2014) reported the detection of 5-methyl-2-thiophenecarboxaldehyde only detected by dynamic headspace/GC-MS as compared to simultaneous distillation analysis in yeast extracts used as flavourings. Similarly, Mahadevan and Farmer (2006) reported the detection of some thiophenes, aliphatic sulphur compounds, and cyclic polysulphur compounds collected by dynamic headspace concentration of yeast extract pastes. On the other hand, static headspace (such as the traditional SPME), combined with two dimensional gas chromatography allowed the identification of 23 thiophenes in roast beef (Cordero et al. 2015) that could be used as aroma compounds in PF.

4.4.3 Thiazoles and thiazolines

Thiazoles and thiazolines are important compounds in roasted or fried meat and have an important role in flavour chemistry. Similar to pyrazines, their content increases with higher cooking temperatures and they are potent flavour ingredients with low aroma thresholds. Most thiazoles contributing to meat flavour are alkyl substituted, and their specific aroma depends on the nature and the number of alkyl moieties attached. One possible route for the formation of thiazoles and thiazolines is via the Maillard reaction involving the action of ammonia and hydrogen sulphide in the presence of α-carbonyls, dicarbonyls, or hydroxyketones (**Figure 2**) as derived from the Strecker reaction. However, lipid-derived aldehydes can also participate in this reaction during cooking (**Figure 4**) and long-chain trialkyl thiazoles have been extracted from meat by dynamic headspace methods and identified by GC-MS (Elmore et al. 1997; Shahidi et al. 2014). In PF made from yeast extracts, many thiazoles were detected using the same technique (DHS/GC-MS) (Ames and

Elmore 1992). Yeast is a rich source of thiamine and thus, it can act as precursor for thiazoles and thiazolines.

5. Sensory analyses and prediction models

5.1 Definition of flavour, taste, odour

The definition of flavour is the overall sensation resulting from the impact of food on the chemical sense receptors in the nose and mouth (Dresow and Böhm 2009). Soluble, non-volatile substances released from food stimulate taste receptor cells, while the volatiles released reach the nose epithelium to give the odour sensation through the so called retro-nasal pathway (Guichard and Salles 2016). The combination of taste and odour is termed flavour. However, the most important contributor to flavour is odour. This becomes evident when a person catches a cold and cannot sense flavour by the nose. Only 200-400 volatiles among 10.000 identified volatile compounds in food are proposed to determine the characteristic odour of a food product (Dunkel et al. 2014). Analyses demonstrated that some very important aromas are not the result of the presence of a unique characterizing compound, but rather result from a reproducible blend of a particular number of components in a specific balance (Dresow and Böhm 2009).

5.2 Sensory analysis and prediction models in food processing

Sugars contribute to sweet taste, and acids to sour taste exclusively, the amino acids and simple peptides can elicit all five primary taste sensations (Sucan and Weerasinghe 2005). Sensory analysis is a scientific discipline that evaluates consumer products by using human senses (sight, smell, taste, touch and hearing) in a complete experimental design (Lawless and Heymann 2010). Panels of trained human assessors are required and, by applying statistical analysis to the results it is possible to make inferences and insights about the products under test.

When designing a food product, it is important to relate the chemical composition with sensory attributes. By knowing the compounds that cause a specific flavour characteristic, we are able to identify the chemical pathway and hence, the flavour precursors that may be 'activated' for product formulation. Thus, a good correlation between the chemical profile and the sensory profile must be determined and hence knowledge of the chemical pathways of flavour formation is needed. While there are some attractive 'proof-of-concept' examples reported for certain (fresh) food products, such chemometric/multivariate modelling approaches have yet to be applied for PF. Prediction modelling of sensory profiles and dynamic modelling of (off-) flavour formation is becoming an important research area in food science. Attempts to relate sensory analysis data to specific chemicals such as volatile compounds have been frequently reported (Lubes and Goodarzi 2017; Seisonen et al. 2016). A non-targeted chemometric approach was used to successfully identify ethyl acetate as a specific off-flavour compound in poor-quality coffee samples (Lindinger et al. 2009). Tikunov et al. (2013) also used a combination of genomic and transcriptomic analysis together with untargeted GC-MS-based metabolomics to identify the gene behind the smoky flavour of tomato. Subsequent blind taste panel and spiking experiments confirmed a causal link between a small number of phenylpropanoid volatiles and the smoky sensory trait. Chemometric analysis has been exploited

frequently in common processed and fresh foodstuffs such as olive oil (Procida et al. 2016) and apple (Aprea et al. 2012; Corollaro et al. 2014). However, these associations can often be weak or difficult to interpret. This may be due to the chemometric method used and/or the natural complexity of food and its flavour. Moreover, current prediction models are also limited to a small number of targeted individual compounds, involve single analytical techniques, or are only focused on specific sensory attributes. As a consequence, they are not robust enough and/or do not allow the generation of novel hypotheses for changing product formulation and manufacturing processes. Expanding the prediction models to a large number of compounds and multiple sensory attributes entails both analytical and statistical challenges.

Chambers and Koppel (2013) reviewed some of the reasons sensory analysis and instrumental measurements often result in poor associations and have identified issues that need to be addressed in future research into understanding the relationships between flavour/aroma phenomena and chemical composition. One difficulty in modelling volatile odour characteristics is that the perception of flavour can vary when volatiles are present in different food matrices, which in itself by definition also remains subjective. For example, hexanal has often been positively associated regarding green/grassy odour. However, negative associations have also been found, with rancidity and oxidized aromatics.

In the case of PF, very little has been done in relation to the volatile chemical composition with the sensory data. Thus, the use of supervised and unsupervised statistical techniques should be put into practice for these more complex flavourings, as was followed by Zielinski et al. 2014.

6. Conclusions

6.1 The importance of (volatile) secondary metabolites in food

Volatiles are only a small proportion of the total number of metabolites in food yet their contribution to food flavour makes them crucial compounds for understanding food science. For example, a rice grain contains 92% starch and 7% protein but the huge difference in market value and consumer preference relates to the remaining 1% including the volatile components (Calingacion et al. 2015). Furthermore cooked food contains many more volatiles than the raw material (Belitz et al. 2008) and it is this complexity which calls for a metabolomics approach to advance our knowledge. Food metabolomics is helping to reveal the mechanisms underlying volatile flavour formation and subsequently the precursor chemical pathways involved. Faster, more sensitive and more comprehensive analysis of food is now possible. Nevertheless, challenges remain – for example, correlating human sensory panel and metabolomic data. In order to innovate, it is important to understand how to make ingredients and food healthier, palatable and more profitable across the entire processing chain. Maintaining positive flavour attributes along the supply chain is one approach to produce better and tastier food.

6.2 The major challenges of food metabolomics

The overall metabolite analysis of the complex matrices typical of processed foods is one of the biggest challenges in food metabolomics. Advances in chromatographic separation and data processing are greatly

advancing our understanding of PF and how these relate to food processing strategies. However, there are still many limitations regarding the separation methods for these complex matrices (Scalbert et al. 2009). Yang et al. (2015) reported a data pre-processing strategy to reduce the masking effect of complex sample matrices. For untargeted analysis, the aim is to detect as many components of matrix as possible. Therefore, metabolite analysis without using solvents results in a much higher (unspecific) coverage of the different chemical groups, as well as a fast, sensitive, solventless and economical alternative. Traditional techniques such as SPME are still being improved thanks to advances in high-throughput instrumentation and data pre-processing (Kataoka et al. 2000). On the other hand, emerging technologies such as SBSE techniques are good alternatives as they require little sample preparation (David and Sandra 2007; Camino-Sánchez et al. 2014; Nogueira 2015). There is no singular method that allows for accurate, sensitive and complete reporting of chemical species but with these new analytical developments increased comprehensiveness is emerging.

Low abundant compounds are difficult to detect. However, in food, these can greatly influence the sensory properties when they also have low aroma thresholds. For instance, volatile sulphur compounds are characterised by their low odour threshold and strong reactivity, making these compounds of great importance to the quality and uniqueness of foodstuffs like cheese, beer, wine and meat (McGorrin 2011). Concentration levels are sometimes below the capacity of the analytical technique indicating that analytical sensitivity can be lower than human detectability. Odour thresholds play an important role in sensory analysis. Some compounds result in intensive aromas even at trace concentrations, while other compounds can change aroma profiles at differing concentrations. Furthermore, food matrices can also influence the perception of individual compounds as compared to the overall flavour of the mixture.

Untargeted food metabolomics is also currently limited by the identification of unknowns. The lack of available libraries with food metabolites makes rapid identification a challenge. In untargeted metabolomics studies, major improvements are still required for spectral deconvolution and automatic metabolite identification. Although major efforts are being made to improve spectral databases, the development of accurate automatic identification algorithms is still subject to the availability of an exhaustive set of reference metabolites (Alonso et al. 2015). The identification of unknowns even present in trace quantities in natural foods is essential for novel processing strategies to deliver superior food flavour and quality. In addition, the identification of new 'natural' flavour compounds in cooked and roasted foods would doubtless be of great interest to the flavour industry.

Chemical and sensory characteristics of soy sauce: A review

Carmen Diez-Simon, Charlotte Eichelsheim, Roland Mumm, Robert D. Hall

Journal of Agricultural and Food Chemistry (2020)
DOI: 10.1021/acs.jafc.0c04274

Abstract

Soy sauce is a fermented product and its flavour is a complex mixture of individual senses which, in combination, create a strong palatable condiment for many Eastern and Western dishes. This review focuses on our existing knowledge of the chemical compounds present in soy sauce and their potential relevance to the flavour profile. Taste is dominated by umami and salty sensations. Free amino acids, nucleotides and small peptides are among the most important taste-active compounds. Aroma is characterized by caramel-like, floral, smoky, malty and cooked potato-like odours. Aroma-active volatiles are chemically diverse including acids, alcohols, aldehydes, esters, furanones, pyrazines and S-compounds. The origin of all compounds relates both to the raw ingredients and starter cultures used as well as the parameters applied during production. We are only just starting to help develop innovative studies where we can combine different analytical platforms and chemometric analysis to link flavour attributes to chemical composition.

Keywords: soy sauce; taste; aroma; liquid chromatography-mass spectrometry (LC-MS); gas chromatography-mass spectrometry (GC-MS); fermented; Maillard reactions

1. Introduction

Soy sauce – also called *jiangyou* in Chinese, *shoyu* in Japanese and *soya sauce* in British English – is the most consumed seasoning in East and Southeast Asian cuisine and is still gaining considerable popularity in Western countries (Gao, Zhang, et al. 2019). Its distinct flavor is characterized by a strong umami, salty and caramel-like character which enhances the overall savory taste and aroma of many kinds of dishes. Soy sauce is used for a wide variety of both cooked (hot and cold) and uncooked foods, such as sushi, sashimi, stir fried noodles and any type of fish, meat and/or vegetable stews. In ancient China, about 2,200 years ago, soy sauce was developed as a way to preserve food – due to its salt content – and also to enhance the flavor of the modest, vegetarian Buddhist diet since salt at that moment was an expensive commodity (Greenberg 2000). Nowadays, soy sauce is used as a food product rather than a food preservative. In fact, its production methods have been developed into advanced technologies and the flavor quality and consistency have improved markedly.

Soy sauce is made using essentially, five basic raw ingredients: soybeans/soybean flakes (as the main protein source), wheat/wheat flour (as the main carbohydrate source), salt, water, and *Aspergillus oryzae*/*Aspergillus sojae*, salt-tolerant yeast and lactic acid bacteria (Devanthi and Gkatzionis 2019; Gao, Zhang, et al. 2019; Lioe et al. 2010). Nonetheless, existing fermentation processes are remarkably diverse and complex, making soy sauce a widely varied product with strong local/cultural heritage. Like balsamic vinegar, soy sauce can be found worldwide, from the cheapest hydrolyzed forms to the most expensive premium, aged brands. There are essentially two general production processes: traditional fermentation and acid hydrolysis, although in China the latter type is since 2018 no longer recognized as an accepted method to produce soy sauce. Acid hydrolysis, also known as “chemical hydrolysis”, uses high concentrations of acid for the initial polymer break down of soybeans and wheat. Traditional fermentation uses starter cultures of microorganisms which secrete enzymes capable of breaking down the proteins, lipids and starches into peptides, free amino acids, volatiles and saccharides (Devanthi and Gkatzionis 2019). Flavor and color development during the production process are attributed to the enzymatic reactions mentioned above, along with additional non-enzymatic Maillard reactions occurring between amino acids and reducing sugars, as well as Strecker degradation of amino acids, lipid oxidation and esterification (Diez-Simon et al. 2019). Maillard reactions, also called browning reactions, occur at higher temperatures and trigger the formation of a wide range of, yet poorly defined, molecules which, in combination, create the characteristic and distinct soy sauce flavor and color.

In this review, we have compiled the existing knowledge of the chemical compounds present in soy sauce and their potential relevance to the flavor profile, differentiating between taste and aroma attributes of the different types of soy sauce currently available on the global market. Some recent seminal papers on advances in our knowledge make this a timely moment to evaluate the literature and discuss future research directions. We also highlight the growing importance of using advanced approaches which move more towards untargeted methods, and which combine analytical platforms and chemometric analysis to link flavor attributes with chemical composition thus bringing us closer to conclusions on causality.

1.1 Different production methods of soy sauce

The production of soy sauce varies greatly in terms of fermentation times, temperatures and ratios of ingredients – soybeans, wheat, salt, water, starter composition, and optionally, other additives. Moreover, the production processes are closely associated with the country of origin. For instance, the Chinese-type uses predominantly more soybeans than wheat (ratios 80:20 and 70:30), whereas the Japanese-type uses equal amounts (ratio 50:50) (Devanthi and Gkatzionis 2019; Wanakhachornkrai and Lertsiri 2003). The presence of wheat is known to influence the final aroma and taste characteristics of the soy sauce. More wheat translates into more carbohydrates, which are degraded to saccharides by the enzymes of the starter microorganisms. These saccharides are the fuel for yeast fermentation, which converts them into various alcohols and other important flavor compounds (Harada et al. 2018). **Table 1** summarizes a list of soy sauces categorized by their origin. In China, soy sauces can be classified into two groups: high-salt liquid-state fermentation soy sauce (HLFSS), considered as the traditional Chinese-type and; low-salt solid-state fermentation soy sauce (LSFSS). LSFSS is made from defatted soybeans and wheat bran, using a shorter ageing period and is therefore cheaper and faster to prepare than HLFSS (Feng, Cai, et al. 2014). For this reason, LSFSS dominates the market of soy sauce in China, although its flavor quality is generally inferior to HLFSS. HLFSS, which is also called Chinese dark soy sauce, is mainly used for slow, long-time cooking, such as stewing. It gives to the dish an attractive caramel color and a slightly sweet undernote. In contrast, LSFSS (Chinese light soy sauce) is used primarily for seasoning light dishes and for dipping, due to its delicate flavor (Chris Liang 2016). In Japan, the most common soy sauces are classified into *koikuchi-shoyu* (Japanese dark soy sauce) *usukuchi-shoyu* (Japanese light soy sauce), *tamari-shoyu* (tamari soy sauce), and *shiro-shoyu* (white soy sauce) (Lioe 2007). Dark-colored Japanese soy sauce (*koikuchi-shoyu*) is produced using the same amounts of soybean and wheat, and is characterized by having a strong aroma and a deep reddish brown color. Light-colored Japanese soy sauce (*usukuchi-shoyu*) is also produced using equal amounts of raw materials, but has a lighter, red-brownish color and milder aroma than dark soy sauce as a result of a more gentle ageing process (Kaneko et al. 2012). White soy sauce (*shiro-shoyu*) is produced using a very high ratio of wheat to soybean and has a light yellow color because it is fermented under conditions that specifically prevent browning reactions (Feng, Su, et al. 2015). Tamari is a type of soy sauce which does not contain wheat (or has less than 10%), resulting in a product with less aroma and a darker color (Kaneko et al. 2012). Tamari is produced mainly in Japan, but also in other countries, although it is less-widely consumed. Other types of soy sauces include sauces with reduced salt or sauces with added caramelized sugar such as *kecap manis* (Indonesian sweet soy sauce), a thick and dark-colored liquid rich in Maillard reaction products (Lioe et al. 2010). Additional soy sauce types which contain other ingredients, such as fermented fish, are not considered in this review.

Table 1. List of the main categories of soy sauces and their description as referred to in this paper.

Origin	Type of soy sauce	Description
Chinese	High-salt liquid-state fermentation soy sauce (HLFSS)	Chinese dark soy sauce
	Low-salt sold-state fermentation soy sauce (LSFSS)	Chinese light soy sauce
Japanese	*Koikuchi-shoyu*	Japanese dark soy sauce
	Usukuchi-shoyu	Japanese light soy sauce
	Shiro-shoyu	White soy sauce (low soybean)
	Tamari-shoyu	Tamari soy sauce (low wheat)
	Saishikomi-shoyu	Double-fermented soy sauce
Indonesian	*Kecap manis*	Indonesian sweet soy sauce
Various	Acid-hydrolyzed	Non-fermented soy sauce

Although traditional Chinese- and Japanese-type soy sauces use different ratios of raw materials and adopt different production conditions, both follow similar steps in their fermentation process. This process can either be based on traditional brewing or acid hydrolysis. A summary of these two methods is schematically shown in **Figure 1**. The two main differences between these processes are the use (or not) of the "starter" used to trigger the degradation of the macromolecules present in the raw ingredients (mainly proteins and carbohydrates, but also lipids and lignin) and secondly the subsequent use of heating treatments and the length of the process (Sano et al. 2007). In traditional fermented soy sauce, a culture containing a complex mixture of different microbial species is added (**Figure 1A**), while in the acid hydrolysis method this is replaced by hydrochloric acid to chemically break down the macromolecules (**Figure 1B**). Some soy sauces found on the market (worldwide) are actually a mixture (blend) of traditional and acid-hydrolyzed soy sauces. This helps to reduce the market price since the latter is much cheaper to produce. The production procedure of traditional soy sauce fermentation, emphasizing the importance of the microbial fermentation steps has been excellently reviewed elsewhere (Devanthi and Gkatzionis 2019). Two steps during the process can be distinguished: koji fermentation (solid-state fermentation, in orange) and moromi fermentation (liquid-state fermentation, in green). Koji fermentation starts with the raw ingredients: soybeans are soaked in water for several hours at ambient temperature to wash out any naturally-occurring fungal inhibitors and impurities, as well as to increase moisture content. Subsequently, they are steamed/autoclaved at high or ambient pressure to kill (or decrease the growth of) any undesired bacteria, yeast and mold. Finally, the steamed soybeans are cooled down to ambient temperature before being mixed with the wheat. This wheat however, is first roasted and ground into flour before the mixing step. The soybeans and wheat flour are then supplemented with 0.05% – 0.3% w/w fungal spores. The most commonly used fungal species are *Aspergillus oryzae*, *A. sojae* and *A. tamarii*. After three days of incubation at ambient temperature, fungal mycelia are already growing on the surface of soybeans. This is called the koji (solid). At this stage, most proteins and carbohydrates have started to break down into peptides, amino acids and saccharides by the enzymes present in the mold.

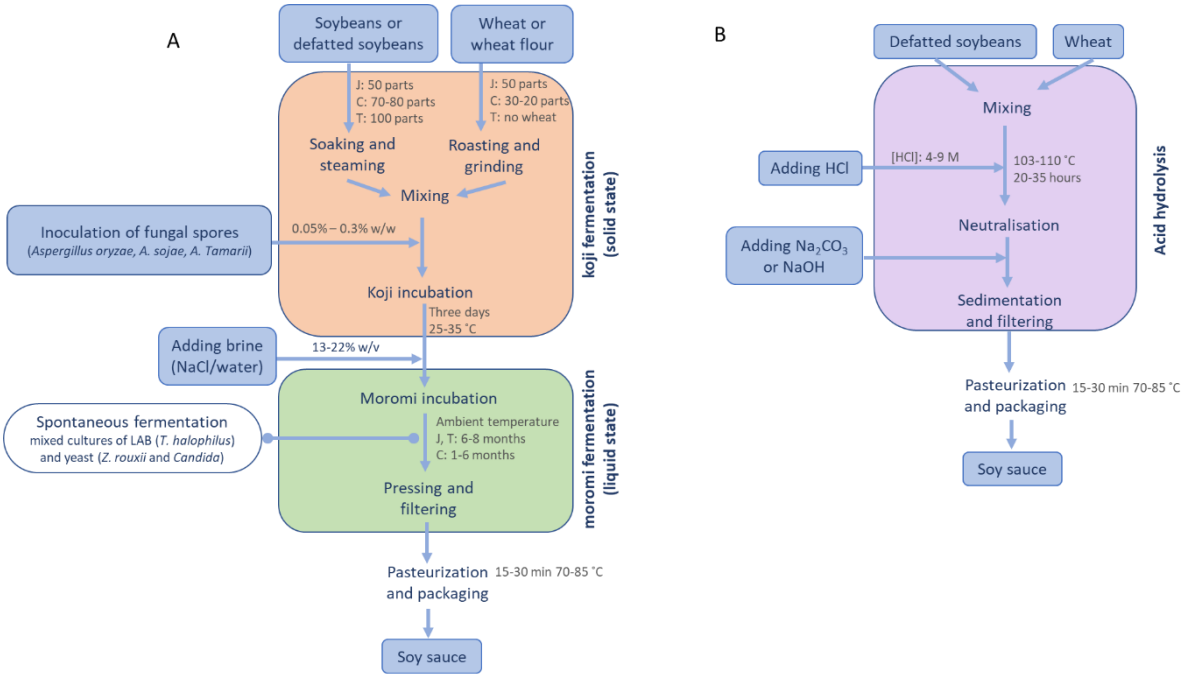


Figure 1. Flow charts of key stages in production of traditionally fermented soy sauce **(A)** with variations: J, Japanese-type; C, Chinese-type; T, Tamari, and acid-hydrolyzed soy sauce **(B)**.

The second step is the moromi fermentation. A brine solution containing 18-22% NaCl is added to the koji. Sometimes, for the premium brands, a raw soy sauce mix is used instead of brine. The resulting product is referred to as double-fermented soy sauce (*Saishikomi-shoyu* in Japan). The high concentration of salt used favors the specific growth of halotolerant microorganisms and prevents (further) growth of the koji molds (*Aspergillus*). At this stage, a second fermentation occurs spontaneously, and only microorganisms tolerant to the levels of salt and pH can grow. The most common microbial cultures found are the indigenous halotolerant lactic acid bacteria (LAB), and certain yeast species (*Z. rouxii* and *Candida spp.*). These microbial cultures use the peptides, amino acids and saccharides formed during the koji stage, as nutrients to grow. Lactic acid fermentation soon decreases the pH to 4-5 which then favors yeast growth, and decreases bacterial growth. These complex microbial cultures play an essential role in the flavor formation during traditional fermentation, influencing thereby final product quality. Consequently, novel developments such as the use of well-defined mixed starter cultures, genetic modification of specific strains, as well as immobilization of cells and enzymes are known to enhance and accelerate flavor formation and give better control of the fermentation process (Devanthi and Gkatzionis 2019; Gao et al. 2020). Traditional fermentation processes can sometimes continue for a long time, from 1 month up to 4 years (Devanthi and Gkatzionis 2019; Wei, Wang, Chen, et al. 2013; Yang et al. 2016). After the second fermentation, the moromi is pressed, filtered, pasteurized, packed and is ready for consumption.

The production of acid-hydrolyzed soy sauce can be performed in just a few days entailing that it is not just a faster but also a cheaper process. This process simply involves first mixing defatted soybeans with wheat gluten and a concentrated solution of hydrochloric acid (4-9 M). The mixture undergoes hydrolysis during prolonged heating (20-35 hours), at temperatures sometimes higher than 100 ^{0}C (Hasnip et al. 2005). It is

subsequently neutralized with Na_2CO_3 or NaOH and impurities are removed by sedimentation and filtration. The raw product is finally pasteurized and packed for consumption (**Figure 1**).

1.2 Compositional differences in different types of soy sauce

Although soy sauce flavor has been studied extensively, only a few authors have compared the flavor (taste and/or aroma) profiles of different types of soy sauces according to the country of origin. One study analyzed the aroma compounds of 27 commercial soy sauces obtained from three different fermentation processes according to origin: Chinese HLFSS, Chinese LSFSS, and Japanese *koikuchi-shoyu* (Feng, Su, et al. 2015). The aim was to identify the effect of the different fermentation processes on the final flavor of the soy sauces. Feng *et al.* found that most volatile compounds were common for all samples. However, their intensities differed, as did the final overall sensory perception. Many esters (e.g. ethyl acetate) and phenols (e.g. guaiacol and 4-ethylguaiacol) were found at relatively high levels in the Japanese soy sauces as compared to the Chinese ones, likely due to the high proportion of wheat used as well as the different parameters applied during fermentation. On the other hand, some acids were only detected in the Chinese LSFSS-type. This may be directly related to the fermentation process where different mold species had been employed. In general, Japanese soy sauces contained higher levels of certain sensory-relevant compounds such as ethanol (alcoholic), 2-phenylethanol (floral, sweet), 4-ethylguaiacol (smoky, bacon), ethyl acetate (fruity) and phenylacetaldehyde (honey-like), whereas Chinese LSFSS scored higher for 3-methylbutanal (malty), methional (cooked potato), dimethyl disulfide (onion or cooked cabbage) and 5-methyl-2-furancarboxaldehyde (almond, spicy, caramel-like). Chinese HLFSS appeared to have a profile somewhere between Chinese LSFSS and Japanese *koikuchi-shoyu*.

With respect to the production process applied (traditional fermentation or acid hydrolysis), soy sauces are also characterized by contrasting flavor profiles, leading to differences in overall taste and aroma. **Figure 2** shows an example of two GC-MS chromatograms of soy sauces: one being traditionally fermented (**Figure 2A**) and an acid-hydrolyzed soy sauce (**Figure 2B**). Both chromatograms have been obtained using the exact same analytical method to enable direct comparison (unpublished data), so the differences observed are due to the production procedure. Both volatile profiles immediately look different which reveals why it is highly relevant to compare both production procedures. One study isolated, analyzed and compared the volatile compounds present in fermented and acid-hydrolyzed soy sauces (Lee et al. 2006). Aroma extract dilution analysis (AEDA) was used to evaluate the odor-active compounds. So far, this appears to be the only article which compares fermented and acid-hydrolyzed soy sauces in terms of their volatile composition. The results have shown that alcohols and esters were dominant in the volatiles of traditionally fermented soy sauces, whereas heterocyclic compounds (e.g. pyrazines and furans) and acids were relatively abundant in the acid-hydrolyzed soy sauces. Odor-active compounds which were seen to be dominant in traditionally fermented soy sauces included furfuryl alcohol (sugar burnt), 2-Ethyl-4-hydroxy-5-methyl-3(2H)furanone (4-HEMF) (caramel/sweet), 2-phenylethanol (floral/sweet), methionol (potato/grassy), 1-octen-3-ol (mushroom), 4-ethylguaiacol (smoky/bacon/soy sauce), and ethanol (alcoholic/solvent-like). In contrast, the odor-active compounds that were more dominant in acid-hydrolyzed soy sauce included guaiacol (burnt woody/medicinal), formic acid (rancid/pungent/metallic), 2,5-dimethylpyrazine (roasted nuts), 2,6-

dimethylpyrazine (nutty), maltol (sweet/caramel), 2,6-dimethoxyphenol (smoky/woody), 2,5-dimethyl-3-ethylpyrazine (potato/woody), and 2-acetyl-5-methylfuran (strong nutty). In the Section "Aroma-active volatile compounds", the potential importance of these odor-active compounds is discussed.

Moreover, a few studies have been performed using targeted LC-MS approaches to determine levels of specific compounds such as levulinic acid and chloropropanols, in acid-hydrolyzed soy sauce. Both (groups of) compounds are considered to be typical markers for acid-hydrolyzed soy sauce. The former is known to be synthesized by the thermal degradation of hexose monosaccharides in the presence of mineral acids such as hydrochloric acid. The latter comprises a group of heat-produced contaminants which are possibly carcinogenic to humans (Hasnip et al. 2005; Sano et al. 2007). These types of targeted analysis are intended to measure defined groups of characterized metabolites, yet hereby will not reveal new compounds that may be specific for the type of soy sauce, and/or influence its final flavor quality.

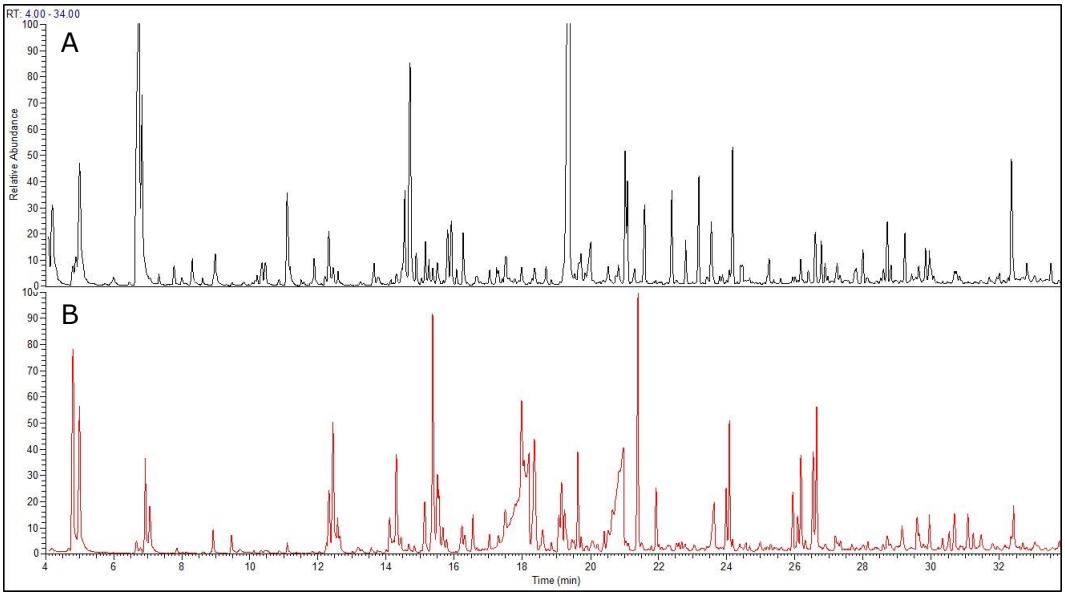


Figure 2. GC-MS volatile profiles of two soy sauce samples extracted by SBSE, being traditionally fermented **(A)** and acid-hydrolyzed soy sauce **(B)**. The untargeted volatile profiles shows quantitative and qualitative differences. (Unpublished data from the authors). The column used was a Zebron ZB-5Msplus with dimensions 30m x 0.25mm x 1.00µm (Phenomenex).

1.3 The flavor of soy sauce and metabolomic analysis

Flavor quality plays an essential role when it comes to consumer acceptance and preference/purchase behavior. Flavor is mainly determined by the taste and aroma characteristics of a product although mouth feel/texture can also play a role. Considering that soy sauce is a complex mixture with a rich flavor, a broad range of flavor attributes have been described in soy sauce. A compilation of the most characteristic attributes that have been described in soy sauce is shown in **Figure 3**. Attributes are represented in a double wheel, separating the taste and aroma characters. Taste attributes have been based on the 5 standard taste categories, but here additionally including *kokumi* and astringent (Nishimura and Kuroda 2019), the former being related to 'mouthfulness' and the latter more to texture/mouth feel. On the other hand, aroma

attributes are much more diverse, and depend often on the chemical groups involved as is regularly described in literature. For illustration, individual examples of the more relevant molecule(s) that impart each of the different attributes are also shown in **Figure 3** (See also **Tables 2** and **3** for more examples). The overall flavor of soy sauce is closely associated with the final chemical composition. Therefore, identifying the compounds that are present in a sample and confirming their sensory relevance is a key step in understanding the link between composition and flavor. Furthermore, being able to chemically identify such compounds will give us immediate insights into the potential flavor formation pathways involved in their appearance and how this relates to the different ingredients and processes employed.

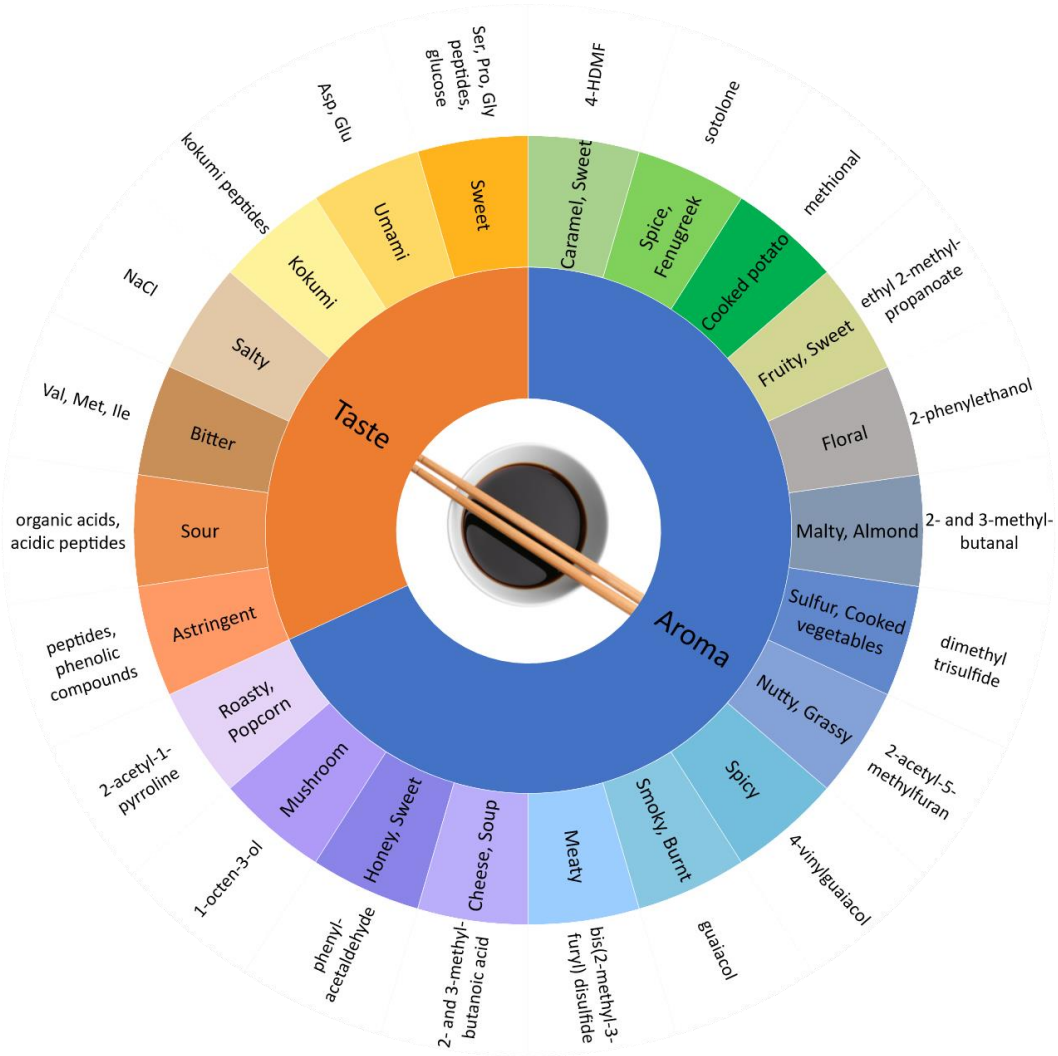


Figure 3. Proposed flavor wheel of soy sauce, constructed on the basis of taste and aroma attributes described in literature (data shown in **Table 2** and **3**). For each attribute, one typical example of a sensory-relevant molecule is also presented. More examples of taste- and aroma-active molecules described in soy sauce can also be found in **Table 2** and **3**, respectively.

Metabolomics is a useful approach to analyze the overall chemical composition of the small molecules that impart the characteristic taste and aroma of plants and food materials (Zhang et al. 2012). Nowadays, metabolomics techniques are based on firstly, chromatographic separation and secondly, metabolite

detection based on mass spectrometry (MS) or NMR. The latter is also particularly valuable for the identification of specific molecules. There is not a single approach that is able to analyze the whole metabolite composition of a product. However, techniques are being developed rapidly in order to increase the coverage of detection, overall sensitivity and reproducibility. Most importantly, these approaches are now being combined with sensory experiments which deliver the flavor perception of the individual compounds detected, and how these correlate with specific attributes. These sensory techniques are mainly based on sensory-driven fractionations, dilution analyses, receptor-based assays, cell-based assays, olfactometry and electronic nose and tongue (Batenburg et al. 2016; Charve et al. 2018; Feng et al. 2017; Riedel et al. 2017; Zou et al. 2018). For soy sauce, very few opportunities have been taken to exploit metabolomics approaches for the analysis either regarding soy sauce flavor or in the development of new high-throughput methods for soy sauce production. Such research would help us generate a clearer picture of sensory-relevant compound precursors and their fate during the processing steps towards the final product. The techniques that have already been used in the detection of taste and aroma compounds in soy sauce are mostly based on gas chromatography mass spectrometry (GC-MS), both for derivatized polar non-volatile compounds (taste) and volatile compounds (aroma) (Sun et al. 2010). Some of these have also been combined with aroma extract dilution analysis (AEDA) and GC-Olfactometry (GC-O) to help develop insights into the potential aroma impact of the compounds detected (Feng et al. 2017). Only occasionally has LC-MS been used, mostly combined with Taste Dilution Analysis (TDA) (Lin et al. 2015), to analyze the semi-polar non-volatile compounds present which are key determinants of taste attributes (Kong et al. 2018). Nevertheless, most studies have focused on more targeted analyses of specific groups of compounds. A broader application for metabolomics approaches in studying the chemical composition without the limitation to focus upon known compounds, has great potential for the analysis of complex mixtures such as soy sauce. This includes the use of advanced strategies for metabolite identification and the use of prediction models to help link chemical composition to sensory relevance and impact (Calingacion et al. 2017; Lubes and Goodarzi 2017; Tikunov et al. 2020). Therefore, soy sauce research may further focus on the optimization and development of untargeted metabolomics techniques as well as to understand the formation pathways and the influence of the different processing steps in the production of soy sauce, within the sensory context. The following sections focus on the non-volatile and volatile compounds characterized in soy sauce, with a particular focus on those influencing taste and aroma properties. The analytical techniques used to characterize the main chemical compound groups are also indicated.

Table 2. (Taste-active) non-volatile compounds reported in different types of soy sauces in literature. Compounds which have been reported to be linked to taste attributes have been considered here including free fatty acids, neutral lipids, phospholipids, polyphenols and others. The type of soy sauce used in the study, as well as the techniques employed are shown.

Compound	Taste attributes	Soy sauce type	Technique	Reference
Free amino acids				
Asp	umami	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
Glu	umami	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
Ser	sweet	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
Pro	sweet	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
Gly	sweet	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
Thr	sweet	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
Ala	sweet	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
Val	bitter	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
Met	bitter	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
Ile	bitter	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
Phe	bitter	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
Lys	bitter	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
Leu	bitter	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
Arg	bitter	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
His	bitter	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
Tyr	bitter	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
Trp	bitter	Chinese / raw soy sauce	HPLC-DAD	(Gao, Zhang, et al. 2019)
Peptide				
Cys-Cys	non-taste	Chinese / HLFSS, LSFSS	HPLC-DAD	(Kong et al. 2018)
γ-Glu-Val-Gly	kokumi	Japanese / *Koikuchi, usukuchi, shiro*	HPLC-MS/MS	(Miyamura and Kuroda 2015)
Synthetic peptides				
ALPEEV	sour, astringent	Chinese / LSFSS	UPLC-MS/MS	(Zhuang et al. 2016)
LPEEV	sour, sweet, umami, astringent	Chinese / LSFSS	UPLC-MS/MS	(Zhuang et al. 2016)
AQALQAQA	sweet, umami, astringent	Chinese / LSFSS	UPLC-MS/MS	(Zhuang et al. 2016)
EQQQQ	sour, salty, umami, astringent	Chinese / LSFSS	UPLC-MS/MS	(Zhuang et al. 2016)
EAGIQ	sour, sweet, salty, astringent	Chinese / LSFSS	UPLC-MS/MS	(Zhuang et al. 2016)
Nucleotides				
5'-GMP	umami enhancement	Chinese / HLFSS, LSFSS	UPLC-DAD	(Kong et al. 2018)
5'-IMP	umami enhancement	Chinese / HLFSS, LSFSS	UPLC-DAD	(Kong et al. 2018)
5'-AMP	umami enhancement (weak)	Chinese / HLFSS, LSFSS	UPLC-DAD	(Kong et al. 2018)
Free fatty acids				
C16:0	not tested	Chinese koji	GC-MS	(Feng, Chen, et al. 2014)
C18:2	not tested	Chinese koji	GC-MS	(Feng, Chen, et al. 2014)
C18:1	not tested	Chinese koji	GC-MS	(Feng, Chen, et al. 2014)
C18:3	not tested	Chinese koji	GC-MS	(Feng, Chen, et al. 2014)
C18:0	not tested	Chinese koji	GC-MS	(Feng, Chen, et al. 2014)
Neutral lipids				
C16:0	not tested	Chinese koji	GC-MS	(Feng, Chen, et al. 2014)

C18:2	not tested	Chinese koji	GC-MS	(Feng, Chen, et al. 2014)
C18:1	not tested	Chinese koji	GC-MS	(Feng, Chen, et al. 2014)
C18:3	not tested	Chinese koji	GC-MS	(Feng, Chen, et al. 2014)
C18:0	not tested	Chinese koji	GC-MS	(Feng, Chen, et al. 2014)
Phospholipids				
C16:0	not tested	Chinese koji	GC-MS	(Feng, Chen, et al. 2014)
C18:2	not tested	Chinese koji	GC-MS	(Feng, Chen, et al. 2014)
C18:1	not tested	Chinese koji	GC-MS	(Feng, Chen, et al. 2014)
C18:3	not tested	Chinese koji	GC-MS	(Feng, Chen, et al. 2014)
C18:0	not tested	Chinese koji	GC-MS	(Feng, Chen, et al. 2014)
Organic acids				
oxalic acid	sour	Chinese / HLFSS, LSFSS	UPLC-DAD	(Kong et al. 2018)
tartaric acid	sour	Chinese / HLFSS, LSFSS	UPLC-DAD	(Kong et al. 2018)
malic acid	sour	Chinese / HLFSS, LSFSS	UPLC-DAD	(Kong et al. 2018)
lactic acid	sour	Chinese / HLFSS, LSFSS	UPLC-DAD	(Kong et al. 2018)
acetic acid	sour	Chinese / HLFSS, LSFSS	UPLC-DAD	(Kong et al. 2018)
pyroglutamic acid	sour	Chinese / HLFSS, LSFSS	UPLC-DAD	(Kong et al. 2018)
succinic acid	sour	Chinese / HLFSS, LSFSS	UPLC-DAD	(Kong et al. 2018)
Polyphenols				
daidzein	not tested	Taiwanese / Chinese	RP-HPLC-UV/DAD	(Gao, Liu, et al. 2019; Lin et al. 2015)
daidzin	not tested	Chinese	RP-HPLC-UV/DAD	(Gao, Liu, et al. 2019)
genistein	not tested	Taiwanese / Chinese	RP-HPLC-UV/DAD	(Gao, Liu, et al. 2019; Lin et al. 2015)
genistin	not tested	Taiwanese / Chinese	RP-HPLC-UV/DAD	(Gao, Liu, et al. 2019; Lin et al. 2015)
glycitein	not tested	Chinese	RP-HPLC-UV/DAD	(Gao, Liu, et al. 2019)
sinapic acid	not tested	Japanese / *Kouikuchi*	RP-HPLC-UV/DAD	(Kaneko et al. 2013)
ferulic acid	not tested	Japanese / *Kouikuchi*	RP-HPLC-UV/DAD	(Kaneko et al. 2013)
vanillic acid	not tested	Japanese / *Kouikuchi*	RP-HPLC-UV/DAD	(Kaneko et al. 2013)
syringic acid	not tested	Japanese / *Kouikuchi*	RP-HPLC-UV/DAD	(Kaneko et al. 2013)
Others				
lovastatin	not tested	Taiwanese	RP-HPLC-UV/DAD	(Lin et al. 2015)
GABA	not tested	Taiwanese	RP-HPLC-UV/DAD	(Lin et al. 2015)
ergothioneine	not tested	Taiwanese	RP-HPLC-UV/DAD	(Lin et al. 2015)

Table 3. Odor-active volatile compounds reported in different types of soy sauces in literature. Only compounds which have been reported to be linked to aroma attributes in literature have been considered here. The type of soy sauce used in the study, as well as the technique employed are shown.

Compound(s)	Odor attributes	Soy sauce type	Technique	Reference
Acids				
2-methylpropanoic acid	cheese-like, fatty	Chinese / LSFSS Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
2-methylbutanoic acid	cheese-like, sweaty	Chinese / LSFSS Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
		Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Steinhaus and Schieberle 2007)
3-methylbutanoic acid	cheese-like, sweaty	Chinese / HLFSS, LSFSS	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015)
		Japanese	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013; Steinhaus and Schieberle 2007)
benzoic acid	fruity, floral	Chinese / HLFSS, LSFSS Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
2-phenylacetic acid	honey-like, sweet, hot chocolate	Japanese	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013; Steinhaus and Schieberle 2007)
butanoic acid	cheese-like	Chinese / HLFSS, LSFSS Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
		Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Steinhaus and Schieberle 2007)
acetic acid	sour	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Steinhaus and Schieberle 2007)
		Chinese / HLFSS, LSFSS Japanese / *Koikuchi* Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
3-methylpentanoic acid	sour herb, slightly green grass	Chinese / LSFSS	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014)
4-methylpentanoic acid	rancid, sour	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
pentanoic acid	sweaty	Japanese / Kikkoman Chinese / raw soy sauce	SAFE-GC-MS/SAFE-AEDA/GC-O SPME-GC-(O)-MS/SPME-AEDA	(Steinhaus and Schieberle 2007)
hexanoic acid	sweaty, pungent	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
nonanoic acid	fatty, green	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
decanoic acid	rancid, fatty	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
phenylacetic acid	sour, honey-like	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
Alcohols				
ethanol	alcoholic, solvent-like	Chinese / HLFSS, LSFSS Japanese / *Koikuchi* Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
2-methyl-1-propanol	bitter, solvent-like	Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
2-phenylethanol	floral, sweet	Chinese / HLFSS, LSFSS Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015)
		Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Steinhaus and Schieberle 2007)
3-methyl-1-butanol	malty, alcoholic	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Steinhaus and Schieberle 2007)
		Chinese / HLFSS, LSFSS Japanese / *Koikuchi*	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)

		Chinese / raw soy sauce		
2-methyl-1-butanol	malty, alcoholic	Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
1-octen-3-ol	mushroom-like	Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
2-ethyl-1-hexanol	rosy	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
2,3-butanediol	fruity	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
Aldehydes				
phenylacetaldehyde	honey-like	Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013; Steinhaus and Schieberle 2007)
		Chinese / HLFSS, LSFSS Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Gao, Liu, Zhang, et al. 2020)
		Japanese and Chinese	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Su, et al. 2015)
benzaldehyde	burnt sugar, caramel-like	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
2-methylbutanal	malty, almond	Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
		Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013; Steinhaus and Schieberle 2007)
3-methylbutanal	malty, almond	Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
		Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013; Steinhaus and Schieberle 2007)
2-methylpropanal	malty, nutty	Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
octanal	green, fruity	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
nonanal	green, fatty	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
trans-4,5-epoxy-(E)-2-decenal	metallic	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013; Steinhaus and Schieberle 2007)
		Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012)
4-hydroxy-3-methoxybenzaldehyde	vanilla-like	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013; Steinhaus and Schieberle 2007)
Esters				
methyl benzoate	floral, honey	Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
ethyl benzoate	fruity, floral	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
ethyl 2-methylbutanoate	fruity	Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013; Steinhaus and Schieberle 2007)
ethyl 2-methylpropanoate	fruity	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
		Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013; Steinhaus and Schieberle 2007)
ethyl butanoate	fruity	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
ethyl 3-methylbutanoate	fruity	Japanese / *Koikuchi* Chinese / raw soy sauce	SAFE-GC-MS/SAFE-AEDA/GC-O SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020; Kaneko et al. 2012, 2013)
ethyl acetate	fruity	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
		Japanese and Chinese	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013)
ethyl propanoate	fruity	Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
isoamyl acetate	fruity, banana-like	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
isoamyl benzoate	fruity, sweet	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)

ethyl hexanoate	fruity, wine-like	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
ethyl octanoate	fruity, sweet	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(X. Gao, Liu, Zhang, et al. 2020)
methyl 2-methylpropanoate	fruity	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Steinhaus and Schieberle 2007)
ethyl phenylacetate	fruity, sweet	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
2-phenylethylacetate	honey, rosy	Chinese / HLFSS, LSFSS Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
ethyl 2-hydroxypropanoate	sweet, fatty	Chinese / HLFSS Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Gao, Liu, Zhang, et al. 2020)
Furan(one)s				
5-methylfurfural	almond, spicy, caramel	Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
furfural	bread, almond, sweet	Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
dihydro-2-methyl-3(2H)-furanone	caramel-like	Chinese / HLFSS, LSFSS	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014)
dihydro-5-methyl-3(2H)-furanone	caramel-like, sweet	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
2,5-dimethyl-4-hydroxy-3(2H)-furanone (4-HDMF)	caramel-like	Japanese / *Koikuchi* Chinese / raw soy sauce	SAFE-GC-MS/SAFE-AEDA/GC-O SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020; Kaneko et al. 2012, 2013; Steinhaus and Schieberle 2007)
5-ethyl-4-hydroxy-2-methyl-3(2H)-furanone (4-HEMF)	caramel-like	Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012; Steinhaus and Schieberle 2007)
		Chinese / HLFSS Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
3-methyl-2(5H)-furanone	caramel-like	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013)
4-hydroxy-5-methyl-3(2H)-furanone	caramel-like, sweet	Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013)
4-decanolide	fatty, milky	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013)
2-furanmethanol	fermented sugar	Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
3-phenylfuran	green bean-like	Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
5-decanolide	milky	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013)
2-acetyl-5-methylfuran	nutty, cocoa-like	Chinese / HLFSS, LSFSS Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Gao, Liu, Zhang, et al. 2020)
3-hydroxy-4,5-dimethyl-2(5H)-furanone (sotolone)	seasoning-like, caramel-like	Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013; Steinhaus and Schieberle 2007)
2,5-dimethylfuran	solvent-like	Japanese and Chinese	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015)
2-acetylfuran	smoky, balsamic	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
2-pentylfuran	green bean, pungent	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
Ketones				
3-ethyl-1,2-cyclopentanedione	caramel-like	Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013)
3-methyl-1,2-cyclopentanedione	caramel-like	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013)
(E)-β-damascenone	cooked apple, honey	Japanese / Kikkoman Chinese / raw soy sauce	SAFE-GC-MS/SAFE-AEDA/GC-O SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020; Steinhaus and Schieberle 2007)
2,3-butanedione	milky	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013)
2-butanone	cheese-like, chemical	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)

2,6-dimethyl-4-heptanone	sweet, fruity	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
3-hydroxy-2-butanone	milky	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013)
1-octen-3-one	mushroom-like	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Steinhaus and Schieberle 2007)
Phenols				
4-ethylguaiacol	smoky, bacon, soy sauce	Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013; Steinhaus and Schieberle 2007)
		Japanese / Chinese	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Su, et al. 2015)
		Chinese / HLFSS, LSFSS Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Gao, Liu, Zhang, et al. 2020)
guaiacol	smoky, burnt	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
		Japanese and Chinese	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013; Steinhaus and Schieberle 2007)
4-vinylguaiacol	spicy, burnt	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Steinhaus and Schieberle 2007)
		Japanese and Chinese	SAFE-GC-MS/SAFE-AEDA/GC-O	(Feng, Su, et al. 2015)
		Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013)
		Chinese / HLFSS, LSFSS Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Gao, Liu, Zhang, et al. 2020)
4-vinylsyringol	spicy	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013)
syringol	spicy, burnt	Japanese / *Koikuchi* Chinese / raw soy sauce	SAFE-GC-MS/SAFE-AEDA/GC-O SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020; Kaneko et al. 2012, 2013)
4-ethylphenol	sweet, spicy	Japanese / *Koikuchi* Chinese / raw soy sauce	SAFE-GC-MS/SAFE-AEDA/GC-O SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020; Kaneko et al. 2012)
Pyrazines				
2,3,5-trimethylpyrazine	burnt	Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
2-isobutyl-3-methoxypyrazine	earthy	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013)
2-isobutyl-3-methylpyrazine	green, celery	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
2-isopropyl-3-methoxypyrazine	earthy, pea-like	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013)
2-ethenyl-6-methylpyrazine	fresh hazelnut-like	Chinese / HLFSS, LSFSS	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014)
2-methylpyrazine	nutty, popcorn	Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
2,5-dimethylpyrazine	roasted nuts	Japanese and Chinese	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Su, et al. 2015)
2,6-dimethylpyrazine	roasted, cocoa	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
2,5-dimethyl-3-ethylpyrazine	roasted, potato	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
2,3-diethyl-5-methylpyrazine	roasted potato	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Steinhaus and Schieberle 2007)
2-ethyl-3,5-dimethylpyrazine	roasted	Japanese / *Koikuchi*	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Su, et al. 2015)
			SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013)
2-ethyl-3-methylpyrazine	caramel-like, baked potato	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
2-isoamyl-6-methylpyrazine	rubbery, sweet	Japanese and Chinese	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015)
Sulfur-containing compounds				
S-methyl-3-methylbutanethioate	cheese-like, soup	Chinese / LSFSS	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014)
2-(methylthio)ethanol	meat-like	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
3-(methylthio)propanol (methionol)	potato-like	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Steinhaus and Schieberle 2007)

		Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020; Kaneko et al. 2013)
3-(methylthio)propanal (methional)	cooked potato-like	Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013; Steinhaus and Schieberle 2007)
		Japanese and Chinese	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015)
bis(2-methyl-3-furyl) disulfide	meat-like	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Steinhaus and Schieberle 2007)
methyl 2-methyl-3-furyl disulfide	meat-like	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013)
dimethyl disulfide	onion, cooked cabbage	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
		Japanese and Chinese	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013)
2-methyl-3-furanthiol	roasty	Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013)
2-furanmethanethiol	caramel-like, sweet	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
2-methylthiophene	sulfury, cooked vegetables	Chinese / LSFSS Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Gao, Liu, Zhang, et al. 2020)
dimethyl trisulfide	sulfury, cooked onion	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
		Japanese and Chinese	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013)
Others				
3-methylindole	animal-like	Japanese / Kikkoman	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2013)
3-hydroxy-2-methyl-4-pyranone (maltol)	caramel-like	Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013; Steinhaus and Schieberle 2007)
		Japanese and Chinese Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014; Feng, Su, et al. 2015; Gao, Liu, Zhang, et al. 2020)
2'-aminoacetophenone	grape-like	Japanese / *Koikuchi*	SAFE-GC-MS/SAFE-AEDA/GC-O	(Kaneko et al. 2012, 2013)
2-acetyl-1-pyrroline	roasty, popcorn-like	Japanese / Kikkoman Chinese / raw soy sauce	SAFE-GC-MS/SAFE-AEDA/GC-O SPME-GC-(O)-MS/SMPE-AEDA	(Gao, Liu, Zhang, et al. 2020; Steinhaus and Schieberle 2007)
3-acetyl-1-pyrroline	roasty, popcorn-like	Chinese / HLFSS, LSFSS	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014)
di-epi-cedrene	sweet	Chinese / HLFSS, LSFSS	SPME-GC-(O)-MS/SPME-AEDA	(Feng, Cai, et al. 2014)
styrene	balsamic, gasoline	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
limonene	citrus, mint	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
naphthalene	mothball-like	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
dodecane	alkane-like	Chinese / raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)
hexadecane	alkane-like	Chinese/ raw soy sauce	SPME-GC-(O)-MS/SPME-AEDA	(Gao, Liu, Zhang, et al. 2020)

2. Taste-active non-volatile compounds

Non-volatile compounds are often important flavor molecules that are primarily linked to the five basic taste categories: sweet, salty, sour, bitter, and umami. Furthermore, they can also contribute to the specific taste nuances of different products. Non-volatile compounds refer mainly to water-soluble molecules with relatively low molecular weights and include for example, amino acids, nucleotides, mono- and disaccharides, phenolic compounds, flavonoids and fatty acids. In soy sauce, non-volatile compounds are widely present and mostly include free amino acids, 5'-nucleotides, small peptides, soluble saccharides and polyols (Kong et al. 2018). These compounds can both be present in different ratios in the raw ingredients used (e.g. soybeans, wheat, *Aspergillus*) and they can also be formed during the production process by the enzymatic and thermal reactions occurring. **Table 2** provides a summary of the most important taste-active non-volatile compounds which have been identified in different types of soy sauces. In the following subsections, the most relevant sensory characteristics are highlighted, along with the chemical compounds which are responsible for the specific taste attributes.

2.1 Umami

Umami taste is described as meaty, broth-like or savory, but it can also increase other taste intensities such as saltiness. Umami is considered as being the 5th basic taste sensation to be described since umami-specific taste receptors are now known (Zhao et al. 2003). Monosodium glutamate (MSG), commonly-used in the Asian cuisine, was the first molecule reported to have an umami taste. MSG is a naturally-occurring component of many soy sauces, but it can also be added as supplement. Examples of other, naturally-occurring umami compounds in Chinese-type soy sauce are aspartate, and the nucleotides 5'-IMP and 5'-GMP as reported by Kong et al. (2018). According to their results, these soy sauces of high grade were described as containing higher levels of taste-active amino acid and nucleotide compounds. In addition to these molecules, several peptides which contribute to umami taste, as well as a number of water-soluble Amadori compounds formed from Maillard reactions, have also been reported (Lioe et al. 2010; Zhuang et al. 2016). Zhuang *et al.* separated five peptides from soy sauce which either had umami taste, or they showed enhancement of the umami sensation (Zhuang et al. 2016). However, the low concentration of these peptides in soy sauce suggests that their overall contribution to umami is likely to be relatively low (Zhao et al. 2016). Nevertheless, many studies suggest that these compounds are able to enhance other basic tastes (Frerot and Chen 2013), and therefore have a strong synergistic activity with glutamate or salt (Shiga et al. 2014). In soy sauce, the highest umami intensity is therefore likely the final result of the presence of multiple key components, even if they do not individually elicit a strong umami taste (Lioe et al. 2010).

2.2 *Kokumi*

In conjunction with the five basic tastes, a new flavor attribute known as '*koku*' or '*kokumi*' was proposed and characterized by Japanese scientists in 1990s (Ueda et al. 1990). *Kokumi* has been considered by some researchers to be the next/sixth basic taste since *kokumi* compounds are also now known to be perceived by specific receptors in the mouth, in this case, through the calcium-sensing receptor (CaSR) (Yang et al.

2019; Zhao et al. 2016). *Kokumi* flavor imparts heartiness/mouthfulness, complexity and long-lasting taste (Nishimura and Kuroda 2019; Zhao et al. 2016), and at the same time, enhances the intensity of other basic tastes, such as umami (Maruyama et al. 2012). Fermented foods like soy sauce, cheese and fermented meat in particular are characterized by imparting *kokumi* taste (Toelstede et al. 2009). This may be due to the metabolic activity of the existent microorganisms generating richness and complexity in flavor, which in turn enhances palatability and deliciousness (Nishimura and Kuroda 2019). Other dishes like curries, stews and ramen noodles are also typified by 'mouthfulness´, complexity and long-lasting taste. However, little is known about the mechanistic aspect of *kokumi* taste. It was demonstrated that some compounds extracted from garlic enhanced lingeringness and 'thick flavor' when added to an umami solution (Ueda et al. 1990), and consequently they proposed to refer to these molecules as *kokumi* compounds. These compounds were identified as being sulfur-containing compounds, however the most important molecules imparting *kokumi* taste, which were subsequently discovered, are γ-glutamyl peptides. They are defined as short-chain oligopeptides, but are now commonly known as *kokumi* peptides (Yang et al. 2019). γ-Glutamyl peptides are naturally found in bacteria, plants (including soybean and wheat) and mammals, and they are components of the glutathione cycle. Glutathione, as well as several γ-glutamyl peptides, can activate the human CaSR receptor. Some examples of *kokumi* active peptides are γ-Glu-Ala, γ-Glu-Val, γ-Glu-Cys, γ-Glu-α-aminobutyryl-Gly and γ-Glu-Val-Gly (Kuroda and Miyamura 2015; Maruyama et al. 2012), which have been identified in many foodstuffs including legumes, *Allium spp.*, cheese, soy sauce, fermented fish and yeast extract (Yang et al. 2019). In other studies on soy sauce (Frerot and Chen 2013; Kuroda et al. 2013), researchers identified and quantitated several other γ-glutamyl peptides: γ-Glu-Glu, γ-Glu-Ile, γ-Glu-Leu, and γ-Glu-Phe. Kuroda *et al.* also identified and quantitated another potent *kokumi* peptide, γ-Glu-Val-Gly (Kuroda et al. 2013), in different brands of commercial Japanese-type dark, light, and white soy sauces. Results indicated that γ-Glu-Val-Gly was more abundant in dark soy sauces, and least abundant in white soy sauces. In addition, among the dark soy sauces, the γ-Glu-Val-Gly content in the ultra-super grade was highest compared to super grade and ordinary grade (grading was based on the percentage of total nitrogen content). It was concluded that differences in γ-Glu-Val-Gly content were derived from differences in the raw material and fermentation process. As mentioned before, Japanese dark soy sauces are produced using almost equal amounts of soybean and wheat, whereas light soy sauces use a very high ratio of wheat to soybean and are fermented under conditions that prevent color development (Kuroda et al. 2013). Consequently, the study revealed that the crude protein contents were found to be significantly positively correlated with γ-Glu-Val-Gly contents, suggesting that the presence of the *kokumi* peptide is directly related to the protein content of the raw ingredients and perhaps therefore also to one of the main storage proteins present in the soybeans.

Apart from the so-called *kokumi* peptides, lipid breakdown products may also influence flavor attributes including *kokumi*. Researchers studied the changes in lipid composition during koji fermentation as well as changes in the sensory characteristics of soy sauce when adding exogenous lipases (Feng, Chen, et al. 2014). Changes in taste related to the addition of lipases were related to high scores in *kokumi* taste, which suggests that lipid metabolism will influence the final *kokumi* characteristics of soy sauce, due to the many metabolite breakdown products formed from lipid degradation. In addition, fatty acid breakdown products formed

during fermentation and heating are also linked to interactions with Maillard reaction products, which results in the formation of many aroma-related (volatile) substances (Diez-Simon et al. 2019). These aroma compounds will be described in the following section.

2.3 Sweet and sour

Saccharides are known to contribute to the sweet taste in many foods while in contrast, organic acids are responsible for the sourness. Typical sweetness and sourness in soy sauce have been explained by a respective high and low ratio of saccharide to acid concentrations. In a study (Lioe 2007), three types of Japanese soy sauces were analyzed for their chemical and sensory characteristics, being dark, white and tamari soy sauces. A slight note of sweetness was perceived in dark and white soy sauces, whereas sourness was perceived in dark and tamari soy sauces (Lioe 2007). This result was correlated with the lower pH found for dark and tamari soy sauces (4.8 and 4.7 respectively) as compared to white soy sauces (5.2). The reason was considered to be the high ratio of saccharide to acid concentrations in white soy sauce, and the low ratio in tamari. The main organic acid found was lactic acid being more abundant in tamari, whereas the main saccharide was glucose, which had the highest concentration in white soy sauce (Lioe 2007). Apart from saccharides, amino acids such as alanine, glycine and serine have also been liked to sweet taste, being more predominant in dark and tamari soy sauces. In another study, lactic acid and pyroglutamic acid were the predominant organic acids contributing to the acidity (sourness) and ensuring a balance in taste of Chinese-type soy sauces (Kong et al. 2018). Specific aroma compounds may also contribute to the sour notes in soy sauce, and these molecules will be highlighted in the following section.

2.4 Bitter and astringent

Apart from peptides, amino acids, fatty acids and lipids, soy sauce also contains traces of other non-volatile secondary components such as polyphenols. Polyphenols have been associated with many health benefitting properties related to digestion and brain function, as well as acting as antioxidants in combating chronic diseases (Khan et al. 2019). Interestingly, some types of polyphenols have also been associated with bitterness, sourness and astringency, in many types of food (Lesschaeve and Noble 2005). Astringency is perceived as a drying or puckering mouthfeel, typically engendered by polyphenols which are naturally present in many plants including wheat and soybeans. Soybeans contain e.g. isoflavones in the form of glucosides, acetyl glucosides and malonyl glucosides (Lin et al. 2015). After fermentation, acetyl glucosides and malonyl glucosides are hydrolyzed to glucosides, and some of these glucosides are also hydrolyzed to become aglycones. In a comparison of different grades of Taiwanese soy sauces (Lin et al. 2015), Lin *et al.* found two isoflavone aglycones (daidzein and genistein) and genistin as a glycosylated isoflavone (**Table 2**). Another study using Chinese-type soy sauce reported daidzin and glycitein, in addition to those mentioned before (Gao, Liu, et al. 2019). Apart from the isoflavones, other phenolic acids (e.g. vanillic acid, syringic acid, ferulic acid, and sinapic acid) have been reported in Japanese soy sauces (Kaneko et al. 2013). These phenolic acids belong to the family of hydroxycinnamic and hydroxybenzoic acids, formed predominantly from the degradation of lignin during grain roasting, and during koji fermentation by *Aspergillus* enzymes (Phillips and Goss 1932), as well as the phenylpropanoyl pathway. These acids have

also been shown to elicit bitterness, sourness and astringency. However, during yeast fermentation and thermal treatment, phenolic acids can degrade by a decarboxylation process to generate volatile methoxyphenol derivatives. For example, vanillic acid degrades to the volatile 2-methoxyphenol (guaiacol), which is instead responsible for smoky and burnt sensory characteristics (See following section).

2.5 Analytical approaches for non-volatile analyses

In recent years, many advanced extraction techniques and modern instrumental analytical applications have been developed for the characterization of non-volatile flavor compounds in different food matrices (**Figure 4**). Here we briefly describe the main techniques specifically in the context of the most important groups of non-volatile soy sauce-relevant metabolites.

For the analysis of free amino acids, organic acids and nucleotides in soy sauce, high-performance liquid chromatography (HPLC) methods with internal standards for targeted detection were used (Kong et al. 2018). Ten soy sauces were compared and quantitative analysis of 17 free amino acids was successfully performed using internal standard calibration curves. In addition, nine organic acids and four nucleotides were detected by UPLC. It was concluded that amino acids, organic acids and nucleotides strongly influenced the taste of the soy sauces and that the analytical techniques used had been successful in quantitating them. However, this study has focused on identified targeted compounds, and the use of standards for their quantitation. As such, there is a total reliance on pre-existing knowledge on pre-defined molecules as well as chemical standard availability. All other compounds, including potentially sensory-relevant (yet chemically unknown) compounds are by definition undetected or ignored. A trend moving towards untargeted metabolomics approaches using e.g. UPLC-MS/MS for profiling analysis is gaining more attention as these allow also the evaluation of a broader range of both known and as yet unknown molecules for comparative analysis (Alonso et al. 2015). This 'untargeted metabolomics' workflow can enable assessment of the potential importance of non-volatile components in a profile even before they have been identified. However, thus far this approach has hardly been exploited for soy sauce analyses to discover additional molecules also not yet recognized to impact the overall flavor.

For the analysis of peptides, techniques such as ultrafiltration and gel filtration chromatography followed by reversed-phase high-performance liquid chromatography (RP-HPLC) are being used to isolate and characterize the different taste-active peptide fractions; sometimes this is combined with a taste dilution analysis (TDA) to define the link to sensory relevance and importance (Lioe et al. 2010). TDA has proven in many studies to be a powerful method to trace the taste-active compounds and determine their impact on sensory experience (Ottinger and Hofmann 2003; Sebald et al. 2018). Characterization of the peptide content has been performed for many types of soy sauce, such as Japanese (Lioe 2007), Indonesian (Lioe et al. 2010), and more recently Chinese types (Zhuang et al. 2016). Zhuang *et al.* characterized five taste-relevant peptides from Chinese soy sauces (Zhuang et al. 2016), and the techniques used to separate the different peptide fractions were sensory-guided fractionation and RP-HPLC. The peptides were subsequently identified by UPLC-MS/MS. Combining different platforms in order to link flavor and chemical composition is gaining more attention and should be the focus for soy sauce flavor studies. Techniques such

as receptor-based assays and cell-based assays (Batenburg et al. 2016; Soares et al. 2018), as a mean to discover new taste-active compounds, have not yet been used in soy sauce.

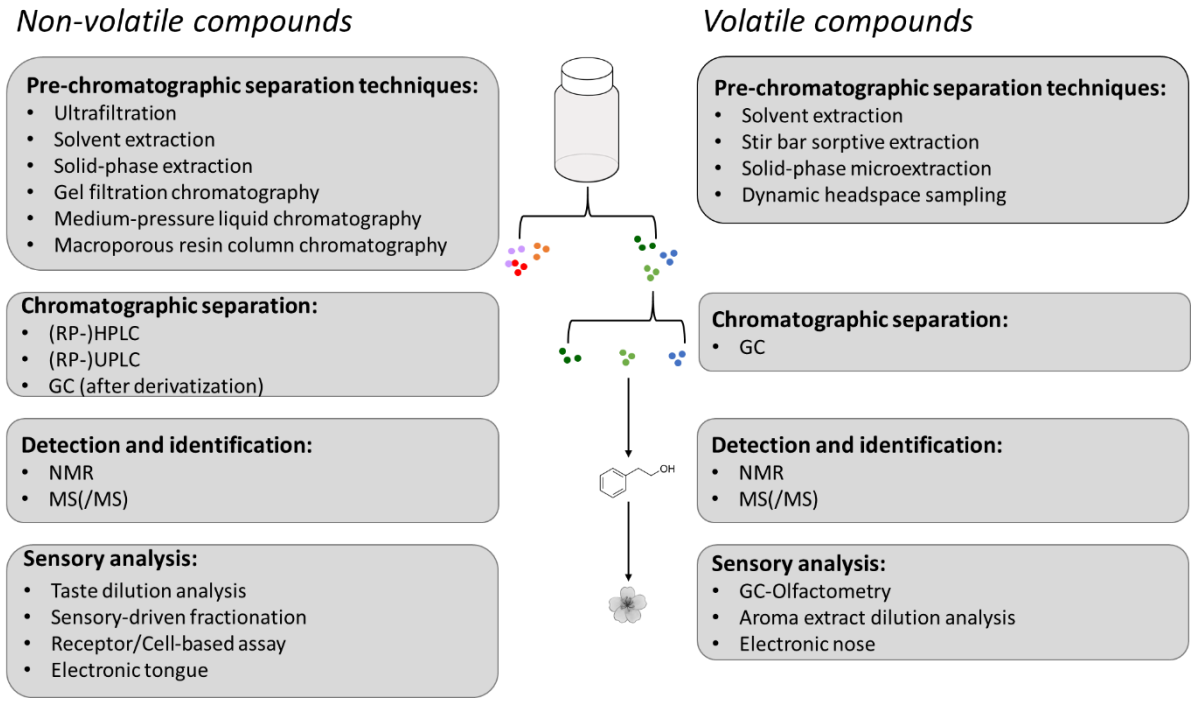


Figure 4. Analytical approaches commonly used to characterize non-volatile and volatile compounds in metabolomics applications. Different techniques can be selected for the various extraction, separation and detection steps, as well as to eventually link specific compounds to flavor characteristics.

For the analysis of γ-glutamyl peptides, a range of other techniques has been applied, such as separation by solid-phase extraction (SPE) prior to UPLC-MS (Frerot and Chen 2013), and derivatization of peptides followed by HPLC-MS/MS (Kuroda et al. 2013). Although the analyses were successful in detecting the compounds, once again the techniques were based on a targeted approach, as mentioned above.

Lipids and the products of lipid degradation processes during koji fermentation are also known to influence the final flavor of soy sauce. For the analysis of lipids (mainly triacylglycerols, TAGs, but also phospholipids, PLs) and fatty acids, liquid-liquid extractions are ordinarily performed using organic solvent mixtures such as chloroform-methanol. The extracted fatty acids are usually then derivatized to fatty acid methyl esters (FAMEs) and analyzed using GC-MS. On the other hand, the organic lipid extract can be analyzed by a RP-HPLC method. Quantitation is done by internal standard spiking and compounds are identified by matching to mass spectral libraries (e.g. NIST).

3. Aroma-active volatile compounds

The importance and mode of action of volatile compounds in food sensory perception have been well covered elsewhere (Belitz et al. 2004). Soy sauce is particularly rich in various volatiles, including alcohols, acids, esters, aldehydes, ketones, phenols, furan(one)s, pyrazines, pyrones and sulfur-containing compounds. **Table 3** provides an overview of the most important aroma-active volatile compounds that

have been identified in different types of soy sauce. Many of these have been linked to the main odor qualities of soy sauce, namely malty, caramel-like, cooked potato-like, floral, alcoholic, sour, smoky, seasoning-like and fruity (Steinhaus and Schieberle 2007).

As indicated above, different localities and different manufacturers often produce contrasting soy sauce types. This results in each containing a unique combination of volatile and semi-volatile compounds, many of which have specific aroma qualities and intensities. In addition to concentration, the odor threshold per compound is of great importance in determining the sensory impact. Those compounds with a low odor threshold therefore still frequently dominate the overall odor despite perhaps only being present in trace amounts. In the following subsections, the most relevant sensory attributes are highlighted, along with the description of the volatile compounds responsible for the specific aroma.

3.1 Malty

2-Methylbutanal and 3-methylbutanal have been found to be responsible for the malty aroma of soy sauce. These short-chain, branched aldehydes can be generated from the amino acids isoleucine and leucine via Strecker (Maillard) reactions or by microbial catabolism (Smit et al. 2009). Strecker degradation reactions are controlled by temperature, as well as substrate concentrations and chemical properties. On the other hand, isoleucine and leucine catabolism, referred to as the Ehrlich pathway, occurs in yeast. Both aldehydes are intermediates and can be further oxidized to the corresponding acids or reduced to the corresponding alcohols, 2- and 3-methyl-1-butanol. These alcohols can also present malty notes, although their odor threshold is much higher than that of the aldehydes and hence they are expected to have potentially lower impact on the final sensory experience (Smit et al. 2009). The origin of 2- and 3-methyl-1-butanol in yeast was demonstrated (Lee et al. 2013), who revealed that samples treated during brine fermentation with *Z. rouxii*, a yeast important in the production of soy sauce (Noda et al. 1980), yielded higher levels of these compounds compared to those treated with other microorganisms. Considering acid-hydrolyzed soy sauce undergoes vigorous heating procedures, while fermented soy sauce production is centered mainly on microbial activity, the aldehydes 2- and 3-methylbutanal can be found in both (Lee et al. 2006). The relative content of these compounds was also found to increase progressively during extended moromi fermentation (150 days) (Gao et al. 2010). Levels of other aroma-impact compounds including alcohols, esters, aldehydes, ketones and furan(on)es also increased and this was accompanied by the enhancement of alcoholic, fruity and caramel-like attributes in sensory evaluation. This indicates that prolonged moromi fermentation is necessary for optimal aroma formation in traditional soy sauce. Furthermore, it was demonstrated that subsequently heating raw soy sauce at 80 °C for 30 min, a step commonly used for pasteurization, also increased the levels of 2- and 3-methylbutanal (Kaneko et al. 2013), in addition to other aroma-relevant compounds.

3.2 Caramel-like aroma

Furanones, such as 2,5-dimethyl-4-hydroxy-3(2H)-furanone (4-HDMF) and the tautomers 2(or 5)-ethyl-4-hydroxy-5(or 2)-methyl-3(2H)-furanone (4-HEMF) have been identified as intense caramel-like aroma

compounds in soy sauce (Steinhaus and Schieberle 2007). 4-HDMF is known to contribute to the aroma of biscuits, dark beer and coffee, while 4-HEMF is characterized by being a sweet, butterscotch, and candy tasting compound, and has been reported, apart from soy sauce, in cheese, coffee, wine, miso, etc. (Dongen and Wiggers). These furanones can be formed through Maillard reactions involving pentoses upon heating (Blank and Fay 1996). However, 4-HDMF and 4-HEMF have also been demonstrated to occur naturally in different yeast species and are mainly associated with carbohydrate metabolism in *Z. rouxii* (Lee et al. 2013; Yanfang and Wenyi 2009). The pyranone 3-hydroxy-2-methyl-4H-pyran-4-one (maltol) is another potent caramel aroma compound found frequently in soy sauce (Sun et al. 2010). Maltol can be formed from saccharides through 2,3-enolization during heating or directly from the Amadori product (Yaylayan and Mandeville 1994). It has been detected in steamed soybean, but not in raw soybean, and is therefore considered to originate from heating the soybeans prior to koji fermentation (Feng et al. 2013). Acid-hydrolyzed soy sauce is richer in maltol, due to the more vigorous heating procedures used (Lee et al. 2006). Interestingly, a study suggests that the caramel-like sensory attribute can be influenced by spiking soy sauce with individual odorants (Feng, Cai, et al. 2014), even if these themselves do not necessarily possess a caramel-like aroma in water. For example, adding 3-methylbutanal alone to soy sauce not only imparted an intense malty aroma, but also increased caramel-like aroma tones. This reflects how complex flavors can be and how a matrix effect and combinations of different aroma compounds can elicit a different final overall aroma.

3.3 Cooked potato-like aroma

One of the most potent odorants in soy sauces presenting a cooked potato-like aroma is 3-(methylthio)propanal (methional) (Steinhaus and Schieberle 2007). This S-containing compound has also been implicated as a major aroma impact compound in various types of cheese (Frank et al. 2004). Despite being present at low concentrations, it readily exceeds its orthonasal odor threshold which is particularly low for this compound (Kaneko et al. 2012; Sun et al. 2010). Furthermore, these authors detected methional as a common factor among five contrasting types of Japanese soy sauce and proposed the cooked potato-like note to be one of the most important aroma characteristics of Japanese soy sauce. Methional is a product of the thermally-induced Strecker degradation from methionine and is generated during both fermentation and pasteurization (Pripis-Nicolau et al. 2000). When heating raw soy sauce, the levels of methional were increased and higher sensory scores were obtained for this cooked potato-like attribute (Kaneko et al. 2013). To a lesser extent, the alcohol 3-(methylthio)-1-propanol (methionol) may also contribute towards the aroma of cooked potato in soy sauce (Feng, Cai, et al. 2014; Steinhaus and Schieberle 2007). Lastly, 2-ethyl-3,5-dimethylpyrazine, which also has a potato-like note, is known to appear after heating soy sauce (at 145 °C for 20 min) and thus may actually only be formed after it is used as a seasoning in domestic cooking (Steinhaus and Schieberle 2007).

3.4 Floral

The alcohol 2-phenylethanol has a rose-like odor and is one of the main compounds which has been linked to the floral notes of soy sauce (Feng, Cai, et al. 2015; Steinhaus and Schieberle 2007). It is produced by

yeast only during traditional fermentation from the aromatic amino acid phenylalanine and is more abundant when the production involves long moromi fermentation periods, such as those typical for Chinese-type HLFSS soy sauces (Feng, Cai, et al. 2014; Lee et al. 2006). Although considered of lesser importance, methyl and ethyl benzoate esters in soy sauce may also exhibit a floral aroma (Feng et al. 2017; Feng, Su, et al. 2015). Additionally, phenylacetaldehyde, known for its hyacinth- or honey-like notes, has a significant contribution to the overall aroma of soy sauce (Steinhaus and Schieberle 2007). It has been identified in many fermented soybean foods and its concentration generally peaks during the later stages of soy sauce moromi fermentation (Feng, Cai, et al. 2014; Wei, Wang, Chen, et al. 2013).

3.5 Alcoholic

Ethanol is the major contributor to the alcoholic attribute of soy sauce (Feng et al. 2017; Steinhaus and Schieberle 2007). During the fermentation period, enzymes present in koji convert starch to various saccharides which are subsequently fermented to yield volatile components including ethanol. Both alcoholic fermentation by yeast (*Z. rouxii* and *Candida* species) and lactic acid fermentation by different bacteria species are responsible for the ethanol content in conventionally brewed soy sauce which sometimes reaches 2-3 % (Hamada et al. 1989; Luh 1995). In contrast, tamari soy sauce, which is produced without or with only small amounts of wheat, lacks a major carbohydrate source and, as a result, contains little or no ethanol (Mengru Liu et al. 2014). Moreover, when the preparation of soy sauce involves high temperatures, such as those typical for Chinese LSFSS soy sauces, ethanol can be lost during production through evaporation (Feng, Cai, et al. 2014).

3.6 Sour

Sour notes in soy sauce are primarily caused by acetic acid, which has, amongst the volatiles, one of the highest concentrations (Steinhaus and Schieberle 2007; S. Y. Sun et al. 2010). Nevertheless, its odor threshold is considerably higher than that of other odorants, generally resulting in lower sensory scores for the sour attribute (Feng, Su, et al. 2015; Steinhaus and Schieberle 2007). Acetic acid is produced by lactic acid bacteria during moromi fermentation. This explains why researchers detected higher levels of this compound in soy sauce samples which had been inoculated with *T. halophilus* (Lee et al. 2006), a lactic acid bacterium often employed in the fermentation of soy sauce (Noda et al. 1980). In addition, 2- and 3-methylbutanoic acid, two branched-chain acids produced from amino acid metabolism in yeast through oxidation of the corresponding aldehyde intermediates (Hazelwood et al. 2008), also deliver sour notes (Kaneko et al. 2012). Most studies however report this odor rather as being cheese-like or sweaty (Feng, Su, et al. 2015; Lee et al. 2013). Particularly 3-methylbutanoic acid is relatively abundant in Korean, Japanese and Chinese soy sauces, where it can be responsible for a sweaty odor note (Lee et al. 2006; Steinhaus and Schieberle 2007; Sun et al. 2010). Short-chain fatty acids such as propanoic acid and butanoic acid can originate from the lipolysis of microorganisms and provide vinegar, cheese-like or rancid odor notes (Lee et al. 2013). On the other hand, certain acidic compounds, including 4-oxo pentanoic acid (levulinic acid), can be derived from thermal degradation of hexose monosaccharides and are detected in high amounts in acid-hydrolyzed soy sauce (Lee et al. 2006). Finally, sorbic acid is commonly used as an

added preservative in certain soy sauces and its presence will inevitably contribute to the sour note in these products (Feng, Cai, et al. 2014).

3.7 Smoky

Methoxyphenols are components of great importance for smoky attributes in soy sauce and other fermented foods such as dry-cured ham (Jerković et al. 2007; Lee et al. 2006). Among them, 2-methoxyphenol (guaiacol), 4-ethyl-2-methoxyphenol (4-ethylguaiacol), 2,6-dimethoxyphenol and methoxy-4-vinylphenol (4-vinylguaiacol) have been reported in various studies to be the odor-active compounds involved (Feng, Su, et al. 2015). Moreover, 4-ethylguaiacol was found to ameliorate the salty taste of soy sauce (Yokotsuka and Sasaki 1998). These phenolic compounds are formed from the hydroxycinnamic and hydroxybenzoic acids produced during koji fermentation. At this stage, most methoxyphenols do not appear to be present (Feng et al. 2013). However, ferulic acid, the hydroxycinnamic acid being the major phenolic constituent of koji, is converted into 4-ethylguaiacol during fermentation through decarboxylation by *Candida (Torulopsis)* yeast species, but not by *Z. rouxii* (Yokotsuka and Sasaki 1998). In addition to the yeast fermentation pathway, methoxyphenols may also be formed through thermal decomposition of the corresponding precursors (Fiddler et al. 1967; Gao et al. 2010; Kaneko et al. 2013). In all cases, the formation of methoxyphenols was accompanied by higher sensory scores for smoky and burnt notes.

3.8 Spice-like, fenugreek-like

3-Hydroxy-4,5-dimethyl-2(5H)-furanone (sotolone), a spice-like, or more precisely fenugreek-like aroma compound, has been proposed to be one of the most intense odorants in soy sauce, and indeed, also in sake and sherry wine (Kaneko et al. 2012; Steinhaus and Schieberle 2007). Although its concentration in soy sauce is low, it is known to have a particularly low odor threshold and hence can have significant impact even at low levels. When sotolone is present at low concentrations, it is known to have a caramel, burnt sugar aroma (Effenberger et al. 2019). Different precursors and reaction pathways have been proposed for its (bio)synthesis but its formation mechanism in soy sauce has not yet been elucidated (Pham et al. 1995). Sotolone has been described as the flavor principle of food prepared from plant protein hydrolysates (Effenberger et al. 2019). Furthermore, sotolone rapidly decomposes at temperatures above 80 °C (Dagan et al. 2006), which explains its significant decrease and the loss of the seasoning-like aroma note when heat treatments (cooking) have been used (Steinhaus and Schieberle 2007).

3.9 Fruity

Various esters are responsible for the fruity note. This note was observed to be more pronounced in Japanese soy sauces than in traditional soy sauces from China (Feng, Su, et al. 2015). In addition, esters are particularly important for masking the presence of potentially unpleasant odors in fermented foods (Qin and Ding 2007). For instance, dimethyl disulfide and dimethyl trisulfide, essentially arising from the metabolism of sulfur-bearing precursors such as methionine (Landaud et al. 2008), are commonly reported

constituents of soy sauce and may give rise to off-odors (Kaneko et al. 2013). Among the most abundant esters in soy sauce are ethyl 2-methylpropanoate, ethyl 2-methylbutanoate and ethyl 3-methylbutanoate, which impart sweet and fruity notes (Feng, Su, et al. 2015). Esters are associated with the metabolism of lipids by yeast, which provides a large number of acids and alcohols that may subsequently undergo esterification to yield a variety of esters (Maarse and Visscher 1989). As these are typical fermentation products they are not found in acid-hydrolyzed soy sauce (Lee et al. 2006). Esters accumulate predominantly at the intermediate stage of fermentation (Wei, Wang, Lv, et al. 2013), and are more abundant in production methods involving longer ageing periods (Feng, Chen, et al. 2014). High-molecular weight fatty acid esters, including ethyl palmitate, ethyl linoleate and ethyl oleate may also exist at high concentrations in soy sauce. They are generated from long-chain fatty acids in the presence of fungal lipases after a prolonged period of constant temperature fermentation (Zheng et al. 2013). Nevertheless, they are seldom considered to contribute significantly to the characteristic aroma of soy sauce.

3.10 Roasted, nutty and mushroom-like

Among the remaining aroma-active volatile compounds in soy sauce are various pyrazines, including 2,6-dimethylpyrazine, 2-methylpyrazine and 2,3,5-trimethylpyrazine, which provide nutty, baked and roasted notes to soy sauce samples (Feng et al. 2017; Feng, Su, et al. 2015). These compounds are mainly formed via Maillard reactions and were found to be major components in the volatiles of acid-hydrolyzed soy sauces which undergo heating (Lee et al. 2006). These pyrazines however, possess relatively high odor thresholds, thereby potentially imparting limited impact on the sensory properties of soy sauce (Feng, Su, et al. 2015; Kaneko et al. 2013). Finally, 1-octen-3-ol and 1-octen-3-one have been identified as aroma-relevant compounds which possess a mushroom-like note (Feng, Su, et al. 2015; Lee et al. 2006; Steinhaus and Schieberle 2007). These volatiles are generated from lipid oxidation by fungal species such as *Aspergillus oryzae* and thus are found at relatively high levels in the fermented soy sauces (Feng et al. 2013).

3.11 Analytical approaches for volatile analyses

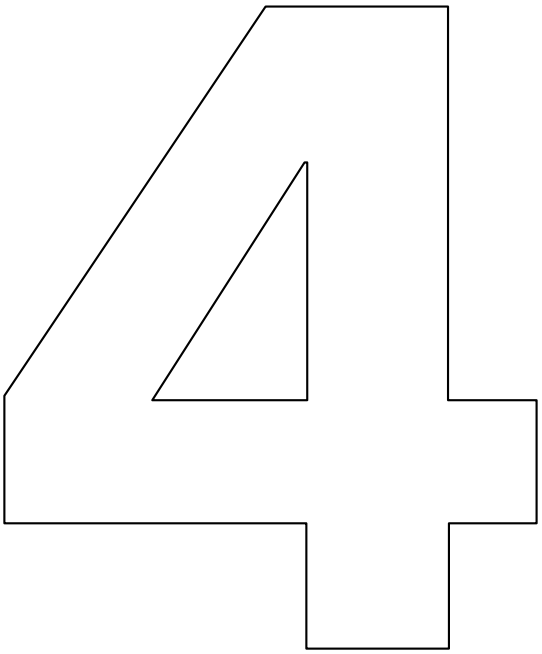
For the analysis of volatile compounds gas chromatography-mass spectrometry (GC-MS) approaches are generally used, preceded by a sample preparation procedure to eliminate interfering substances and improve detection sensitivity by enriching solutes from the aqueous matrix (Parliament 1997). These sample preparation techniques for analysis always influence the volatile profile obtained by discriminating volatiles on the basis of their polarity, volatility, etc. The more common techniques used for extracting volatiles and semi-volatiles are based on liquid-liquid extractions, also called solvent extractions, and/or sorptive-based extractions (**Figure 4**) (Pico et al. 2016). Liquid-liquid extraction uses organic solvents to collect the volatile compounds from the sample matrix. A very common technique to analyze aroma volatiles in soy sauce has been solvent-assisted flavor evaporation (SAFE) (Jo et al. 2011; Kaneko et al. 2013; Steinhaus and Schieberle 2007; Zhu and Cadwallader 2019). SAFE is considered to be one of the best solvent extraction methods that captures the complete aroma extract of a sample, including the most labile aroma compounds. In addition, SAFE also avoids the formation of thermally generated artefacts during thermal desorption and GC-MS analyses (Zhu and Cadwallader 2019). However, SAFE is time consuming, labor

intensive and uses high amounts of (toxic) organic solvents. Sorptive-based extractions, on the other hand, use ad- and/or absorbent polymers to collect the volatile fraction from the samples, and do not need the use of any organic solvents. Moreover they are also fast, easy to manipulate and cost-effective (Nogueira 2015). Headspace extraction (e.g. SPME: Solid-phase microextraction) or in-liquid extraction (e.g. SBSE: Stir bar sorptive extraction) are the most popular among all the techniques proposed in recent years. All these pre-extraction techniques have specific advantages and disadvantages in terms of analyte coverage, reproducibility and sensitivity (Diez-Simon et al. 2019), and hence, different techniques, prior to GC-MS, have been used in the analysis of the volatile composition of soy sauces. This has generated inconsistencies between the volatiles detected, as well as the concentrations measured in different studies, when using different approaches. For example, furanone compounds are important molecules impacting the caramel-like aroma in soy sauce samples. Furanones, such as 4-HDMF and 4-HEMF, are commonly detected when using solvent extraction methods. However, due to their low volatility, they can be lost when other methods (e.g. headspace extraction) are used (Lee et al. 2013; Lee et al. 2006). Nevertheless, these semi-volatile compounds can also be detected by LC-MS (Li et al. 1998). Similar to furanone compounds, maltol possesses low volatility and is therefore rarely detected when headspace trapping techniques such as dynamic headspace sampling (DHS) have been used (Wanakhachornkrai and Lertsiri 2003). Although maltol could be detected with HS-SPME (Feng et al. 2017), this was only in trace concentrations. Feng *et al.* carried out SPME optimization and found that fibers containing the polymers divinylbenzene (DVB) and/or polyacrylate (PA) had a higher affinity for heterocyclic compounds such as maltol, compared to those with more apolar stationary phases like polydimethylsiloxane (PDMS) (Feng et al. 2017). Elevated temperatures and longer extraction times also increased the chromatographic response of maltol, as well as other compounds with high molecular mass and boiling point. Lastly, sotolone often remains undetected when the more standard sample extraction methods such as solid-phase micro extraction (SPME) are used (Feng, Su, et al. 2015). This can be attributed to limitations of SPME, such as the retention capacity and selectivity. Nevertheless, sotolone is also semi-volatile and can be detected by LC-MS.

Consequently, the method of choice may influence the volatile profile obtained, and carefully weighed choices must be taken when performing untargeted metabolomic analyses. To overcome the main limitations of SPME, stir bar sorptive extraction (SBSE) is increasingly being used to analyze volatile (and less volatile) compounds in many types of samples (Ochiai et al. 2018). SBSE uses a larger sorptive phase volume (24 µL compared to 0.5 µL in SPME) and consequently, higher sensitivities can be achieved. Moreover, the sorptive phase in SBSE is in direct contact with the sample which makes it more efficient for trapping less volatile, more polar compounds. However, where automation of the SPME procedure has been successfully achieved, SBSE still requires manual actions. In addition, and a limited number of coatings are commercially available, as compared to SPME. SBSE has not yet been used for the analysis of soy sauce, although it has been used as an effective and reliable method for quantitation of volatile compounds in Chinese soy sauce aroma type liquor (Fan et al. 2011). Consequently, new methods for the analysis of soy sauce aroma can still be exploited in order to improve the quality and comprehensiveness of the analysis.

4. Conclusions and future perspectives

Soy sauce is a popular food commodity used directly either as a condiment or as a cooking ingredient and its popularity has moved from Asia to right across the world. The two very different production processes (microbial vs chemical) as well as all the additional cultural nuances that can be introduced, both in the balance of the initial raw ingredients as well as in the subsequent details of the production process steps, result in a broad range of contrasting products being available on the global market. Each of these soy sauce products are known to have unique/contrasting flavor profiles, and there are parallels to be found with e.g. balsamic vinegar where extensive differences in product price reflect production method differences and the period of aging which define the final sensory experience as perceived by the consumer. Recent studies have advanced our knowledge of soy sauce composition allowing us to identify more sensory-relevant compounds. However, we are still far from having a full mechanistic understanding of how final chemical composition determines sensory experience - i.e. determining which components are most causal. This also entails that we are still limited in our capacity to link differences in the final product with the different raw ingredients and processes used. Nevertheless, our ability to predict and potentially confirm the pathways between raw ingredient precursors and the final chemical mixture is improving and with this knowledge we should be better able to design directed improvements in production processes to deliver more robust and consistent soy sauces in a more (cost-)efficient way. As indicated here, increased future use of so-called untargeted metabolomics analyses of both volatile and non-volatile metabolites can offer us valuable new routes of exploration. These approaches facilitate the analysis of a much broader range of compounds, even before their chemical structures have been determined. Consequently, when these analyses are combined with sensory analysis (e.g. aroma extract dilution analysis, AEDA), this allows us to both identify sensory–relevant compounds as well as filter out the most important ones from these mixtures which deserve direct attention for metabolic annotation and the identification of novel compounds.



Comparison of different volatile trapping techniques for the comprehensive analysis of food flavourings by Gas Chromatography-Mass Spectrometry

Carmen Diez-Simon, Brenda Ammerlaan, Marco van den Berg, John van Duynhoven, Doris M. Jacobs, Roland Mumm, Robert D. Hall

Journal of Chromatography A (2020)
DOI: 10.1016/j.chroma.2020.461191

Abstract

Trapping volatiles is a convenient way to study aroma compounds but it is important to determine which volatile trapping method is most comprehensive in extracting the most relevant aroma components when investigating complex food products. Awareness of their limitations is also crucial. (Un)targeted metabolomic approaches were used to determine the volatile profiles of two commercial flavourings. Four trapping techniques were tested as was the addition of salt to the mixture. Comprehensiveness and repeatability were compared and SBSE proved particularly suitable for extracting components such as polysulfides, pyrazines and terpene alcohols, and provided an overall broader chemical spectrum. SPME proved to be more suitable in extracting sesquiterpenes and DHS in extracting monoterpenes. Adding salt to the sample had only quantitative effects on volatiles as detected by SPME. These results help clarify the advantages and limitations of different trapping techniques and hence deliver a valuable decision tool for food matrix analysis.

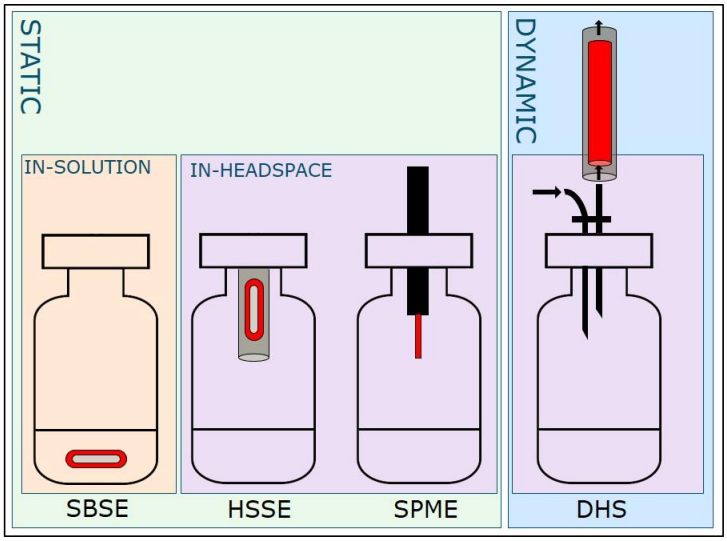
Keywords: process flavours; volatiles; headspace techniques; stir bar sorptive extraction (SBSE); gas chromatography-mass spectrometry (GC-MS); Maillard reaction

1. Introduction

Volatile and semi-volatile compounds play a central role in food quality as they are often the fragrant or bioactive compounds that are primarily important in imparting positive sensory attributes. However, such components can also be instrumental in imparting negative sensory attributes through being so-called, 'off-flavours'. Off-flavours (or taints) are undesirable sensory notes often arising as a result of incorrect preparation methods, natural product degradation or the use of incorrect storage procedures (Ridgway 2015). For many years the more standard techniques used for extracting volatiles and semi-volatiles involved some kind of liquid-liquid extraction. However, a recent trend is a progression towards simplification, miniaturization and minimization of organic (toxic) solvents used in order to be more sustainable and also reduce waste (Blasco et al. 2004; Raynie 2006). There are now several sorptive-based methods which are faster and avoid the use of organic solvents for the analysis of volatile compounds (Richter et al. 2017). These methods commonly use ab- and adsorptive materials in which the volatiles are collected either in or above (headspace) a (liquid) food matrix. Headspace extraction (e.g. SPME: Solid-phase micro extraction; DHS: Dynamic headspace; HSSE: Headspace sorptive extraction) or in-liquid extraction (e.g. SBSE: Stir bar sorptive extraction) are the most popular among all the techniques proposed in recent years (Nogueira 2015).

Figure 1 shows a schematic set up of the four trapping techniques used in this study. Stir bar sorptive extraction in solution (SBSE) and in headspace (HSSE), are based on the trapping of volatiles onto a polymer coated on a magnetic stir bar (Ochiai et al. 2018). Both SBSE and HSSE are techniques that were developed 20 years ago and have shown great capacity for the static sorptive extraction at (ultra-)trace levels of non-polar to medium-polar solutes with volatile to semi-volatile characteristics in complex food systems (Nogueira 2015). The stir bar can be either placed in the liquid sample or in the so called headspace above (**Figure 1**). The latter requires that the volatiles are driven out of the sample material into the headspace. On the other hand, SPME has become a more widely used headspace technique to analyse volatiles arising from many types of food sample (Kataoka et al. 2000), particularly due to its easy automation and the wider variety of ab-/adsorbent polymers available. SPME uses a small fused-silica fiber that is coated with one or more polymers to trap the volatiles. The other commonly used method to trap volatiles in foodstuffs is done by a dynamic headspace system (DHS). Its main difference compared to static techniques is that DHS traps volatiles through flushing the sample with a flow of gas. This helps accumulate the analytes more efficiently in the adsorbent phase which can be directly connected to the vial through a needle. DHS also provides a wide variety of adsorbents. All four techniques have their own advantages and disadvantages and hence the specific research question should determine which is optimal. With respect to sensitivity and selectivity, for instance, the availability of different coating polymers and coating volumes can influence the performance of the technique for certain compounds. With SBSE, as compared to SPME, more sorptive phase volume is used (24 µL and 0.5 µL, respectively) and consequently, higher sensitivities can be achieved (David and Sandra 2007). On the other hand, SPME fibres are currently available in several combinations of different types of coating while for SBSE and HSSE the coatings are limited to two (PDMS and EG-Silicone). Consequently, the selectivity of the technique for certain compounds differs depending on the coating used.

Additionally, static and dynamic modes have shown different chemical profiles of the same food samples (Petretto et al. 2017). DHS does not depend on the equilibrium between the gas and liquid phase. Therefore, DHS is able to concentrate the analytes at ultra-trace levels and thus sensitivity also improves (Ochiai et al. 2014). The goal of this study was to compare the four trapping techniques for comprehensiveness and repeatability by analysing two process flavours (PF) with high diversity in chemical composition.



Trapping technique		Ab-/adsorbent phase
SBSE	Stir-bar sorptive extraction	PDMS
HSSE	Headspace sorptive extraction	PDMS
SPME	Solid-phase microextraction	PDMS/DVB/CAR
DHS	Dynamic headspace	Tenax TA

Ab-/adsorbent phase

Figure 1. A schematic representation of the four volatile trapping techniques used in this study. The ab-/adsorbent coatings, used specifically in this study, are highlighted in red. We can differentiate between static in-solution, static in-headspace and dynamic in-headspace approaches. **SBSE**: Stir-bar sorptive extraction; **HSSE**: Headspace sorptive extraction; **SPME**: Solid-phase microextraction; **DHS**: Dynamic headspace system.

Process flavours (PF), also referred to as 'reaction flavours', are composed of a complex mixture of ingredients that are thermally processed under controlled conditions during manufacture. They are important flavouring ingredients regularly added to improve taste and enhance specific sensory attributes in savoury food products, such as soups, snacks, sauces or ready-to-eat meals (Diez-Simon et al. 2019). One of the most prominent groups of PFs now used as flavouring ingredients are based on yeast autolysates and extracts (In et al. 2005). They enhance flavour by imparting cheesy, meaty or savoury notes, and can be used in meat substitutes for vegetarian food applications. Two examples of commercial PFs are Maxagusto™ (DSM Food Specialties, the Netherlands) and Flavour Yeast Extract (Hubei Angel Yeast Co. Ltd, China). By varying the yeast strain, the yeast processing conditions, such as temperature and time, and the combination with other basic ingredients, different savoury notes can be obtained to create a wide palette of high quality taste and aroma supplements (Ames 1994; Mahadevan and Farmer 2006; Münch et al. 1997). Studying which volatile compounds might contribute to the aroma and taste of PFs is of great importance for food formulation and flavour studies. The reactions that lead to the formation of these volatiles during the production of PFs are mainly related to lipid oxidation, Maillard reactions and thermal degradation of sugars, proteins, ribonucleotides, pigments and vitamins (Parker 2015b). Of particular interest are Maillard reactions, which occur between a nitrogen-containing compound and a reducing sugar (Shahidi et al. 2014). When catalysed by high temperatures, a cascade of chemical reactions are triggered to potentially form a

vast range of volatile aroma-related compounds. These compounds are chemically diverse and display many different physico-chemical properties (for an overview see Diez-Simon et al. (2019)). Consequently, characterizing the volatile composition of PFs will give us insights into strategies for improving flavour quality, and approaches for investigating the link of certain volatiles to desirable sensory attributes.

This study aims to detect the highest number of volatile compounds present in process flavours (PFs) in a reproducible manner. We have focused on the most novel sorption-based techniques (SBSE, HSSE, SPME and DHS) as they do not need the use of any (toxic) organic solvents and they are fast, easy to manipulate and cost-effective, among many other reasons (Nogueira 2015). We hypothesized that, from all sorptive techniques, SBSE will cover a higher range of volatiles as it is in direct contact with the liquid and thus, more polar and less volatile compounds will (additionally) be trapped as compared with the most commonly used techniques (SPME and DHS).

To test our hypothesis, we have developed and compared four volatile trapping techniques coupled to GC-MS in order to study which technique gives the most comprehensive overview of the volatile composition of two different PFs (Maxagusto G28 and Maxagusto S99). Repeatability of the techniques was also tested. Furthermore, the use of salt in sample preparation was also tested as some salts are known to release more volatiles into the gas phase.

2. Materials and Methods

2.1 Food materials

Two yeast extract-based process flavours (PF) from the series Maxagusto were obtained from DSM Food Specialties (Delft, the Netherlands). These are commercially available as Maxagusto G28 and Maxagusto S99 (www.dsm.com). Maxagusto G28 is an aromatic, pungent garlic flavouring reported to have a distinct fresh, fried garlic profile. Maxagusto S99, in contrast, has a natural roasted spice base for Asian-type recipes with the flavour of cooked vegetables (www.dsm.com). Both were obtained as 50 g dry powder in sealed foil bags. After the samples were aliquoted, they were stored at -80°C until analysis.

Prior to analysis, Maxagusto samples were suspended in tap water at a concentration of 2 mg/mL. This concentration is equivalent to the dosage level commonly used in the food application. Once the powders were fully suspended, they were sonicated for 10 min to break up any remaining aggregates or small solid particles. For the method involving salt (Section 3.4), a saturating amount of salt at a final concentration of 5 M (for $CaCl_2$) and 6 M (for NaCl) was added after the Maxagusto samples had been dissolved. The vials were stirred and immediately closed after the addition of salt. Screw-cap glass vials (10 mL and 20 mL) with silicone/PTFE septa (Supelco, PA, USA) were used.

2.2 Chemicals

A range of reference chemicals were used: an *n*-Alkane (C_8-C_{22}) series was purchased from Sigma-Aldrich (Milwaukee, WI, USA). The standards used for metabolite identification were: Benzaldehyde, Diallyl-disulfide, Dimethyl-disulfide, 2,5-Dimethyl-pyrazine, Dimethyl-trisulfide, *p*-Eugenol, Furfural, Hexanal, D-

Limonene, Methyl-propyl-trisulfide, Nonanal, Octanal, alpha-Pinene, alpha-Terpineol and alpha-Terpinolene. All compounds were purchased from Sigma-Aldrich. Methanol (Biosolve BV, NL) was used as solvent for the preparation of the standard solutions. Calcium chloride ($CaCl_2$) and sodium chloride (NaCl) were purchased from Sigma-Aldrich and Honeywell-Fluka (Seelze, Germany), respectively.

2.3 Trapping techniques

Four different trapping techniques for volatile compounds were tested in this study. For all four, samples were analysed using the same GC-MS instrument with the same settings by thermally desorbing them using a multi-purpose sampling robot (MPS-2, Gerstel, Mülheim, Germany). Prior to the comparison of these four methods an extensive series of preliminary trials was performed for each in order to determine which specific set of parameters and settings gave the best results. Different combinations of adsorbent types, extraction times, temperature regimes, etc. were tested as described in Supplementary **Table S1**. The best result was considered to be those conditions which gave the most comprehensive overview of the components present. The optimal procedure for each method is described below.

2.3.1 Stir bar sorptive extraction in solution (SBSE) and in headspace (HSSE)

For SBSE and HSSE, we used a set of 20 Gerstel 10 mm x 0.5 mm PDMS stir bars (Gerstel, Germany). Prior to sampling, the stir bars were conditioned for 60 min at 260 °C under a continuous stream of helium gas (grade 5.0) in empty glass tubes. After conditioning, they were individually stored in clean closed screw-cap glass vials until use. Clean tweezers or a magnetic bar were used for handling the stir bars. Each sample had its individual and traceable stir bar. For the preliminary tests shown in Supplementary **Table S1**, PDMS MonoTrap (Monolithic Material Sorptive Extraction (MMSE), GL Sciences, Japan) and EG-Silicone stir bars (Gerstel, Germany) were also tested and as the results did not show a comparable or better coverage of volatiles as PDMS stir bars, we did not proceed further with these approaches.

The following methodology was applied for the sample series: A volume of 3 mL sample solution (2 mg/mL) in a 20 mL glass vial was used for volatile trapping, which translates into a phase ratio of 125 (sample volume/PDMS volume, thereby 3000 µL/24 µL = 125). The use of a small phase ratio is fundamental for a uniform extraction of a wide range of compounds. For SBSE analysis, the stir-bar was placed either in the Maxagusto solution or positioned in the headspace above the sample using a glass insert (Gerstel, Mülheim, Germany) for HSSE (**Figure 1**). In the case of $SBSE^2$ analysis (Werkhoff et al. 2001), two stir bars were placed in the same vial, one for SBSE and one for HSSE (**Figure 4**). After inserting the stir bars, vials were immediately closed and incubated in a water bath at 60 °C for 10 min. To extract volatiles, samples were placed on a multipoint magnetic stirring plate (Thermo Scientific Variomag, Waltham, USA) at room temperature (RT), 450 rpm for a further 80 min. After the full extraction time (90 min, 60°C+RT), the stir bars were removed from the samples using tweezers or a magnetic bar, rinsed for 2-3 seconds in ultraclean water, dried with a clean tissue and put into empty clean glass liners for desorption. The analyses were done right after sampling. Volatiles were first desorbed from the stir bars in a Thermal Desorption Unit (TDU, Gerstel) connected to the Cooled Injection System (CIS) of the GC-MS. Desorption in the TDU was done in

splitless mode at a temperature of 30 °C for 0.5 min, and a ramp of 120 °C/min to reach a final temperature of 250 °C (with a 5 min hold) using helium (Grade 5.0) as carrier gas. The desorbed volatiles were focused in the CIS on a glass liner packed with Tenax TA at -10 °C which was then flushed for 0.2 min at -10 °C with helium flow of 35 mL/min (solvent vent mode). The volatiles in the CIS were transferred to the analytical column by rapidly firing the trap from -10 °C to 250 °C with an increase of 720 °C/min after which the temperature was held at 250 °C for 5 min. A split of 1:5 was used. The desorption and injection was fully automated for all samples using a Gerstel MPS-2 autosampler and operated using Gerstel MAESTRO software version 3.2.

2.3.2 Headspace solid-phase micro extraction (SPME)

For SPME, a PDMS/DVB/CAR (Polydimethylsiloxane / Divinylbenzene / Carboxen) 50/30 µm diameter, 1 cm length (Supelco, PA, USA) fiber was used. Prior to analysis, the fiber was conditioned as recommended by the manufacturer.

The following methodology was applied for the sample series: A volume of 1 mL sample solution (2 mg/mL) in a 10 mL glass vial was used to trap volatiles. The vials were incubated at 60 °C for 10 min with agitation. Subsequently, volatiles were trapped by exposing the fiber to the headspace of the vial for 20 min at 60 °C without agitation (**Figure 1**). The fiber was then thermally desorbed in the CIS containing an empty glass liner (1 mm ID) with a helium flow of 1 mL/min at 250 °C for 2 min onto the GC column, in splitless mode. Trapping and injection was fully automated using a Gerstel MPS-2 autosampler and operated using Gerstel MAESTRO software version 3.2.

2.3.3 Dynamic headspace extraction (DHS)

The trapping of the volatiles was done using the DHS module by Gerstel mounted onto a Gerstel MPS-2 autosampler and operated using Gerstel MAESTRO software version 3.2. A glass tube packed with 60 mg of Tenax TA (Gerstel, Germany) was used as adsorbent. Before sampling, the Tenax tube was conditioned for 60 min at 285 °C with a constant flow of helium (Grade 5.0) as carrier gas.

The following methodology was applied for the sample series: A volume of 3 mL sample solution (2 mg/mL) was added to a 20 mL glass vial. The samples were first incubated at 30 °C for 10 min, with agitation. For DHS analyses, 30°C was chosen instead of 60°C due to problems with water trapped on the Tenax phase. After incubation, volatiles were collected on the Tenax cartridge by purging the vial with a continuous flow of helium at 30 mL/min for 10 min (**Figure 1**). The temperature of the vial was maintained at 30 °C while the temperature in the Tenax trap tube was set to 20 °C. After collection, the Tenax tube was purged with a helium flow of 10 mL/min for 5 min at 28 °C in order to remove moisture and remained oxygen. Cartridges were directly desorbed as described in Section 2.3.1 for SBSE and HSSE. The volatiles in the CIS were desorbed in splitless mode for the first 4 min after which a split 1:40 was applied.

2.4 Gas-Chromatography Mass-Spectrometry (GC-MS)

For all four trapping techniques, the same instrument and settings were employed. All analyses were conducted on an Agilent GC7890A coupled to a 5975C quadrupole mass spectrometer. The column used was a Zebron ZB-5Msplus with dimensions 30m x 0.25mm x 1.00µm (Phenomenex). The column oven was temperature programmed starting at 45 °C for 2 min, then increased at a rate of 5 °C/min to 250 °C and then maintained at 250 °C for 5 min. The carrier gas was helium, at a flow of 1 mL/min. The column effluent was ionised by electron impact at 70 eV, in the scan range m/z 33–500. The interface temperature was 280°C.

2.5 Experimental procedure and data analysis

We analysed eight replicates for both Maxagusto products. The analysis order was kept the same for all the techniques. A series of n-alkanes were analysed at each sequence for calculating retention indices (RI) using a third order polynomial function.

After visual inspection of the GC-MS total ion current chromatograms using vendor software, raw data were processed using an untargeted metabolomics approach. Baseline correction and alignment of mass signals ($s/n \geq 3$) were performed using MetAlign software (Lommen 2009). Mass signals present in ≤4 replicates were discarded. Signal redundancy was removed and mass spectra were reconstructed using MSClust (Tikunov et al. 2012). Metabolites were identified by matching the mass spectra and retention indices to authentic reference standards or those in the NIST17 Mass Spectral library (v.2.3).

For statistical analysis, we compared and visualized the main tendencies of the generated data by principal components analysis (PCA) after log 10 transformation and Pareto scaling of the samples using SIMCA 15.0.2. software (Sartorius Stedim Data Analytics AB, Umeå, Sweden). Graphs were also produced using Microsoft Excel 365.

3. Results and discussion

This study was initiated to investigate the potential of four analytical techniques for trapping a range of different classes of volatiles present in PFs with contrasting chemical compositions. Maxagusto PFs were chosen for their diversity in chemical groups relevant for defining flavour and aroma profiles. The four techniques were chosen for their simplicity, easy manipulation and robustness in extracting a high range of volatile compounds directly from the headspace or liquid mixture without having to use toxic organic solvents (Nogueira 2015). Each technique was separately pre-optimized for the best trapping conditions in terms of comprehensiveness before the trapping procedures were compared. The parameters that were optimized are summarized in the Supplementary **Table S1**.

3.1 Untargeted volatile comparison of the four different techniques

Metabolomics aims to characterize comprehensively a broad range of small molecules in a biological sample. Most importantly, it helps to compare accurately the global metabolite profile between groups of samples and thus to identify discriminatory compounds. However, the method of extraction has a major influence on the range of metabolites detected. In this study, we aimed to demonstrate that the use of different trapping techniques delivers distinctly different profiles for volatile compounds of PFs, both in terms of comprehensiveness and repeatability. The results were first compared in an unbiased way by looking at the complete volatile compound spectrum using a metabolomics approach. Raw GC-MS data were processed using an untargeted workflow pipeline and processed data were tested for their repeatability and selectivity. For both Maxagusto types, SBSE revealed the highest number of compounds as compared to the other three trapping techniques (**Table 1**). SPME also revealed a high number of compounds, while DHS appeared to detect the smallest number of compounds which was not completely unexpected. With DHS, we experienced technical limitations due to water interfering with the trapping of volatiles when a sampling temperature of 60^0C was used, even after using very large purging volumes. Hence, 30°C was used and therefore it would be incorrect to compare results from the SPME technique directly with those from DHS. Nevertheless, we have decided to include the DHS findings due to the strong qualitative differences that were observed for certain sensory-relevant chemical groups. The repeatability represented by the coefficient of variation (CV) of all compounds was lower for the SBSE and SPME data compared to that of HSSE and DHS. HSSE appeared to deliver the least repeatable data. Our results confirm published data where SBSE, although not widely used, has demonstrated its effectiveness in trapping predominantly non-polar and semi-polar compounds from liquid food samples (Ochiai et al. 2018). SBSE has also delivered more comprehensive profiles than SPME in liquid matrices, such as wine (Caven-Quantrill and Buglass 2011) and coffee (Bicchi et al. 2002).

Table 1. Number of compounds detected by an untargeted study in Maxagusto samples G28 and S99 using the four techniques. Eight replicates were measured for the statistical analysis. SBSE: Stir-bar sorptive extraction; HSSE: Headspace sorptive extraction; SPME: Solid-phase microextraction; DHS: Dynamic headspace system.

	G28				S99			
	SBSE	**HSSE**	**SPME**	**DHS**	**SBSE**	**HSSE**	**SPME**	**DHS**
Number of compounds detected	158	73	122	68	164	89	134	76
Number of known compounds	51	31	44	27	59	46	49	36
Coefficient of variation (%)	1.66	18.60	2.99	3.38	3.57	10.35	4.40	7.45

Principal components analysis (PCA) was performed to visualise how the volatile composition differed between the four techniques (**Figure 2**). Based on all detected volatile compounds the first two PCs show a clear separation of the profiles of the four different techniques for both Maxagusto G28 (**Figure 2A**) and S99 (**Figure 2B**). The first two PCs explained more than 80% of the total variance for both Maxagusto types respectively. Variation of the samples for each approach was smaller than between techniques indicating

good repeatability. The biggest difference was found for SBSE in both sample types (G28 and S99) as PC1 separates these from all other techniques. Furthermore, SPME samples also seem to have a more distinct profile compared to the other methods. Interestingly, despite the different sorbent (PDMS vs Tenax) and means of trapping the volatiles (one being a dynamic and one a static method), there was a strong similarity in the volatile composition from DHS and HSSE. The reasons behind this are not yet clear, however, the variables involved could be related to the type of adsorbent phase used and the dimensions of the coating phase, as well as to the previously mentioned lower temperature used for DHS.

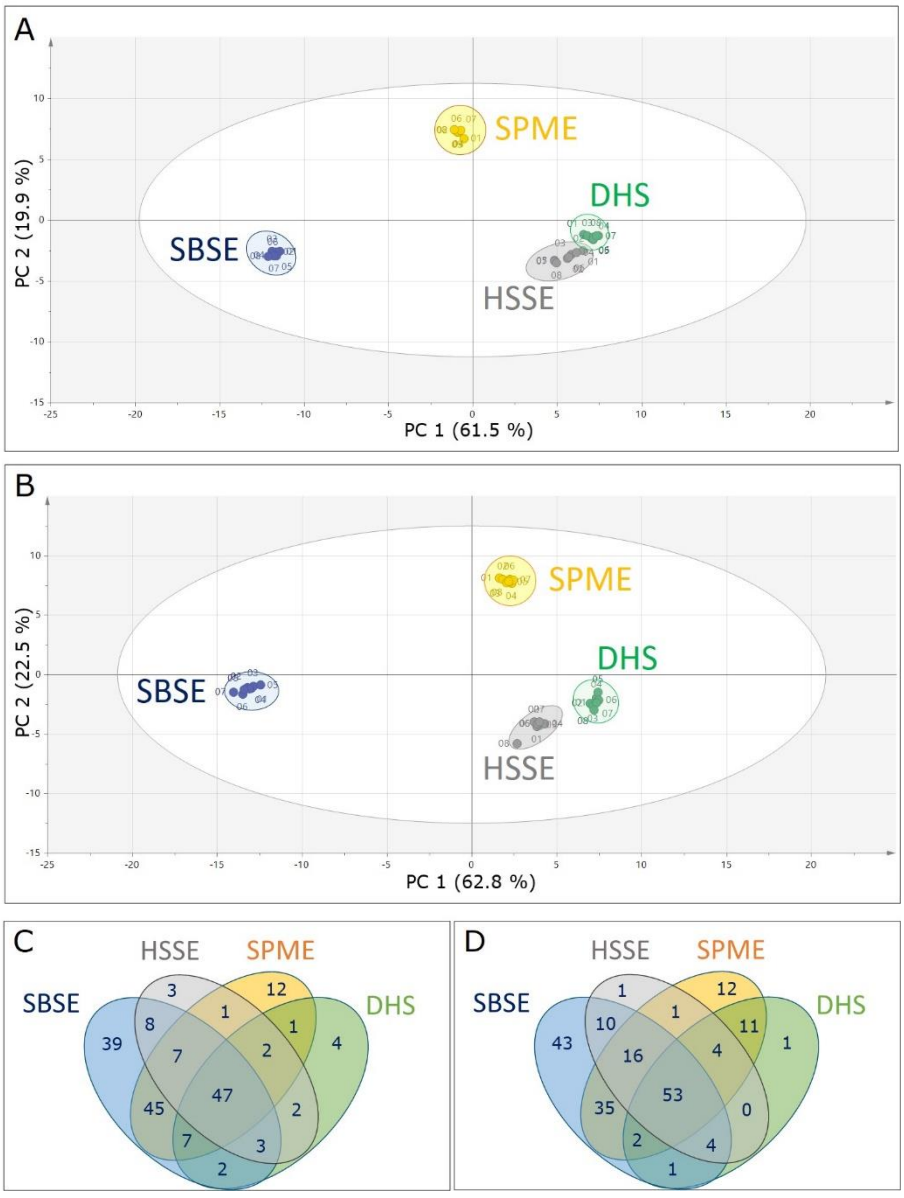


Figure 2. Principal components analysis (PCA) score plot of the volatile profiles of Maxagusto G28 **(A)** and S99 **(B)** using four trapping techniques. Eight replicates for each technique are represented. The first and second PC explain the corresponding percentage of variation shown on each axis. Venn diagram representing the total number of metabolites detected by SBSE, HSSE, SPME and DHS in Maxagusto samples G28 **(C)** and S99 **(D)**. SBSE (blue): Stir-bar sorptive extraction; HSSE (grey): Headspace sorptive extraction; SPME (orange): Solid-phase microextraction; DHS (green): Dynamic headspace. Compounds considered at least present in 5 out of the 8 replicates.

The differences between the techniques in the PCA were reflected by the overlap of compounds visualized by a Venn diagram showing the total number of volatiles detected for both Maxagusto samples (**Figure 2C** and **D**). In both samples, the subset of compounds commonly detected across all four techniques formed the largest group. In total, 47 compounds were commonly detected in G28 (**Figure 2C**) and 53 compounds in S99 (**Figure 2D**). However, SBSE also trapped an additional, almost equal number of compounds (39 and 43, respectively) that were unique for this technique while the other techniques uniquely trapped only between 1-12 additional metabolites. Both SPME and SBSE also shared many compounds that DHS and HSSE could not trap. As a conclusion, SBSE followed by SPME were the techniques giving the most comprehensive profiles in terms of numbers. However, qualitative differences should also be considered of importance as detailed further below.

3.2 Targeted volatile compounds trapped by SBSE, HSSE, SPME and DHS

To study the differences between the trapping techniques in more detail, we looked at the number of compounds identified in both G28 and S99 samples (Supplementary **Tables S2** and **S3**, respectively). Compounds were identified based on comparison of the retention index (RI) and the mass spectra of authentic reference standards or from commercial and in-house libraries including NIST. The level of identification given follows the guidelines of the Metabolomics Standards Initiative (Sumner et al. 2007). All compounds with no or a lower level of identification reliability (levels 3 and 4) were here considered as 'unknowns' (Supplementary **Table S4** and **S5**). For evaluation, the volatiles detected were divided into the following chemical groups: aromatics, polysulfides, pyrazines, aldehydes, sesquiterpenes/monoterpenes hydrocarbons/alcohols, and 'others'. Detailed analyses of the groups of identified compounds found in the different trapping profiles revealed that some metabolite classes are better represented by some trapping techniques than others. Polysulfides, pyrazines and aromatics were more abundant in G28, whereas S99 revealed primarily sesquiterpenes and monoterpenes as well as polysulfides and pyrazines. The total number of identified compounds was 55 for G28 and 60 for S99. To visualize the main trend of the identified compounds across the four techniques, a graph was made to show the relative contribution of each compound in each technique (**Figure 3**). A noticeable observation was the high contribution of SBSE to trapping pyrazines, as compared to the other techniques. This was observed for both Maxagusto G28 (**Figure 3A**) and S99 (**Figure 3B**) samples. DHS trapped a considerable number of monoterpene hydrocarbons whereas sesquiterpene hydrocarbons were more prominent in SPME data than in those of the other techniques. Sesquiterpene alcohols were trapped mostly by SBSE, as well as some other aromatics. The aromatics, aldehydes and the group 'others' appeared to show the greatest variation between compound regarding trapping across the four techniques. These clear differences are also important when making choices for analysing food samples that may be richer in specific chemical groups of compounds. Or when the purpose of the analysis focuses on characterizing just one class of compounds. In the following sections some specific observations for the different chemical groups are described.

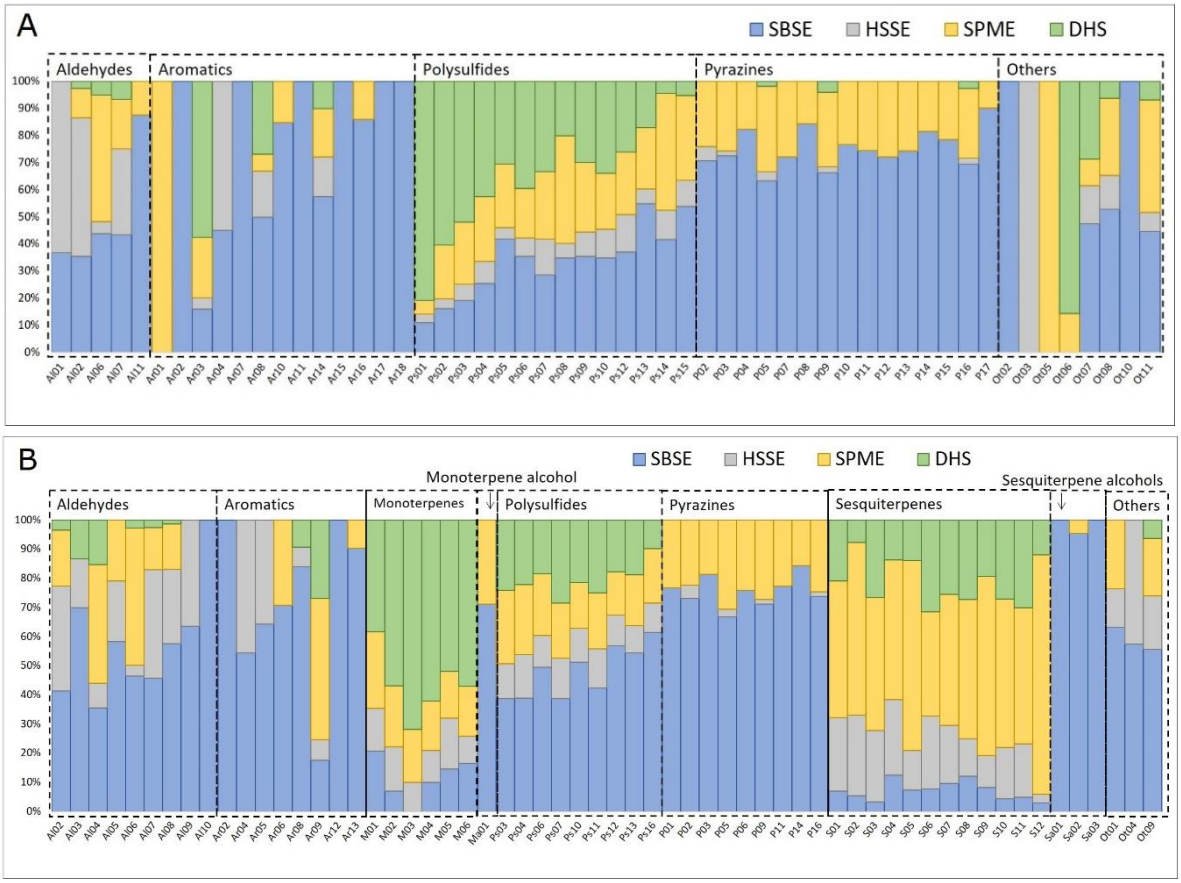


Figure 3. Relative proportion of identified metabolites trapped for each techniques grouped per compound class. **(A)** Maxagusto G28 and **(B)** Maxagusto S99. Metabolite names correspond to metabolites in Supplementary **Table S2** and **S3**.

3.2.1 Polysulfides

Polysulfides are important (savoury) components in food flavour research. They are characterized by having an alliaceous, sulphur, roasted garlic type flavour (www.Foodb.ca). In this study, a total of 14 and 9 polysulfides were identified in G28 and S99 samples respectively (Supplementary **Figure S1**). Seven were common to both samples. Sulfides or polysulfides (disulfides, trisulfides, etc.) with allyl- and methyl- groups were seen to be the most abundant S-containing compounds. Interestingly, most polysulfides were detected by all the techniques. However, the distribution in the four techniques is not the same for sample G28 and S99 which might be due to the absolute abundance of polysulfides. Polysulfides in G28 were 30 times more abundant than in S99 (Supplementary **Figure S2**) and, hence, a slightly different distribution between techniques was observed in both samples. In sample G28, polysulfides with low molecular weight (low RI) were more prominent when using DHS, whereas those with high molecular weight (MW) were more prominent when using SBSE. In the case of diallyl trisulfide (Ps14), highest levels were detected when using SPME. HSSE appeared least successful in trapping polysulfides. Overall, polysulfides with allyl- and methyl-groups were the most abundant sulphur compounds in Maxagusto PFs and all four techniques could be used for the analysis of this type of aroma compound. It should be noted that S-containing compounds can be

thermally labile and may oxidize to form polysulfides, thus, when analysing by GC-MS, artefact formation can be observed (Hofmann et al. 1996; Lestremau et al. 2004).

3.2.2 Pyrazines and other aromatics

The other class of compounds better represented when using SBSE is the aromatics. Aromatic heterocyclic compounds can be found in foods, often at low concentrations. However, they can be highly influential to the overall flavour by contributing to aroma complexity of food. Heterocyclic compounds have been strongly linked to 'roast meat' flavour formation during heating and are important compounds in processed foods and food flavouring (Diez-Simon et al. 2019). They are formed from many degradation pathways, the most important concerning Maillard reactions between amino acids and sugars. However, some can be formed from lipids or from lipid degradation products (Shahidi et al. 2014). This group of compounds is chemically diverse so the distribution between the techniques for the trapping of these volatiles is more varied (Supplementary **Figure S2**). Again, while SBSE appeared the best overall trapping method, other approaches, such as SPME or DHS, were sometimes better for individual molecules suggesting that trapping success is perhaps more structure-dependant within this diverse compound class.

Sulphur-containing heterocycles (1,2-dithiole, Ar08; and 2-vinyl-1,3-dithiine, Ar14) were the most abundant in G28 (Supplementary **Figures S2 C** and **D**). SBSE and DHS were the techniques that trapped the S-containing heterocycles more effectively. On the other hand, thiophenes and N-containing heterocycles, such as pyridines and pyrroles, were better trapped by SBSE with the exception of 3-methyl thiophene (Ar03) which was more prominent in DHS.

3.2.3 Terpenes

In plants, terpenoids (monoterpenes and sesquiterpenes) play important highly diverse roles in nature as plant hormones, defence compounds, insect / animal attractants and repellents, etc. (Bohlmann 2015). In flavour science, terpenoids are important when formulating new flavouring ingredients as they confer a wide range of aroma characters, such as sweet, herbal, spicy and woody (Behrens et al. 2011). A total of 6 monoterpene hydrocarbons, 1 monoterpene alcohol, 12 sesquiterpene hydrocarbons and 3 sesquiterpene alcohols were identified by one or more of the four trapping techniques (Supplementary **Figure S3**). In sample G28, only 1 terpene (p-Eugenol, Supplementary **Table S2**) was identified. In S99, the major terpene compound was the sesquiterpene caryophyllene (Supplementary **Table S3**) but this has for practical reasons been excluded from **Figures 3** and **S3** as its abundance was at least 10 times higher than the others and the peak was saturated. All monoterpene hydrocarbons observed were detected by all the techniques although DHS revealed relatively higher levels than SBSE, HSSE or SPME each of which showed similar results. The same was observed for other terpene compounds, which were better recovered by DHS as compared to SPME (Ochiai et al. 2012). One monoterpene alcohol (Ma01) was clearly visible using SBSE but was not detected when using HSSE and DHS. In the case of sesquiterpene hydrocarbons and sesquiterpene alcohols, all the techniques appeared able to trap all types. SPME was the most effective in trapping sesquiterpene hydrocarbons while SBSE was better (or the only one) able to trap sesquiterpene alcohols

than HSSE, SPME or DHS. Based on standard deviation values, SBSE appeared the most repeatable for monoterpenes and sesquiterpenes in sample S99. The high trapping efficiency of SBSE for sesquiterpene alcohols may be due to the high molecular weight (and hence are less volatile) and high polarity of these compounds. This could entail that they are poorly released into the headspace (Richter et al. 2017). The SBSE technique has previously been characterized for its effectiveness in trapping semi-polar and polar compounds which SPME and/or DHS are unable to do so (David et al. 2019). Moreover, modifications of the SBSE technique by e.g. pre-treating the stir bars with organic solvents provide interesting potential to broaden the trapping of more polar and less volatile compounds, as compared to normal SBSE, offering improvements for the analysis of food flavourings, including PFs. Important polar aroma compounds, such as short chain fatty acids (C3-C5), were detected by the modified SBSE method (Gilart et al. 2014; Macnamara et al. 2018; Ochiai et al. 2018; Picard et al. 2018). However, it also obscures other compounds which co-elute or interfere with the polar metabolites (Macnamara et al. 2018). Overall, all four techniques are able to trap most of the terpenoids observed. SBSE traps a considerable amount of sesquiterpene alcohols that other techniques cannot detect while SPME is able to trap a considerable amount of sesquiterpene hydrocarbons. Interestingly, DHS profiles were rich in monoterpene hydrocarbons making this technique the preferred choice when this group is of specific interest. Their MW is much lower than for sesquiterpenes, potentially making them more suitable for dynamic headspace techniques.

3.3 Combining headspace and in-solution trapping (SBSE^2)

HSSE seems, in general, the technique least suited for the volatile analysis of Maxagusto PFs. Nevertheless, for some aldehydes and aromatics it, together with SBSE, could make an important trapping technique when both are combined together. Therefore, a dual stir bar sorptive extraction (SBSE^2) was carried out to test the comprehensiveness of combining HSSE and SBSE in one desorption/chromatogram (Werkhoff et al. 2001). The analysis was done by placing two stir bars in the same vial, one located in the headspace above the liquid and the second stir bar immersed in the aqueous solution (**Figure 4**). After trapping the volatiles, both stir bars were placed in the same desorption glass tube and desorbed together onto the GC-MS to deliver a single chromatogram. Results showed that SBSE^2 extraction was able to trap all the volatiles that were extracted by SBSE and HSSE separately. A few examples of polysulfides, aldehydes, aromatics and pyrazines are shown in **Figure 4**. Compounds that are mostly (or only) trapped by one mode, such as pyrazines in SBSE, are altogether trapped in the dual mode. However, for the PFs analysed in this study, the number of compounds trapped only by HSSE was very small, showing no significant additional information. Thus, the combination of both methods, for this type of PFs, does not bring an improved level of comprehensiveness compared to individual SBSE or HSSE sampling. Nonetheless, performing SBSE^2 was proven to cover the range of volatiles that were more selective for both modes of trapping. Suggesting that a dual combination approach will increase the coverage of volatile metabolites in samples that contain a high diversity of compound classes.

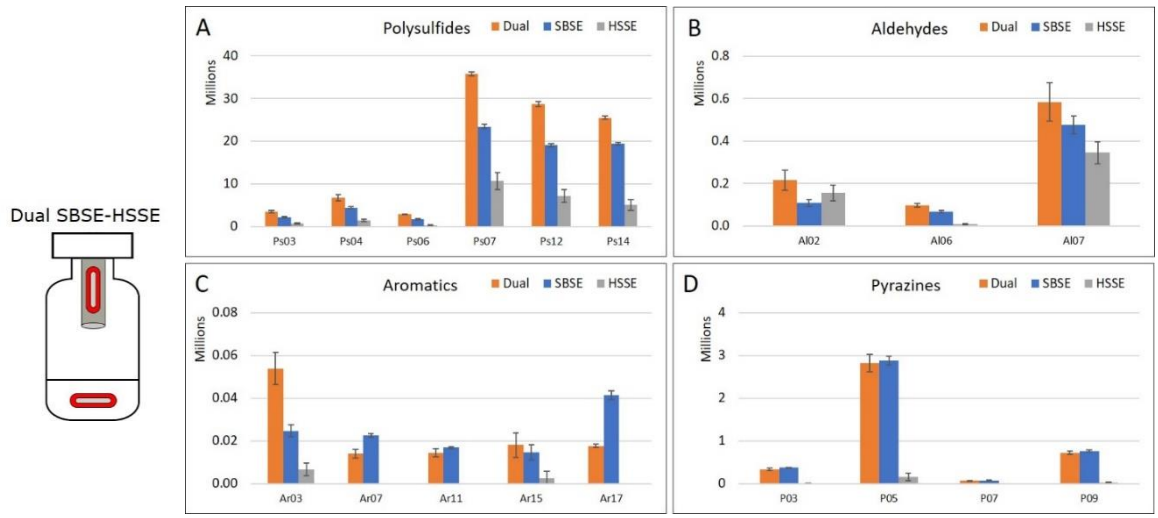


Figure 4. SBSE[2] sampling. A few distinguished compounds are shown as examples: Polysulfides **(A)**; Aldehydes **(B)**; Aromatics **(C)**; Pyrazines **(D)**. Abundance of those compounds is expressed in Total Ion Current (TIC) * 10^6, for the different trapping techniques (Dual, SBSE and HSSE). Mean and sd values of three replicates are also shown. Metabolite names correspond to metabolites in Supplementary **Table S2**.

3.4 Effect of salt addition in sample preparation

The addition of salt changes the physico-chemical properties of the sample solution which can sometimes help to release volatiles from the sample matrix into the gas phase when headspace techniques are used. Some salts, like NaCl, have been shown to enhance the release of certain volatile molecules from a liquid into the headspace (Blasco et al. 2004). In doing so, they increase the sensitivity of the analytical technique. However, in PFs, this effect might not be wholly beneficial for all chemical groups of interest which are present in the sample.

The total abundance of volatiles was analysed using SPME-GC-MS, after samples (Maxagusto G28 and S99) had been prepared with the addition of sodium chloride (NaCl), calcium chloride ($CaCl_2$) or just water. Polysulfide, pyrazine and aromatic groups were more abundant in G28 (**Figure 5A**), whereas S99 revealed primarily sesquiterpenes and monoterpenes as well as polysulfides and pyrazines (**Figure 5B**). The addition of either NaCl or $CaCl_2$ increased the abundance of most pyrazines, aromatics and aldehydes in the volatile profiles. Likewise, polysulfides were also slightly enhanced. However, the abundance of volatile sesquiterpenes and monoterpenes was observed to decrease when salts were added (**Figure 5**). This contrast might be explained by differences in the solubility and polarity between the different classes of compounds. The 'salting-out' effect is proposed to enrich for the more hydrophilic compounds (log $K_{O/W}$ < 3) in the headspace (Nogueira 2015). Pyrazines (and other aromatics) are highly water soluble compounds and, depending on the functional groups attached to them, their surface polarity can also be high. Consequently, their concentration in the headspace can be expected to increase in samples following salt addition. Terpenes, however, are apolar molecules that possess low water solubility. Consequently, their abundance in the headspace decreases when compared to solutions without salt. Monoterpenes are more water soluble than sesquiterpenes because of their lower MW and slightly higher solubility, and this might explain why sesquiterpene abundance in the headspace decreased significantly on the addition of both NaCl

and $CaCl_2$ whereas monoterpenes were much less affected (**Figure 5**). Polysulfides and aldehydes are also polar and soluble in water, but their solubility and polarity are lower than for pyrazines. Therefore, when salt was added, their abundance increased in both sample types but less dramatically.

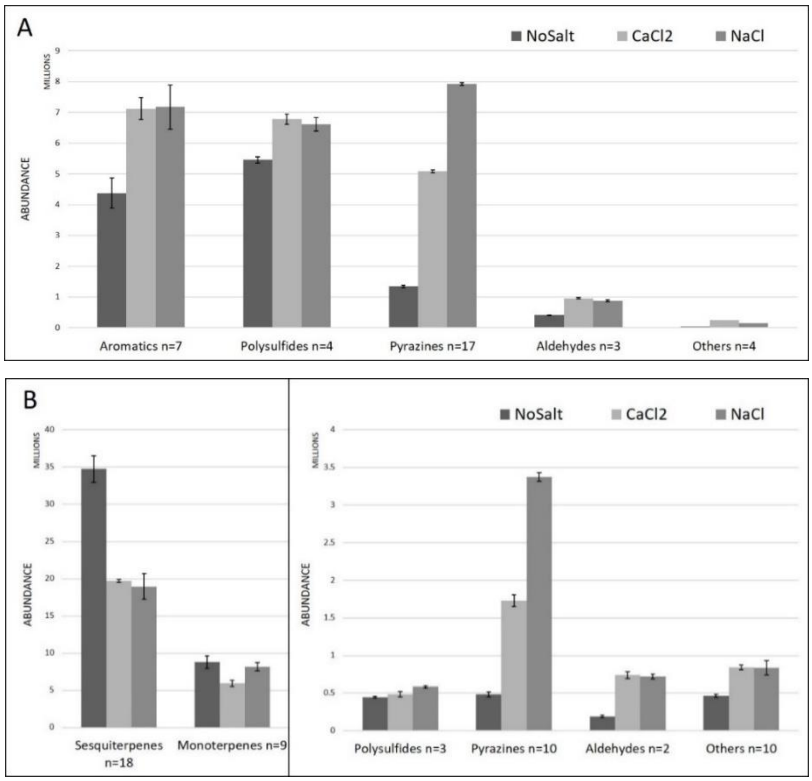


Figure 5. Effect of salt addition during sample preparation on the relative abundance of different volatile chemical groups, as detected for Maxagusto G28 **(A)** and S99 **(B)** respectively, analyzed by SPME-GC-MS. Mean and sd of five replicates per treatment are given. Addition of sodium chloride (NaCl; dark grey), calcium chloride ($CaCl_2$; light grey) or just water (NoSalt; black). N = number of identified metabolites found per group.

A difference between using $CaCl_2$ or NaCl was only observed for pyrazines. Furthermore, we observed the lowest standard deviations (sd) for the pyrazines and the highest for sesquiterpenes (**Figure 5**). This may relate to differences in the chemical structure or dynamic range of the compounds. There are no differences in the sd values when comparing samples without salt, with $CaCl_2$ or NaCl.

In conclusion, the addition of salt increases the abundance of volatiles in the headspace trapped by SPME, with the exception of terpene compounds. Adding salt primarily had more quantitative than qualitative influence on volatile profiles. These results demonstrate that the addition of salt(s), although increasing the abundance of certain compound groups might not necessarily be the method of choice for an untargeted analysis of headspace volatiles in food flavourings. In this study, salt was not used in the final methods due to its negative influence in the abundance of terpenes.

4. Conclusions

In this study, it was demonstrated that four different trapping approaches, SBSE, HSSE, SPME and DHS indeed provide different volatile profiles for the Maxagusto PFs used. Under the experimental conditions applied, SBSE proved to be most suitable in extracting volatiles such as polysulfides, pyrazines and terpene alcohols, and generally provided the broadest spectrum of compounds. Moreover, SBSE trapped a significant extra number of compounds absent from the other profiles. This suggests that SBSE would be the most convenient starting point for the comprehensive analysis of volatile compounds in similar food matrices. Moreover, modifications of the SBSE approach provides interesting potential to broaden the trapping of semi-polar and semi-volatiles offering beneficial choices for the analysis of food flavourings, including PFs. On the other hand, SPME and DHS techniques were the most successful in extracting sesquiterpenes and monoterpenes hydrocarbons, respectively. This entails that, should there be particular interest in this compound class, SPME (and DHS) would (also) be suitable for the analyses. Furthermore, both SPME and SBSE were the most repeatable techniques for generating data on the water soluble PFs used here.

Few studies have compared directly these kind of trapping techniques for their robustness and comprehensiveness. The comparison of these techniques is crucial when some parameters cannot be maintained constant, as each of the techniques differs in many properties. Here it can be concluded that the extraction method has a significant impact on the volatile profile of PFs, based on the volatility, solubility and polarity of the compounds targeted. In fact, this is a crucial step when the focus of the analysis is the contribution of these volatiles to the aroma and taste of PFs. Carefully weighed choices must therefore be made regarding the best combination of analytical procedures to employ. For broadest comprehensiveness more than one protocol might be needed depending on the chemical complexity of the specific samples to be characterised. For this study, a combination of SBSE and SPME would give the best result. However, this entails extra labour and input costs and therefore SBSE would be the individual method of choice.

Supplementary material

Tables S1. Combinations of the different parameters that were tested for the different trapping techniques (A: SBSE and HSSE; B: SPME; C: DHS). The final selection of the best method used for comparative purposes was based on the number of compounds detected and the reproducibility between replicates (comprehensiveness and reproducibility). The settings in bold were the final conditions used in this study, as described in the Materials and Methods.

Table S1 A

Technique	Sample preparation		Trapping method			Desorption GC-MS
	Sample concentration (mg/mL)	Sample Volume (mL)	Adsorbent type	Trapping time (min)	Trapping temperature (°C)	Desorption mode
SBSE	2	3	PDMS	120	RT	Split5
SBSE	**2**	**3**	**PDMS**	**10+80**	**60+RT**	**Split5**
SBSE	2	3	PDMS	60	RT	Split5
SBSE	50	3	PDMS	30	60	Split20
SBSE	50	3	PDMS	30	60	Split50
SBSE	2	3	PDMS	30	60	Split5
SBSE	2	3	PDMS	30	60	Splitless
SBSE	2	3	EGSilicone	120	RT	Split5
SBSE	2	3	Monotrap	30	60	Split20
SBSE	50	1	PDMS	30	60	Split20
HSSE	**2**	**3**	**PDMS**	**10+80**	**60+RT**	**Split5**

Table S1 B

Technique	Sample preparation			Trapping method				Desorption GC-MS
	Sample concentration (mg/mL)	Salt addition	pH	Adsorbent type	Incubation time (min)	Trapping time (min)	Trapping temperature (°C)	Desorption mode
SPME	2	Nothing	6.7	PDMS/DVB	10	20	50	Split 20
SPME	2	Nothing	6.7	PDMS/DVB/CAR	10	20	50	Split20
SPME	2	CaCl2	6.7	PDMS/DVB/CAR	10	20	50	Split20
SPME	2	NaCl	6.7	PDMS/DVB/CAR	10	20	50	Split20
SPME	2	Nothing	6.7	PDMS/DVB/CAR	10	20	50	Splitless
SPME	**2**	**Nothing**	**6.7**	**PDMS/DVB/CAR**	**10**	**20**	**60**	**Splitless**
SPME	2	CaCl2	6.7	PDMS/DVB/CAR	10	20	60	Splitless
SPME	2	NaCl	6.7	PDMS/DVB/CAR	10	20	60	Splitless
SPME	2	Nothing	6.7	PDMS/DVB/CAR	10	60	60	Splitless
SPME	2	Nothing	6.7	PDMS/DVB/CAR	10	60	30	Splitless
SPME	2	Nothing	6.7	PDMS/DVB/CAR	10	20	30	Splitless
SPME	2	Nothing	5.4	PDMS/DVB/CAR	10	20	60	Splitless
SPME	2	Nothing	7.9	PDMS/DVB/CAR	10	20	60	Splitless
SPME	2	Nothing	6.7	PDMS/DVB/CAR	20	20	60	Splitless
SPME	2	Nothing	6.7	PDMS/DVB/CAR	10	20	90	Splitless
SPME	0.4	NaCl	6.7	PDMS/DVB/CAR	10	20	50	Splitless
SPME	0.4	Nothing	6.7	PDMS/DVB/CAR	10	20	50	Splitless
SPME	10	Nothing	6.7	PDMS/DVB/CAR	10	20	50	Splitless
SPME	10	NaCl	6.7	PDMS/DVB/CAR	10	20	50	Splitless

Table S1 C

Technique	Sample preparation			Trapping method			Desorption GC-MS
	Sample concentration (mg/mL)	Sample Volume (mL)	Salt addition	Trapping time (min)	Trapping temperature (°C)	Drying phase time (min)	Split/splitless mode
DHS	50	1	Nothing	5	60	1	Split20
DHS	50	1	NaCl	5	60	1	Split20
DHS	50	1	CaCl2	5	60	1	Split20
DHS	2	1	Nothing	5	60	10	Splitless + Split40
DHS	2	1	Nothing	5	60	5	Splitless + Split40
DHS	2	1	Nothing	10	60	10	Splitless + Split40
DHS	2	1	Nothing	5	30	5	Splitless + Split40
DHS	2	1	Nothing	10	30	5	Splitless + Split40
DHS	2	1	Nothing	10	30	20	Splitless + Split40
DHS	**2**	**3**	**Nothing**	**10**	**30**	**5**	**Splitless + Split40**

Table S2. Volatile compounds identified in Maxagusto sample G28 using four trapping techniques. SBSE: Stir-bar sorptive extraction; HSSE: Headspace sorptive extraction; SPME: Solid-phase microextraction; DHS: Dynamic headspace.

Name	Compound	CAS #	Molecular Formula	RI[a]	ID[b]	SBSE[c]		HSSE[c]		SPME[c]		DHS[c]	
						Mean (TIC)	C.V. %	Mean (TIC)	C.V. %	Mean (TIC)	C.V. %	Mean (TIC)	C.V. %
	Aldehydes												
Al01	2-Butenal, 3-methyl-	107-86-8	C5H8O	767	2	3105	87	5327	95	ND	ND	ND	ND
Al02	Hexanal	66-25-1	C6H12O	784	1	107023	14	155300	22	32423	6	8113	69
Al06	Benzeneacetaldehyde	122-78-1	C8H8O	1054	2	66625	7	6867	31	71121	9	7591	30
Al07	Nonanal	124-19-6	C9H18O	1112	1	474965	8	344713	14	199947	5	72248	18
Al11	2-Hexenal, 5-methyl-2-phenyl-	21834-92-4	C13H16O	1493	2	117251	2	ND	ND	16555	11	ND	ND
	Aromatics												
Ar01	Thiazole	288-47-1	C3H3NS	698	2	ND	ND	ND	ND	24454	11	ND	ND
Ar02	Pyridine	110-86-1	C5H5N	727	2	26624	14	ND	ND	ND	ND	ND	ND
Ar03	Thiophene, 3-methyl-	616-44-4	C5H6S	770	2	24812	11	6768	42	34645	8	89649	26
Ar04	Furfural	98-01-1	C5H4O	819	1	15817	18	19305	22	ND	ND	ND	ND
Ar07	Pyridine, 3-ethyl-	536-78-7	C7H9N	959	2	22544	4	ND	ND	ND	ND	ND	ND
Ar08	3H-1,2-Dithiole	288-26-6	C3H4S2	965	2	17720919	3	6009782	18	2185074	7	9579876	2
Ar10	Pyridine, 2-ethyl-5-methyl-	18113-81-0	C8H11N	1032	2	113141	3	ND	ND	20419	3	ND	ND
Ar11	Thiazole, 4-ethyl-5-methyl-	52414-91-2	C6H9NS	1045	2	16938	3	ND	ND	ND	ND	ND	ND
Ar14	4H-1,3-Dithiine, 2-vinyl-	80028-57-5	C6H8S2	1239	2	5002583	2	1267620	21	1556568	14	878283	15
Ar15	Benzene, hexyl-	1077-16-3	C12H18	1274	2	14703	23	ND	ND	ND	ND	ND	ND
Ar16	Indane-4-carboxaldehyde	51932-70-8	C10H10O	1283	2	66962	5	ND	ND	11045	3	ND	ND
Ar17	Pyridine, 3-phenyl-	1008-88-4	C11H9N	1481	2	41357	5	ND	ND	ND	ND	ND	ND
Ar18	2-Furancarbothioamide, N-butyl-	42383-72-2	C9H13NOS	1584	2	9036	4	ND	ND	ND	ND	ND	ND
	Polysulfides												
Ps01	Sulfide, allyl methyl	10152-76-8	C4H8S	689	2	195954	15	55844	21	87712	32	1436976	11
Ps02	Disulfide, dimethyl	624-92-0	C2H6S2	735	1	89692	12	19863	21	109396	11	333309	20
Ps03	Sulfide, diallyl	592-88-1	C6H10S	851	2	2166577	6	686098	22	2602508	6	5892863	14
Ps04	Disulfide, methyl allyl	2179-58-0	C4H8S2	917	2	4433639	5	1460936	19	4162740	5	7451026	11
Ps05	Disulfide, methyl 1-propenyl, I-	23838-19-9	C4H8S2	941	2	155235	6	15515	54	86672	8	112862	20
Ps06	Trisulfide, dimethyl	3658-80-8	C2H6S3	980	1	1743406	5	331674	26	894440	5	1942911	9
Ps07	Disulfide, diallyl	2179-57-9	C6H10S2	1095	1	23407353	2	1066763 2	18	20387338	3	2718424 4	3
Ps08	Disulfide, 1-propenyl propyl, I-	23838-21-3	C6H12S2	1107	2	67924	4	10768	27	77840	6	39139	5
Ps09	Disulfide, allyl 1-propenyl, I-	122156-03-0	C6H10S2	1109	2	438418	4	111280	26	319715	6	370544	5
Ps10	Disulfide, allyl 1-propenyl, (Z)-	122156-02-9	C6H10S2	1115	2	1800833	3	546513	24	1071698	6	1758206	4
Ps12	Trisulfide, allyl methyl	34135-85-8	C4H8S3	1159	2	19040307	2	7202967	19	11852300	4	1346722 5	5
Ps13	Trisulfide, methyl propyl	17619-36-2	C4H10S3	1172	1	158357	7	15706	46	65395	9	49202	15
Ps14	Trisulfide, diallyl	2050-87-5	C6H10S3	1322	2	19416377	1	5068210	23	20085902	4	2058869	27
Ps15	Trisulfide, allyl propyl	33922-73-5	C6H12S3	1334	2	1268444	3	226629	32	735809	4	123731	12
	Pyrazines												
P02	Pyrazine, 2,5-dimethyl-	123-32-0	C6H8N2	908	1	344104	2	24811	63	117731	4	ND	ND
P03	Pyrazine, 2-ethyl-6-methyl-	13925-03-6	C7H10N2	1002	2	374822	3	9002	48	132855	2	ND	ND
P04	Pyrazine, 2-ethenyl-6-methyl-	13925-09-2	C7H8N2	1023	2	45141	2	ND	ND	9673	11	ND	ND
P05	Pyrazine, 3-ethyl-2,5-dimethyl-	13360-65-1	C8H12N2	1084	2	2883461	3	162074	51	1435559	2	79988	53
P07	Pyrazine, 2-ethyl-3,5-dimethyl-	13360-65-1	C8H12N2	1092	2	77680	4	ND	ND	30032	4	ND	ND
P08	Pyrazine, 2-methyl-5-propyl-	29461-03-8	C8H12N2	1105	2	94675	6	ND	ND	17693	1	ND	ND
P09	Pyrazine, 3,5-diethyl-2-methyl-	18138-05-1	C9H14N2	1164	2	765996	3	25124	37	316867	2	47089	27
P10	Pyrazine, 2,5-dimethyl-3-propyl-	18433-97-1	C9H14N2	1167	2	148891	4	ND	ND	45313	10	ND	ND
P11	Pyrazine, 2,5-dimethyl-3-(2-methylpropyl)-	32736-94-0	C10H16N2	1209	2	73556	3	ND	ND	25147	2	ND	ND
P12	Pyrazine, 2,3-dimethyl-5-(1-methylpropyl)-	32263-00-6	C10H16N2	1242	2	71552	8	ND	ND	27604	13	ND	ND
P13	Pyrazine, 2,3,5-trimethyl-6-propyl-	92233-82-4	C10H16N2	1250	2	185271	3	ND	ND	63790	1	ND	ND
P14	Pyrazine, 2-isoamyl-6-methyl-	91010-41-2	C10H16N2	1261	2	128338	2	ND	ND	29135	3	ND	ND
P15	Pyrazine, 2,5-dimethyl-3-(2-methylbutyl)-	72668-36-1	C11H18N2	1310	2	118206	4	ND	ND	32424	5	ND	ND
P16	Pyrazine, 2,5-dimethyl-3-(3-methylbutyl)-	18433-98-2	C11H18N2	1321	2	948640	1	31082	21	349966	2	35987	16
P17	Pyrazine, 2-(3-methylbutyl)-3,5-dimethyl	111150-30-2	C11H18N2	1499	2	42104	2	ND	ND	4608	38	ND	ND
	Others												
Ot02	1-Pyrrolidinamine, N-ethylidene-	60144-27-6	C6H12N2	1017	2	19768	24	ND	ND	ND	ND	ND	ND
Ot03	3-Cyclohexene-1-carbonitrile	100-45-8	C7H9N	1028	2	ND	ND	47529	43	ND	ND	ND	ND
Ot05	Dithioacetate, allyl	27249-83-8	C5H8S2	1097	2	ND	ND	ND	ND	8916	7	ND	ND
Ot06	Decane, 3,7-dimethyl-	17312-54-8	C12H26	1207	2	ND	ND	ND	ND	10913	14	64677	28
Ot07	1,2-Dithi-4-ene, 3-ethenyl-	62488-52-2	C6H8S2	1214	2	6685619	2	2001534	20	1384673	11	4046218	5
Ot08	4H-1,2,3-Trithiine	290-30-2	C3H4S3	1230	2	2906596	7	690083	24	1562128	18	346139	10
Ot10	p-Eugenol	97-53-0	C10H12O2	1363	1	80199	4	ND	ND	ND	ND	ND	ND
Ot11	1,2,3,4-Tetrathiane, 5-methyl-	116664-30-3	C3H6S4	1406	2	5556244	4	872365	25	5159901	6	845029	8

[a] Calculated retention index (RI)

[b] Level of identification (Level of ID) (Sumner et al. 2007): 1, identification using internal standards; 2, identification by coincidence with the RI and mass spectrum of standards from the NIST library; 3, identification by coincidence with only mass spectrum of standards from the NIST library; 4, tentatively identification by mass spectrum.

[c] Abundance expressed as the mean of the total ion current (TIC) of eight replicates and the Coefficient of Variation (CV) expressed in percentage (%).

Table S3. Volatile compounds identified in Maxagusto sample S99 using four trapping techniques. SBSE: Stir-bar sorptive extraction; HSSE: Headspace sorptive extraction; SPME: Solid-phase microextraction; DHS: Dynamic headspace.

Name	Compound	CAS #	Molecular Formula	RI[a]	ID[b]	SBSE[c]		HSSE[c]		SPME[c]		DHS[c]	
						Mean (TIC)	C.V. %	Mean (TIC)	C.V. %	Mean (TIC)	C.V. %	Mean (TIC)	C.V. %
	Aldehydes												
Al02	Hexanal	66-25-1	C6H12O	784	1	295822	42	255887	82	136995	2	25120	9
Al03	2-Heptenal, I-	18829-55-5	C7H12O	957	2	45044	16	10800	106	ND	ND	8559	13
Al04	Benzaldehyde	100-52-7	C7H6O	967	2	379797	5	90237	6	432242	2	163021	10
Al05	Octanal	124-13-0	C8H16O	1006	1	361324	21	128689	29	129458	5	ND	ND
Al06	Benzeneacetaldehyde	122-78-1	C8H8O	1054	2	101381	8	7979	22	102046	5	6079	5
Al07	Nonanal	124-19-6	C9H18O	1112	1	442594	34	359031	20	138179	13	25786	23
Al08	Decanal	112-31-2	C10H20O	1215	2	433288	8	191630	20	116780	4	10348	5
Al09	2-Decenal, (Z)-	2497-25-8	C10H18O	1271	2	141805	75	81126	165	ND	ND	ND	ND
Al10	Benzeneacetaldehyde, alpha-ethylidene-	4411-89-6	C10H10O	1283	2	21251	8	ND	ND	ND	ND	ND	ND
	Aromatics												
Ar02	Pyridine	110-86-1	C5H5N	727	2	22030	31	ND	ND	ND	ND	ND	ND
Ar04	Furfural	98-01-1	C5H4O2	819	1	27088	53	22586	42	ND	ND	ND	ND
Ar05	Furfuryl alcohol	98-00-0	C5H6O2	839	2	36247	67	20120	67	ND	ND	ND	ND
Ar06	Furan, 2-acetyl-	1192-62-7	C6H6O2	903	2	18713	14	ND	ND	7688	3	ND	ND
Ar08	3H-1,2-Dithiole	288-26-6	C3H4S2	965	2	106292	12	8496	44	ND	ND	11772	24
Ar09	Benzene, tert-butyl-	98-06-6	C10H14	1028	2	8597	17	3408	78	23562	15	13127	14
Ar12	Thiophene-2-carboxaldehyde	98-03-3	C5H4OS	1060	2	4321	78	ND	ND	ND	ND	ND	ND
Ar13	Pyrrole, 2-acetyl-	1072-83-9	C6H7NO	1067	2	59047	7	ND	ND	6321	6	ND	ND
	Monoterpenes												
M01	p-Cymene	99-87-6	C10H14	1035	2	239124	4	168060	11	300376	12	439007	9
M02	alpha-Pinene	80-56-8	C10H16	939	1	136373	8	285726	7	396138	13	1080131	7
M03	Camphene	79-92-5	C10H16	959	2	ND	ND	8685	13	15674	24	62020	13
M04	gamma-Terpinene	99-85-4	C10H16	1016	2	200159	13	218626	19	334139	13	1233334	11
M05	D-Limonene	5989-27-5	C10H16	1040	1	3120978	7	3717728	8	3426914	12	11054032	4
M06	alpha-Terpinolene	586-62-9	C10H16	1099	1	124289	12	69499	26	128100	15	427463	14
	Monoterpene alcohol												
Ma01	alpha-Terpineol	98-55-5	C10H18O	1211	1	71793	11	ND	ND	29116	7	ND	ND
	Polysulfides												
Ps03	Sulfide, diallyl	592-88-1	C6H10S	852	2	54737	10	17041	23	35575	19	34017	9
Ps04	Disulfide, methyl allyl	2179-58-0	C4H8S2	917	2	267183	7	101776	15	164317	15	151775	12
Ps06	Trisulfide, dimethyl	3658-80-8	C2H6S3	979	1	95084	6	21089	18	40566	18	35451	15
Ps07	Disulfide, diallyl	2179-57-9	C6H10S2	1094	1	810915	4	292100	15	394406	13	597506	8
Ps10	Disulfide, allyl 1-propenyl, (Z)-	122156-03-0	C6H10S2	1115	2	127268	4	28767	20	38688	16	53092	9
Ps11	Disulfide, dipropyl	629-19-6	C6H14S2	1123	2	140807	5	44063	20	63736	17	82430	10
Ps12	Trisulfide, allyl methyl	34135-85-8	C4H8S3	1159	2	727055	5	134701	18	190210	13	226615	4
Ps13	Trisulfide, methyl propyl	17619-36-2	C4H10S3	1172	1	178623	7	30746	21	56778	13	61839	5
Ps16	Trisulfide, dipropyl	6028-61-1	C6H14S3	1347	2	138388	5	22901	30	41846	8	22178	17
	Pyrazines												
P01	Pyrazine, methyl-	109-08-0	C5H6N2	811	2	24719	10	ND	ND	7483	4	ND	ND
P02	Pyrazine, 2,5-dimethyl-	123-32-0	C6H8N2	909	1	243643	5	14870	48	74396	5	ND	ND
P03	Pyrazine, 2-ethyl-6-methyl-	13360-64-0	C7H10N2	1002	2	166832	7	ND	ND	38352	4	ND	ND
P05	Pyrazine, 3-ethyl-2,5-dimethyl-	13360-65-1	C8H12N2	1084	2	971491	4	37399	35	444519	2	ND	ND
P06	Pyrazine, 2,6-diethyl-	13067-27-1	C8H12N2	1091	2	35277	9	ND	ND	11148	2	ND	ND
P09	Pyrazine, 3,5-diethyl-2-methyl-	18138-05-1	C9H14N2	1164	2	267936	3	6108	17	102720	2	ND	ND
P11	Pyrazine, 2,5-dimethyl-3-(2-methylpropyl)-	32736-94-0	C10H16N2	1209	2	95388	2	ND	ND	27906	3	ND	ND
P14	Pyrazine, 2-isoamyl-6-methyl-	91010-41-2	C10H16N2	1261	2	88455	5	ND	ND	16285	3	ND	ND
P16	Pyrazine, 2,5-dimethyl-3-(3-methylbutyl)-	18433-98-2	C11H18N2	1321	2	672951	2	14965	6	224014	3	ND	ND
	Sesquiterpenes												
S01	delta-Elemene	20307-84-0	C15H24	1351	2	333910	13	1184409	16	2209440	4	988735	17
S02	alpha-Cubebene	17699-14-8	C15H24	1363	2	7714	9	38435	21	82243	5	10826	23
S03	alpha-Copaene	3856-25-5	C15H24	1395	2	124152	17	889126	16	1655974	4	966056	17
S04	beta-Elemene	515-13-9	C15H24	1405	2	162449	7	337552	19	620122	4	178030	21
S05	beta-cis-Caryophyllene	118-65-0	C15H24	1426	2	221661	10	397684	24	1930094	12	409937	18
S06	beta-Caryophyllene	87-44-5	C15H24	1443	2	4225197	9	1340708	15	1910325	2	16907747	10
S07	alpha-Caryophyllene	6753-98-6	C15H24	1478	2	360662	8	743015	19	1675206	3	947810	15
S08	alpha-Curcumene	644-30-4	C15H22	1490	2	414421	5	432195	21	1615053	2	924447	15
S09	alpha-Zingiberene	495-60-3	C15H24	1503	2	69754	6	90244	29	512593	9	160591	36
S10	beta-Selinene	17066-67-0	C15H24	1511	2	70940	14	281845	21	812179	2	431246	19
S11	alpha-Selinene	473-13-2	C15H24	1516	2	102265	14	377199	21	960440	3	616232	19
S12	cis-Calamenene	483-77-2	C15H22	1536	2	8796	5	8957	38	237513	3	34417	37
	Sesquiterpene alcohols												
Sa01	Caryophyllene oxide	1139-30-6	C15H24O	1606	2	759727	3	ND	ND	ND	ND	ND	ND
Sa02	Isospathulenol	88395-46-4	C15H24O	1643	2	951805	2	ND	ND	45282	8	ND	ND
Sa03	Xanthorrhizol	30199-26-9	C15H22O	1749	2	26770	12	ND	ND	ND	ND	ND	ND
	Others												
Ot01	5-Hepten-2-one, 6-methyl-	110-93-0	C8H14O	984	2	104514	11	21683	15	38846	13	ND	ND
Ot04	1-Hexanol, 2-ethyl-	104-76-7	C8H18O	1031	2	34322	16	25426	29	ND	ND	ND	ND
Ot09	2-Undecanone	112-12-9	C11H22O	1299	2	88793	3	29063	74	31291	8	10058	17

[a] Calculated retention index (RI)

[b] Level of identification (Level of ID) (Sumner et al. 2007): 1, identification using internal standards; 2, identification by coincidence with the RI and mass spectrum of standards from the NIST library; 3, identification by coincidence with only mass spectrum of standards from the NIST library; 4, tentatively identification by mass spectrum.

[c] Abundance expressed as the mean of the total ion current (TIC) of eight replicates and the Coefficient of Variation (CV) expressed in percentage (%).

Table S4. Total number of volatile compounds detected in Maxagusto sample G28 using four trapping techniques. For all codes see Supplementary **Table S2**.

Name	Compound	Molecular Formula	Group	RI[a]	Level of ID[b]	SBSE[c]		HSSE[c]		SPME[c]		DHS[c]	
						Mean (TIC)	C.V. (%)	Mean (TIC)	C.V. (%)	Mean (TIC)	C.V. (%)	Mean (TIC)	C.V. (%)
U01	Unknown			614	3	25859	7	53117	16	ND	ND	151706	41
U02	Unknown			622	4	37641	53	ND	ND	ND	ND	ND	ND
U03	Unknown			654	3	ND	ND	ND	ND	87442	41	ND	ND
U04	Unknown			660	4	15411	4	ND	ND	78819	40	ND	ND
U05	Unknown			661	3	ND	ND	ND	ND	ND	ND	22076	78
U06	Unknown			662	4	103193	5	13326	30	14339	35	ND	ND
U07	Unknown			671	3	ND	ND	24736	12	ND	ND	ND	ND
U08	Unknown			686	4	ND	ND	ND	ND	19550	77	ND	ND
Ps01	Sulfide, allyl methyl	C4H8S	Polysulfide	689	2	195954	15	55844	21	87712	32	1436976	11
Ar01	Thiazole	C3H3NS	Aromatic	698	2	ND	ND	ND	ND	24454	11	ND	ND
Ar02	Pyridine	C5H5N	Aromatic	727	2	26624	14	ND	ND	ND	ND	ND	ND
U09	Unknown			731	4	3196	7	2929	9	3336	7	8178	44
Ps02	Disulfide, dimethyl	C2H6S2	Polysulfide	735	1	89692	12	19863	21	109396	11	333309	20
U10	Unknown			752	3	13135	5	15508	4	6598	10	31410	23
Al01	2-Butenal, 3-methyl-	C5H8O	Aldehyde	767	2	3105	87	5327	95	ND	ND	ND	ND
U11	Unknown			769	4	14622	9	5562	40	23502	2	4017	96
Ar03	Thiophene, 3-methyl-	C5H6S	Aromatic	770	2	24812	11	6768	42	34645	8	89649	26
Al02	Hexanal	C6H12O	Aldehyde	784	1	107023	14	155300	22	32423	6	8113	69
Ar04	Furfural	C5H4O	Aromatic	819	1	15817	18	19305	22	ND	ND	ND	ND
U12	Unknown			836	3	199162	5	112300	13	202321	2	141977	22
U13	Unknown			839	4	235144	7	55347	26	ND	ND	2185	39
U14	Unknown			839	4	2708	5	2060	13	2950	2	22683	13
Ps03	Sulfide, diallyl	C6H10S	Polysulfide	851	2	2166577	6	686098	22	2602508	6	5892863	14
U15	Unknown			865	4	47901	5	13972	32	33212	5	33785	39
U16	Unknown			886	4	2993	5	ND	ND	19199	9	ND	ND
U17	Unknown			888	3	ND	ND	26310	12	13734	32	ND	ND
U18	Unknown			896	4	61555	13	49337	21	17087	7	ND	ND
P02	Pyrazine, 2,5-dimethyl-	C6H8N2	Pyrazine	908	1	344104	2	24811	63	117731	4	ND	ND
U19	Unknown			910	3	25828	4	ND	ND	17555	9	ND	ND
Ps04	Disulfide, methyl allyl	C4H8S2	Polysulfide	917	2	4433639	5	1460936	19	4162740	5	7451026	11
U20	Unknown			925	4	20195	9	ND	ND	26580	14	4103	22
U21	Unknown			929	3	16924	2	ND	ND	ND	ND	ND	ND
U22	Unknown			931	4	ND	ND	ND	ND	20581	25	ND	ND
U23	Unknown			932	4	15553	9	ND	ND	9087	15	9203	27
Ps05	Disulfide, methyl 1-propenyl, I-	C4H8S2	Polysulfide	941	2	155235	6	15515	54	86672	8	112862	20
U24	Unknown			957	4	11890	15	ND	ND	ND	ND	ND	ND
Ar07	Pyridine, 3-ethyl-	C7H9N	Aromatic	959	2	22544	4	ND	ND	ND	ND	ND	ND
Ar08	3H-1,2-Dithiole	C3H4S2	Aromatic	965	2	17720919	3	6009782	18	2185074	7	9579876	2
U25	Unknown			966	3	ND	ND	ND	ND	72166	9	ND	ND
U26	Unknown			966	4	826854	5	513127	17	113937	10	1533229	3
U27	Unknown			974	4	18559	10	9268	28	3870	17	3730	63
Ps06	Trisulfide, dimethyl	C2H6S3	Polysulfide	980	1	1743406	5	331674	26	894440	5	1942911	9
U28	Unknown			985	4	17303	21	8562	38	ND	ND	ND	ND
U29	Unknown			992	4	102005	3	20120	16	60716	3	39927	19
U30	Unknown			995	4	20528	7	28325	14	15861	9	41449	21
U31	Unknown			995	4	32366	6	2220	39	7849	18	5095	11
U32	Unknown			995	3	ND	ND	55339	15	25282	15	77827	23
P03	Pyrazine, 2-ethyl-6-methyl-	C7H10N2	Pyrazine	1002	2	374822	3	9002	48	132855	2	ND	ND
U33	Unknown			1006	4	347930	5	106177	13	145372	1	13001	69
U34	Unknown			1012	4	ND	ND	ND	ND	68517	22	ND	ND
Ot02	1-Pyrrolidinamine, N-ethylidene-	C6H12N2	Other	1017	2	19768	24	ND	ND	ND	ND	ND	ND
P04	Pyrazine, 2-ethenyl-6-methyl-	C7H8N2	Pyrazine	1023	2	45141	2	ND	ND	9673	11	ND	ND
Ot03	3-Cyclohexene-1-carbonitrile	C7H9N	Other	1028	2	ND	ND	47529	43	ND	ND	ND	ND
Ar10	Pyridine, 2-ethyl-5-methyl-	C8H11N	Aromatic	1032	2	113141	3	ND	ND	20419	3	ND	ND
U35	Unknown			1040	4	56771	33	23185	27	20306	13	4694	84
Ar11	Thiazole, 4-ethyl-5-methyl-	C6H9NS	Aromatic	1045	2	16938	3	ND	ND	ND	ND	ND	ND
Al06	Benzeneacetaldehyde	C8H8O	Aldehyde	1054	2	66625	7	6867	31	71121	9	7591	30
U36	Unknown			1065	3	5813	43	ND	ND	ND	ND	ND	ND
U37	Unknown			1068	3	40023	14	ND	ND	ND	ND	ND	ND
U38	Unknown			1076	4	10280	14	ND	ND	1610	38	ND	ND
P05	Pyrazine, 3-ethyl-2,5-dimethyl-	C8H12N2	Pyrazine	1084	2	2883461	3	162074	51	1435559	2	79988	53
U39	Unknown			1087	4	25547	7	ND	ND	16458	8	ND	ND
P07	Pyrazine, 2-ethyl-3,5-dimethyl-	C8H12N2	Pyrazine	1092	2	77680	4	ND	ND	30032	4	ND	ND
Ps07	Disulfide, diallyl	C6H10S2	Polysulfide	1095	1	23407353	2	10667632	18	20387338	3	27184244	3
U40	Unknown			1096	3	25620	3	ND	ND	9931	5	9215	5
Ot05	Dithioacetate, allyl	C5H8S2	Other	1097	2	ND	ND	ND	ND	8916	7	ND	ND
U41	Unknown			1101	4	13786	12	ND	ND	ND	ND	ND	ND
P08	Pyrazine, 2-methyl-5-propyl-	C8H12N2	Pyrazine	1105	2	94675	6	ND	ND	17693	1	ND	ND
Ps08	Disulfide, 1-propenyl propyl, I-	C6H12S2	Polysulfide	1107	2	67924	4	10768	27	77840	6	39139	5
U42	Unknown			1108	3	310177	2	22870	22	162490	4	86353	6
Ps09	Disulfide, allyl 1-propenyl, I-	C6H10S2	Polysulfide	1109	2	438418	4	111280	26	319715	6	370544	5
Al07	Nonanal	C9H18O	Aldehyde	1112	1	474965	8	344713	14	199947	5	72248	18
Ps10	Disulfide, allyl 1-propenyl, (Z)-	C6H10S2	Polysulfide	1115	2	1800833	3	546513	24	1071698	6	1758206	4
U43	Unknown			1122	3	296960	3	85963	20	74311	10	346147	25
U44	Unknown			1142	4	7691	24	3243	17	77520	5	4644	37
Ps12	Trisulfide, allyl methyl	C4H8S3	Polysulfide	1159	2	19040307	2	7202967	19	11852300	4	13467225	5
U45	Unknown			1160	4	801866	4	478927	15	513391	4	1180378	6
P09	Pyrazine, 3,5-diethyl-2-methyl-	C9H14N2	Pyrazine	1164	2	765996	3	25124	37	316867	2	47089	27
P10	Pyrazine, 2,5-dimethyl-3-propyl-	C9H14N2	Pyrazine	1167	2	148891	4	ND	ND	45313	10	ND	ND
Ps13	Trisulfide, methyl propyl	C4H10S3	Polysulfide	1172	1	158357	7	15706	46	65395	9	49202	15
U46	Unknown			1174	4	20944	6	ND	ND	6033	16	ND	ND
U47	Unknown			1176	4	ND	ND	ND	ND	14131	59	ND	ND
U48	Unknown			1177	4	32860	7	ND	ND	8090	9	ND	ND
U49	Unknown			1179	3	18436	13	ND	ND	6759	38	ND	ND
U50	Unknown			1184	3	5501748	2	841774	22	3810421	11	1129751	6
U51	Unknown			1188	3	20389	1	ND	ND	5875	4	ND	ND
U52	Unknown			1189	4	8627	2	ND	ND	3801	3	ND	ND
U53	Unknown			1193	3	ND	ND	ND	ND	ND	ND	56621	12
U54	Unknown			1194	4	26818	2	ND	ND	5736	12	ND	ND

U55	Unknown			1196	3	4502	20	ND	ND	ND	ND	ND	ND
U56	Unknown			1198	3	ND	ND	ND	ND	10606	2	ND	ND
Ot06	Decane, 3,7-dimethyl-	C12H26	Other	1207	2	ND	ND	ND	ND	10913	14	64677	28
P11	Pyrazine, 2,5-dimethyl-3-(2-methylpropyl)-	C10H16N2	Pyrazine	1209	2	73556	3	ND	ND	25147	2	ND	ND
U57	Unknown			1211	3	31465	4	ND	ND	ND	ND	ND	ND
Ot07	1,2-Dithi-4-ene, 3-ethenyl-	C6H8S2	Other	1214	2	6685619	2	2001534	20	1384673	11	4046218	5
U58	Unknown			1222	3	ND	ND	ND	ND	50247	22	ND	ND
Ot08	4H-1,2,3-Trithiine	C3H4S3	Other	1230	2	2906596	7	690083	24	1562128	18	346139	10
U59	Unknown			1233	4	115062	4	ND	ND	26343	5	ND	ND
U60	Unknown			1236	4	29264	4	ND	ND	9804	2	ND	ND
Ar14	4H-1,3-Dithiine, 2-vinyl-	C6H8S2	Aromatic	1239	2	5002583	2	1267620	21	1556568	14	878283	15
P12	Pyrazine, 2,3-dimethyl-5-(1-methylpropyl)-	C10H16N2	Pyrazine	1242	2	71552	8	ND	ND	27604	13	ND	ND
U61	Unknown			1242	4	35857	22	ND	ND	28209	9	65203	15
U62	Unknown			1245	4	64840	5	6485	62	5506	41	6948	11
P13	Pyrazine, 2,3,5-trimethyl-6-propyl-	C10H16N2	Pyrazine	1250	2	185271	3	ND	ND	63790	1	ND	ND
U63	Unknown			1254	3	40836	4	ND	ND	11789	12	ND	ND
P14	Pyrazine, 2-isoamyl-6-methyl-	C10H16N2	Pyrazine	1261	2	128338	2	ND	ND	29135	3	ND	ND
U64	Unknown			1263	4	133545	7	5039	54	14059	15	ND	ND
U65	Unknown			1271	4	13922	37	28049	66	ND	ND	ND	ND
Ar15	Benzene, hexyl-	C12H18	Aromatic	1274	2	14703	23	ND	ND	ND	ND	ND	ND
Ar16	Indane-4-carboxaldehyde	C10H10O	Aromatic	1283	2	66962	5	ND	ND	11045	3	ND	ND
U66	Unknown			1286	4	12561	10	ND	ND	2314	30	ND	ND
U67	Unknown			1289	3	8846	3	ND	ND	ND	ND	ND	ND
U68	Unknown			1291	3	6652	4	ND	ND	ND	ND	ND	ND
U69	Unknown			1295	4	41165	4	ND	ND	14798	17	ND	ND
U70	Unknown			1297	3	7754	4	14600	25	51429	24	51084	17
U71	Unknown			1299	3	9154	4	ND	ND	ND	ND	ND	ND
U72	Unknown			1302	3	15365	2	ND	ND	ND	ND	ND	ND
U73	Unknown			1306	3	ND	ND	4280	87	17983	18	248132	23
P15	Pyrazine, 2,5-dimethyl-3-(2-methylbutyl)-	C11H18N2	Pyrazine	1310	2	118206	4	ND	ND	32424	5	ND	ND
U74	Unknown			1315	4	45793	17	ND	ND	7829	80	ND	ND
U75	Unknown			1318	4	218020	4	7033	70	ND	ND	ND	ND
P16	Pyrazine, 2,5-dimethyl-3-(3-methylbutyl)-	C11H18N2	Pyrazine	1321	2	948640	1	31082	21	349966	2	35987	16
Ps14	Trisulfide, diallyl	C6H10S3	Polysulfide	1322	2	1941637	1	5068210	23	2008590	4	2058869	27
U76	Unknown			1327	4	52321	2	ND	ND	13558	6	3482	11
U77	Unknown			1331	3	52969	3	ND	ND	10726	2	ND	ND
Ps15	Trisulfide, allyl propyl	C6H12S3	Polysulfide	1334	2	1268444	3	226629	32	735809	4	123731	12
U78	Unknown			1343	4	360973	8	60154	43	ND	ND	ND	ND
U79	Unknown			1346	4	71755	11	ND	ND	27802	8	ND	ND
U80	Unknown			1357	4	499007	7	89910	39	19034	19	ND	ND
U81	Unknown			1358	4	ND	ND	ND	ND	41397	26	ND	ND
U82	Unknown			1358	4	230028	1	ND	ND	104156	2	ND	ND
Ot10	p-Eugenol	C10H12O2	Other	1363	1	80199	4	ND	ND	ND	ND	ND	ND
U83	Unknown			1372	3	50919	3	ND	ND	17778	6	ND	ND
U84	Unknown			1384	3	42102	3	ND	ND	16719	9	ND	ND
U85	Unknown			1388	3	41138	11	ND	ND	ND	ND	ND	ND
U86	Unknown			1391	4	100879	3	ND	ND	42236	2	ND	ND
U87	Unknown			1395	3	27927	3	ND	ND	8324	2	ND	ND
U88	Unknown			1399	4	39256	9	ND	ND	ND	ND	ND	ND
U89	Unknown			1404	3	76166	2	ND	ND	ND	ND	ND	ND
Ot11	1,2,3,4-Tetrathiane, 5-methyl-	C3H6S4	Other	1406	2	5556244	4	872365	25	5159901	6	845029	8
U90	Unknown			1407	4	6267	2	ND	ND	ND	ND	ND	ND
U91	Unknown			1410	4	1081480	5	157505	32	965710	12	802549	15
U92	Unknown			1432	4	25107	10	ND	ND	3131	14	ND	ND
U93	Unknown			1434	3	31123	1	ND	ND	10009	8	ND	ND
U94	Unknown			1446	3	40980	2	ND	ND	ND	ND	ND	ND
U95	Unknown			1451	4	20948	13	3001	58	6372	5	ND	ND
U96	Unknown			1452	3	39342	5	ND	ND	18361	2	ND	ND
U97	Unknown			1460	3	65735	6	ND	ND	27993	4	ND	ND
U98	Unknown			1463	4	22086	4	ND	ND	4297	3	ND	ND
U99	Unknown			1466	4	12731	14	ND	ND	ND	ND	ND	ND
U100	Unknown			1472	4	28727	8	ND	ND	2824	5	ND	ND
U101	Unknown			1475	3	17001	50	ND	ND	ND	ND	ND	ND
Ar17	Pyridine, 3-phenyl-	C11H9N	Aromatic	1481	2	41357	5	ND	ND	ND	ND	ND	ND
Al11	2-Hexenal, 5-methyl-2-phenyl-	C13H16O	Aldehyde	1493	2	117251	2	ND	ND	16555	11	ND	ND
P17	Pyrazine, 2-(3-methylbutyl)-3,5-dimethyl	C11H18N2	Pyrazine	1499	2	42104	2	ND	ND	4608	38	ND	ND
U102	Unknown			1500	4	9273	69	15384	59	ND	ND	ND	ND
U103	Unknown			1507	3	95634	48	ND	ND	49287	47	12183	40
U104	Unknown			1528	4	12897	3	ND	ND	ND	ND	ND	ND
U105	Unknown			1549	3	48573	2	ND	ND	ND	ND	ND	ND
U106	Unknown			1563	4	3740	39	ND	ND	240735	9	ND	ND
U107	Unknown			1569	4	16359	8	ND	ND	3217	4	ND	ND
U108	Unknown			1579	3	19511	2	ND	ND	ND	ND	ND	ND
Ar18	2-Furancarbothioamide, N-butyl-	C9H13NOS	Aromatic	1584	2	9036	4	ND	ND	ND	ND	ND	ND
U109	Unknown			1590	4	ND	ND	19586	49	ND	ND	775871	44
U110	Unknown			1625	4	ND	ND	2461	13	ND	ND	15168	19
U111	Unknown			1626	4	10582	5	ND	ND	ND	ND	ND	ND
U112	Unknown			1629	3	68479	4	ND	ND	ND	ND	ND	ND
U113	Unknown			1649	4	ND	ND	ND	ND	ND	ND	39475	60
U114	Unknown			1661	4	13576	6	ND	ND	ND	ND	ND	ND
U115	Unknown			1667	3	54334	5	ND	ND	ND	ND	ND	ND
U116	Unknown			1682	4	241713	10	ND	ND	ND	ND	ND	ND
U117	Unknown			1693	3	166907	4	ND	ND	35218	19	6708	64
U118	Unknown			1696	3	ND	ND	6295	75	ND	ND	ND	ND
U119	Unknown			1732	4	35383	6	ND	ND	ND	ND	ND	ND
U120	Unknown			1784	4	24314	5	ND	ND	ND	ND	ND	ND
U121	Unknown			1810	3	26710	8	ND	ND	ND	ND	ND	ND
U122	Unknown			1845	3	29092	22	21728	35	ND	ND	95827	21
U123	Unknown			1845	4	ND	ND	ND	ND	ND	ND	30431	33
U124	Unknown			1871	3	45516	13	ND	ND	ND	ND	ND	ND
U125	Unknown			1955	4	38522	13	6256	84	ND	ND	ND	ND
U126	Unknown			1973	3	ND	ND	ND	ND	4486	79	ND	ND
U127	Unknown			1994	4	169657	5	ND	ND	ND	ND	2888	51
U128	Unknown			2000	4	171207	3	ND	ND	ND	ND	5179	53

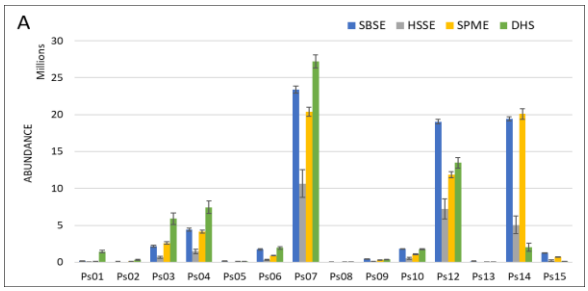
Table S5. Total number of volatile compounds detected in Maxagusto sample S99 using four trapping techniques. For all codes see Supplementary **Table S2**.

Name	Compound	Molecular Formula	Group	RI[a]	Level of ID[b]	SBSE[c]		HSSE[c]		SPME[c]		DHS[c]	
						Mean (TIC)	C.V %	Mean (TIC)	C.V %	Mean (TIC)	C.V %	Mean (TIC)	C.V %
Ar02	Pyridine	C5H5N	Aromatic	727	2	22030	31	ND	ND	ND	ND	ND	ND
Un01	Unknown			730	4	ND	ND	ND	ND	4901	38	12682	14
Al02	Hexanal	C6H12O	Aldehyde	784	1	295822	42	255887	82	136995	2	25120	9
Un02	Unknown			791	3	ND	ND	52752	17	34251	15	ND	ND
P01	Pyrazine, methyl-	C5H6N2	Pyrazine	811	2	24719	10	ND	ND	7483	4	ND	ND
Ar04	Furfural	C5H4O2	Aromatic	819	1	27088	53	22586	42	ND	ND	ND	ND
Ar05	Furfuryl alcohol	C5H6O2	Aromatic	839	2	36247	67	20120	67	ND	ND	ND	ND
Un03	Unknown			848	3	68035	8	31127	18	80395	4	12990	15
Ps03	Sulfide, diallyl	C6H10S	Polysulfide	852	2	54737	10	17041	23	35575	19	34017	9
Un04	Unknown			864	4	35901	4	11101	32	16383	30	ND	ND
Un05	Unknown			872	4	ND	ND	ND	ND	101832	39	ND	ND
Un06	Unknown			880	3	29863	16	10985	33	10891	11	ND	ND
Un07	Unknown			888	3	5588	82	31003	17	13689	13	ND	ND
Un08	Unknown			896	3	62331	61	44730	54	ND	ND	ND	ND
Un09	Unknown			902	3	13692	60	10370	37	ND	ND	ND	ND
Ar06	Furan, 2-acetyl-	C6H6O2	Aromatic	903	2	18713	14	ND	ND	7688	3	ND	ND
Un10	Unknown			903	3	247186	4	61850	10	150973	7	88149	11
P02	Pyrazine, 2,5-dimethyl-	C6H8N2	Pyrazine	909	1	243643	5	14870	48	74396	5	ND	ND
Ps04	Disulfide, methyl allyl	C4H8S2	Polysulfide	917	2	267183	7	101776	15	164317	15	151775	12
Un11	Unknown			924	3	ND	ND	ND	ND	8996	16	ND	ND
Un12	Unknown			933	4	30751	7	7939	16	22549	16	14502	24
M02	alpha-Pinene	C10H16	Monoterpene	939	1	136373	8	285726	7	396138	13	1080131	7
Un13	Unknown			942	4	72554	4	3559	38	15379	17	10759	25
Un14	Unknown			956	3	ND	ND	ND	ND	11168	66	ND	ND
Al03	2-Heptenal, I-	C7H12O	Aldehyde	957	2	45044	16	10800	106	ND	ND	8559	13
M03	Camphene	C10H16	Monoterpene	959	2	ND	ND	8685	13	15674	24	62020	13
Ar08	3H-1,2-Dithiole	C3H4S2	Aromatic	965	2	106292	12	8496	44	ND	ND	11772	24
Al04	Benzaldehyde	C7H6O	Aldehyde	967	2	379797	5	90237	6	432242	2	163021	10
Un15	Unknown			974	3	21042	33	12162	93	ND	ND	ND	ND
Ps06	Trisulfide, dimethyl	C2H6S3	Polysulfide	979	1	95084	6	21089	18	40566	18	35451	15
Ot01	5-Hepten-2-one, 6-methyl-	C8H14O	Other	984	2	104514	11	21683	15	38846	13	ND	ND
Un16	Unknown			988	3	598167	8	874967	8	727394	14	3255090	6
Un17	Unknown			993	4	118051	2	31658	12	64347	7	19407	21
Un18	Unknown			993	3	192807	8	271057	9	212263	15	914428	8
Un19	Unknown			996	4	ND	ND	43596	13	15499	18	94251	12
P03	Pyrazine, 2-ethyl-6-methyl-	C7H10N2	Pyrazine	1002	2	166832	7	ND	ND	38352	4	ND	ND
Al05	Octanal	C8H16O	Aldehyde	1006	1	361324	21	128689	29	129458	5	ND	ND
Un20	Unknown			1012	4	ND	ND	ND	ND	140828	21	ND	ND
M04	gamma-Terpinene	C10H16	Monoterpene	1016	2	200159	13	218626	19	334139	13	1233334	11
Un21	Unknown			1019	3	2186255	8	3511973	8	3469288	12	9963309	4
Un22	Unknown			1023	4	30532	7	3196	13	5733	13	ND	ND
Un23	Unknown			1026	3	ND	ND	ND	ND	13346	15	6456	12
Ar09	Benzene, tert-butyl-	C10H14	Aromatic	1028	2	8597	17	3408	78	23562	15	13127	14
Un24	Unknown			1028	4	ND	ND	10113	23	ND	ND	ND	ND
Ot04	1-Hexanol, 2-ethyl-	C8H18O	Other	1031	2	34322	16	25426	29	ND	ND	ND	ND
M01	p-Cymene	C10H14	Monoterpene	1035	2	239124	4	168060	11	300376	12	439007	9
M05	D-Limonene	C10H16	Monoterpene	1040	1	3120978	7	3717728	8	3426914	12	11054032	4
Un25	Unknown			1045	4	53199	5	6735	34	36636	9	3658	35
Al06	Benzeneacetaldehyde	C8H8O	Aldehyde	1054	2	101381	8	7979	22	102046	5	6079	5
Un26	Unknown			1056	3	ND	ND	ND	ND	17563	23	ND	ND
Ar12	Thiophene-2-carboxaldehyde	C5H4OS	Aromatic	1060	2	4321	78	ND	ND	ND	ND	ND	ND
Un27	Unknown			1065	4	14280	56	ND	ND	ND	ND	ND	ND
Ar13	Pyrrole, 2-acetyl-	C6H7NO	Aromatic	1067	2	59047	7	ND	ND	6321	6	ND	ND
Un28	Unknown			1069	4	40147	7	38382	8	28583	19	79888	10
Un29	Unknown			1075	3	12948	30	ND	ND	ND	ND	ND	ND
P05	Pyrazine, 3-ethyl-2,5-dimethyl-	C8H12N2	Pyrazine	1084	2	971491	4	37399	35	444519	2	ND	ND
P06	Pyrazine, 2,6-diethyl-	C8H12N2	Pyrazine	1091	2	35277	9	ND	ND	11148	2	ND	ND
Ps07	Disulfide, diallyl	C6H10S2	Polysulfide	1094	1	810915	4	292100	15	394406	13	597506	8
Un30	Unknown			1096	3	17937	9	ND	ND	ND	ND	ND	ND
M06	alpha-Terpinolene	C10H16	Monoterpene	1099	1	124289	12	69499	26	128100	15	427463	14
Un31	Unknown			1101	3	11509	4	ND	ND	ND	ND	ND	ND
Un32	Unknown			1102	3	5981	13	ND	ND	26146	13	ND	ND
Un33	Unknown			1104	4	40588	48	23259	50	ND	ND	ND	ND
Un34	Unknown			1108	4	339730	6	36858	9	238706	4	50478	13
Al07	Nonanal	C9H18O	Aldehyde	1112	1	442594	34	359031	20	138179	13	25786	23
Ps10	Disulfide, allyl 1-propenyl, (Z)-	C6H10S2	Polysulfide	1115	2	127268	4	28767	20	38688	16	53092	9
Ps11	Disulfide, dipropyl	C6H14S2	Polysulfide	1123	2	140807	5	44063	20	63736	17	82430	10
Un35	Unknown			1123	4	21697	8	8236	29	10133	17	27344	9
Un36	Unknown			1124	3	ND	ND	ND	ND	ND	ND	ND	ND
Un37	Unknown			1128	3	21991	8	ND	ND	ND	ND	ND	ND
Un38	Unknown			1131	4	31372	4	4090	57	11391	17	10300	22
Un39	Unknown			1132	3	7021	6	ND	ND	ND	ND	ND	ND
Ps12	Trisulfide, allyl methyl	C4H8S3	Polysulfide	1159	2	727055	5	134701	18	190210	13	226615	4
P09	Pyrazine, 3,5-diethyl-2-methyl-	C9H14N2	Pyrazine	1164	2	267936	3	6108	17	102720	2	ND	ND
Un40	Unknown			1168	3	174266	4	ND	ND	52516	5	ND	ND
Ps13	Trisulfide, methyl propyl	C4H10S3	Polysulfide	1172	1	178623	7	30746	21	56778	13	61839	5
Un41	Unknown			1174	3	9594	4	ND	ND	ND	ND	ND	ND
Un42	Unknown			1176	3	8764	6	ND	ND	ND	ND	ND	ND
Un43	Unknown			1183	3	82328	7	ND	ND	37850	16	ND	ND
Un44	Unknown			1192	3	20855	11	ND	ND	15458	3	ND	ND
Un45	Unknown			1199	4	8384	21	ND	ND	2746	4	ND	ND
Un46	Unknown			1208	3	ND	ND	ND	ND	7619	38	13457	25
P11	Pyrazine, 2,5-dimethyl-3-(2-methylpropyl)-	C10H16N2	Pyrazine	1209	2	95388	2	ND	ND	27906	3	ND	ND
Ma01	alpha-Terpineol	C10H18O	Monoterpene alcohol	1211	1	71793	11	ND	ND	29116	7	ND	ND
Al08	Decanal	C10H20O	Aldehyde	1215	2	433288	8	191630	20	116780	4	10348	5
Un47	Unknown			1222	3	ND	ND	ND	ND	31497	12	ND	ND
Un48	Unknown			1224	4	13707	58	ND	ND	38052	12	ND	ND
Un49	Unknown			1230	4	18764	10	ND	ND	2427	4	ND	ND
Un50	Unknown			1233	3	26670	9	ND	ND	3859	58	ND	ND
Un51	Unknown			1236	3	12185	3	ND	ND	ND	ND	ND	ND
Un52	Unknown			1238	4	12846	8	ND	ND	ND	ND	ND	ND
Un53	Unknown			1245	3	22596	3	ND	ND	ND	ND	ND	ND
Un54	Unknown			1246	4	214566	13	ND	ND	ND	ND	ND	ND
Un55	Unknown			1257	4	47648	8	ND	ND	ND	ND	ND	ND
P14	Pyrazine, 2-isoamyl-6-methyl-	C10H16N2	Pyrazine	1261	2	88455	5	ND	ND	16285	3	ND	ND
Al09	2-Decenal, (Z)-	C10H18O	Aldehyde	1271	2	141805	75	81126	165	ND	ND	ND	ND
Un56	Unknown			1276	4	10882	18	ND	ND	2421	5	ND	ND
Al10	Benzeneacetaldehyde, alpha-ethylidene-	C10H10O	Aldehyde	1283	2	21251	8	ND	ND	ND	ND	ND	ND
Ot09	2-Undecanone	C11H22O	Other	1299	2	88793	3	29063	74	31291	8	10058	17

Un57	Unknown			1306	3	ND	ND	ND	ND	15094	16	121378	28
Un58	Unknown			1310	3	46915	3	ND	ND	10750	3	ND	ND
P16	Pyrazine, 2,5-dimethyl-3-(3-methylbutyl)-	C11H18N2	Pyrazine	1321	2	672951	2	14965	6	224014	3	ND	ND
Un59	Unknown			1328	4	18998	65	ND	ND	ND	ND	ND	ND
Un60	Unknown			1334	4	289975	7	45157	20	85256	13	40484	12
Ps16	Trisulfide, dipropyl	C6H14S3	Polysulfide	1347	2	138388	5	22901	30	41846	8	22178	17
Un61	Unknown			1347	3	5455	40	21902	29	57427	4	ND	ND
S01	delta-Elemene	C15H24	Sesquiterpene	1351	2	333910	13	1184409	16	2209440	4	988735	17
Un62	Unknown			1351	4	8223	36	62896	19	126738	4	184419	16
Un63	Unknown			1354	3	163658	11	ND	ND	8046	38	ND	ND
Un64	Unknown			1358	4	ND	ND	ND	ND	40992	19	ND	ND
Un65	Unknown			1359	4	30083	9	2578	46	15299	9	ND	ND
S02	alpha-Cubebene	C15H24	Sesquiterpene	1363	2	7714	9	38435	21	82243	5	10826	23
Un66	Unknown			1367	4	34212	9	2215	51	11989	19	ND	ND
Un67	Unknown			1368	4	6725	5	ND	ND	ND	ND	ND	ND
Un68	Unknown			1371	4	74962	17	ND	ND	ND	ND	ND	ND
Un69	Unknown			1379	3	208807	2	13265	26	76938	2	ND	ND
Un70	Unknown			1383	4	34257	10	ND	ND	9844	37	ND	ND
Un71	Unknown			1391	4	ND	ND	ND	ND	35602	6	31143	23
Un72	Unknown			1391	3	60216	2	ND	ND	23837	2	ND	ND
S03	alpha-Copaene	C15H24	Sesquiterpene	1395	2	124152	17	889126	16	1655974	4	966056	17
Un73	Unknown			1404	3	45942	6	ND	ND	ND	ND	ND	ND
S04	beta-Elemene	C15H24	Sesquiterpene	1405	2	162449	7	337552	19	620122	4	178030	21
Un74	Unknown			1406	4	33775	12	ND	ND	16797	36	ND	ND
S05	beta-cis-Caryophyllene	C15H24	Sesquiterpene	1426	2	221661	10	397684	24	1930094	12	409937	18
Un75	Unknown			1426	4	ND	ND	ND	ND	17396	14	25262	17
Un76	Unknown			1432	4	ND	ND	ND	ND	105648	7	3924	83
S06	beta-Caryophyllene	C15H24	Sesquiterpene	1443	2	4225197	9	13407088	15	19103258	2	16907747	10
Un77	Unknown			1443	3	48723	3	ND	ND	ND	ND	ND	ND
Un78	Unknown			1449	4	ND	ND	ND	ND	412563	5	121363	18
Un79	Unknown			1450	4	18029	50	18114	29	339445	5	119694	15
Un80	Unknown			1450	4	19706	30	4649	17	7336	3	ND	ND
Un81	Unknown			1453	3	28678	4	ND	ND	4107	58	ND	ND
Un82	Unknown			1456	4	ND	ND	ND	ND	ND	ND	8455	19
Un83	Unknown			1463	4	32015	2	ND	ND	ND	ND	ND	ND
Un84	Unknown			1469	3	ND	ND	ND	ND	598062	3	284511	19
Un85	Unknown			1475	3	25759	31	ND	ND	ND	ND	ND	ND
S07	alpha-Caryophyllene	C15H24	Sesquiterpene	1478	2	360662	8	743015	19	1675206	3	947810	15
Un86	Unknown			1487	4	8158	13	ND	ND	24593	31	2548	62
S08	alpha-Curcumene	C15H22	Sesquiterpene	1490	2	414421	5	432195	21	1615053	2	924447	15
Un87	Unknown			1493	4	67241	1	ND	ND	6263	24	ND	ND
Un88	Unknown			1495	3	ND	ND	ND	ND	65150	9	ND	ND
Un89	Unknown			1495	4	23607	3	ND	ND	12642	17	ND	ND
Un90	Unknown			1501	4	12817	51	18437	54	3931	37	ND	ND
S09	alpha-Zingiberene	C15H24	Sesquiterpene	1503	2	69754	6	90244	29	512593	9	160591	36
Un91	Unknown			1503	4	ND	ND	ND	ND	19495	12	34917	47
Un92	Unknown			1507	3	49034	22	ND	ND	77240	31	13085	21
S10	beta-Selinene	C15H24	Sesquiterpene	1511	2	70940	14	281845	21	812179	2	431246	19
S11	alpha-Selinene	C15H24	Sesquiterpene	1516	2	102265	14	377199	21	960440	3	616232	19
Un93	Unknown			1520	3	ND	ND	31413	31	105665	3	39458	30
Un94	Unknown			1528	4	26253	5	ND	ND	ND	ND	8777	50
Un95	Unknown			1531	4	21729	17	95953	27	425560	6	208181	24
Un96	Unknown			1531	4	ND	ND	22403	34	74247	5	184180	18
Un97	Unknown			1533	4	14523	72	41349	27	12478	5	140256	23
Un98	Unknown			1533	3	199694	4	234715	22	518701	3	294675	18
S12	cis-Calamenene	C15H22	Sesquiterpene	1536	2	8796	5	8957	38	237513	3	34417	37
Un99	Unknown			1541	3	ND	ND	ND	ND	19096	3	ND	ND
Un100	Unknown			1548	4	ND	ND	ND	ND	36902	7	9344	34
Un101	Unknown			1558	3	ND	ND	ND	ND	52253	4	ND	ND
Un102	Unknown			1561	4	71495	6	ND	ND	4769	4	ND	ND
Un103	Unknown			1563	3	333579	4	ND	ND	29138	6	ND	ND
Un104	Unknown			1566	3	14165	13	ND	ND	ND	ND	ND	ND
Un105	Unknown			1590	4	1397	70	9694	43	ND	ND	36322	36
Un106	Unknown			1594	4	466083	4	10507	46	118106	32	19537	48
Un107	Unknown			1596	4	5129	22	ND	ND	ND	ND	ND	ND
Un108	Unknown			1599	3	227040	3	ND	ND	ND	ND	ND	ND
Sa01	Caryophyllene oxide	C15H24O	Sesquiterpene alcohol	1606	2	759727	3	ND	ND	ND	ND	ND	ND
Un109	Unknown			1623	3	386877	3	ND	ND	4794	38	ND	ND
Un110	Unknown			1628	4	98956	15	ND	ND	ND	ND	ND	ND
Un111	Unknown			1639	4	213502	7	ND	ND	ND	ND	ND	ND
Sa02	Isospathulenol	C15H24O	Sesquiterpene alcohol	1643	2	951805	2	ND	ND	45282	8	ND	ND
Un112	Unknown			1647	3	29051	4	ND	ND	ND	ND	ND	ND
Un113	Unknown			1648	3	211277	10	ND	ND	ND	ND	ND	ND
Un114	Unknown			1650	4	37637	9	ND	ND	ND	ND	ND	ND
Un115	Unknown			1653	4	55235	7	ND	ND	ND	ND	ND	ND
Un116	Unknown			1658	4	33661	15	ND	ND	4162	11	ND	ND
Un117	Unknown			1676	3	486601	2	ND	ND	28702	9	ND	ND
Un118	Unknown			1686	4	16602	19	ND	ND	ND	ND	ND	ND
Un119	Unknown			1689	4	ND	ND	ND	ND	19706	5	ND	ND
Un120	Unknown			1689	4	16559	9	ND	ND	ND	ND	ND	ND
Un121	Unknown			1693	3	251325	5	ND	ND	ND	ND	ND	ND
Un122	Unknown			1696	4	37024	9	7867	62	ND	ND	ND	ND
Un123	Unknown			1705	3	342943	6	ND	ND	9188	30	ND	ND
Sa03	Xanthorrhizol	C15H22O	Sesquiterpene alcohol	1749	2	26770	12	ND	ND	ND	ND	ND	ND
Un124	Unknown			1769	4	ND	ND	ND	ND	58895	8	ND	ND
Un125	Unknown			1777	4	42623	25	ND	ND	ND	ND	ND	ND
Un126	Unknown			1785	4	96382	6	ND	ND	ND	ND	ND	ND
Un127	Unknown			1794	4	332467	5	ND	ND	13232	11	ND	ND
Un128	Unknown			1804	4	37993	4	ND	ND	ND	ND	ND	ND
Un129	Unknown			1809	4	56940	4	ND	ND	3618	49	ND	ND
Un130	Unknown			1817	4	60990	6	ND	ND	2535	3	ND	ND
Un131	Unknown			1845	4	20819	11	18837	24	ND	ND	19327	23
Un132	Unknown			1859	4	42803	13	ND	ND	ND	ND	ND	ND
Un133	Unknown			1865	4	12569	20	ND	ND	ND	ND	ND	ND
Un134	Unknown			1940	3	348813	61	ND	ND	ND	ND	ND	ND
Un135	Unknown			1955	4	59384	33	18322	9	ND	ND	ND	ND

Figure S1. Abundance of individual polysulfides expressed in Total Ion Current (TIC) * 10^6 for the different trapping techniques in G28 **(A)** and S99 **(B)**. SBSE (blue): Stir-bar sorptive extraction; HSSE (grey): Headspace sorptive extraction; SPME (yellow): Solid-phase microextraction; DHS (green): Dynamic headspace. RI: Calculated retention index.

A	Polysulfides	Molecular Formula	RI
Ps01	Sulfide, allyl methyl	C4H8S	689
Ps02	Disulfide, dimethyl	C2H6S2	735
Ps03	Sulfide, diallyl	C6H10S	851
Ps04	Disulfide, methyl allyl	C4H8S2	917
Ps05	Disulfide, methyl 1-propenyl, I-	C4H8S2	941
Ps06	Trisulfide, dimethyl	C2H6S3	980
Ps07	Disulfide, diallyl	C6H10S2	1095
Ps08	Disulfide, 1-propenyl propyl, I-	C6H12S2	1107
Ps09	Disulfide, allyl 1-propenyl, I-	C6H10S2	1109
Ps10	Disulfide, allyl 1-propenyl, (Z)-	C6H10S2	1115
Ps12	Trisulfide, allyl methyl	C4H8S3	1159
Ps13	Trisulfide, methyl propyl	C4H10S3	1172
Ps14	Trisulfide, diallyl	C6H10S3	1322
Ps15	Trisulfide, allyl propyl	C6H12S3	1334



B	Polysulfides	Molecular Formula	RI
Ps03	Sulfide, diallyl	C6H10S	852
Ps04	Disulfide, methyl allyl	C4H8S2	917
Ps06	Trisulfide, dimethyl	C2H6S3	979
Ps07	Disulfide, diallyl	C6H10S2	1094
Ps10	Disulfide, allyl 1-propenyl, (Z)-	C6H10S2	1115
Ps11	Disulfide, dipropyl	C6H14S2	1123
Ps12	Trisulfide, allyl methyl	C4H8S3	1159
Ps13	Trisulfide, methyl propyl	C4H10S3	1172
Ps16	Trisulfide, dipropyl	C6H14S3	1347

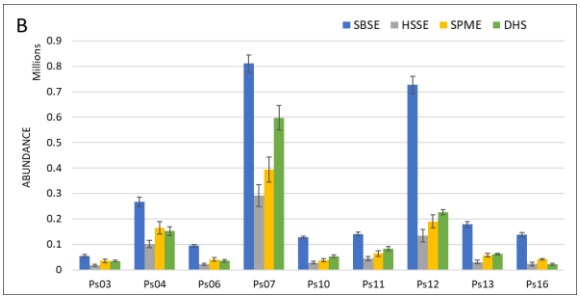
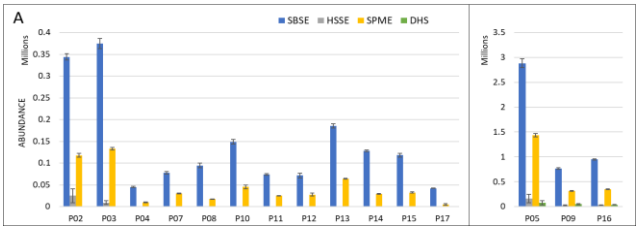
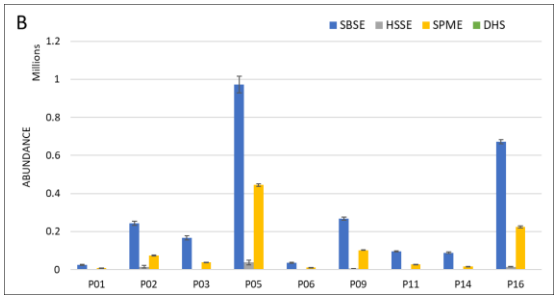


Figure S2. Abundance of individual pyrazines and other aromatics expressed in Total Ion Current (TIC) * 10^6 for the different trapping techniques. **(A)** Pyrazines in G28; **(B)** Pyrazines in S99; **(C)** Aromatics in G28; **(D)** Aromatics in S99.

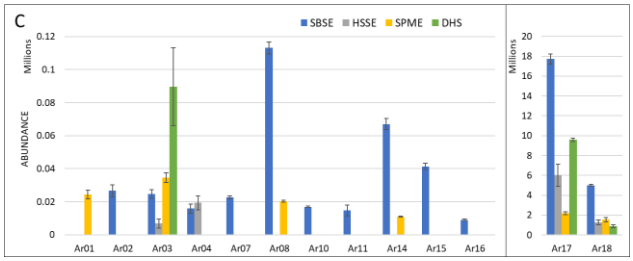
A	Pyrazines	Molecular Formula	RI
P02	Pyrazine, 2,5-dimethyl-	C6H8N2	908
P03	Pyrazine, 2-ethyl-6-methyl-	C7H10N2	1002
P04	Pyrazine, 2-ethenyl-6-methyl-	C7H8N2	1023
P05	Pyrazine, 3-ethyl-2,5-dimethyl-	C8H12N2	1084
P07	Pyrazine, 2-ethyl-3,5-dimethyl-	C8H12N2	1092
P08	Pyrazine, 2-methyl-5-propyl-	C8H12N2	1105
P09	Pyrazine, 3,5-diethyl-2-methyl-	C9H14N2	1164
P10	Pyrazine, 2,5-dimethyl-3-propyl-	C9H14N2	1167
P11	Pyrazine, 2,5-dimethyl-3-(2-methylpropyl)-	C10H16N2	1209
P12	Pyrazine, 2,3-dimethyl-5-(1-methylpropyl)-	C10H16N2	1242
P13	Pyrazine, 2,3,5-trimethyl-6-propyl-	C10H16N2	1250
P14	Pyrazine, 2-isoamyl-6-methyl-	C10H16N2	1261
P15	Pyrazine, 2,5-dimethyl-3-(2-methylbutyl)-	C11H18N2	1310
P16	Pyrazine, 2,5-dimethyl-3-(3-methylbutyl)-	C11H18N2	1321
P17	Pyrazine, 2-(3-methylbutyl)-3,5-dimethyl	C11H18N2	1499



B	Pyrazines	Molecular Formula	RI
P01	Pyrazine, methyl-	C5H6N2	811
P02	Pyrazine, 2,5-dimethyl-	C6H8N2	909
P03	Pyrazine, 2-ethyl-6-methyl-	C7H10N2	1002
P05	Pyrazine, 3-ethyl-2,5-dimethyl-	C8H12N2	1084
P06	Pyrazine, 2,6-diethyl-	C8H12N2	1091
P09	Pyrazine, 3,5-diethyl-2-methyl-	C9H14N2	1164
P11	Pyrazine, 2,5-dimethyl-3-(2-methylpropyl)-	C10H16N2	1209
P14	Pyrazine, 2-isoamyl-6-methyl-	C10H16N2	1261
P16	Pyrazine, 2,5-dimethyl-3-(3-methylbutyl)-	C11H18N2	1321



C	Aromatics	Molecular Formula	RI
Ar01	Thiazole	C3H3NS	698
Ar02	Pyridine	C5H5N	727
Ar03	Thiophene, 3-methyl-	C5H6S	770
Ar04	Furfural	C5H4O	819
Ar07	Pyridine, 3-ethyl-	C7H9N	959
Ar08	3H-1,2-Dithiole	C3H4S2	965
Ar10	Pyridine, 2-ethyl-5-methyl-	C8H11N	1032
Ar11	Thiazole, 4-ethyl-5-methyl-	C6H9NS	1045
Ar14	4H-1,3-Dithiine, 2-vinyl-	C6H8S2	1239
Ar15	Benzene, hexyl-	C12H18	1274
Ar16	Indane-4-carboxaldehyde	C10H10O	1283
Ar17	Pyridine, 3-phenyl-	C11H9N	1481
Ar18	2-Furancarbothioamide, N-butyl-	C9H13NOS	1584



D	Aromatics	Molecular Formula	RI
Ar02	Pyridine	C5H5N	727
Ar04	Furfural	C5H4O2	819
Ar05	Furfuryl alcohol	C5H6O2	839
Ar06	Furan, 2-acetyl-	C6H6O2	903
Ar08	3H-1,2-Dithiole	C3H4S2	965
Ar09	Benzene, tert-butyl-	C10H14	1028
Ar12	Thiophene-2-carboxaldehyde	C5H4OS	1060
Ar13	Pyrrole, 2-acetyl-	C6H7NO	1067

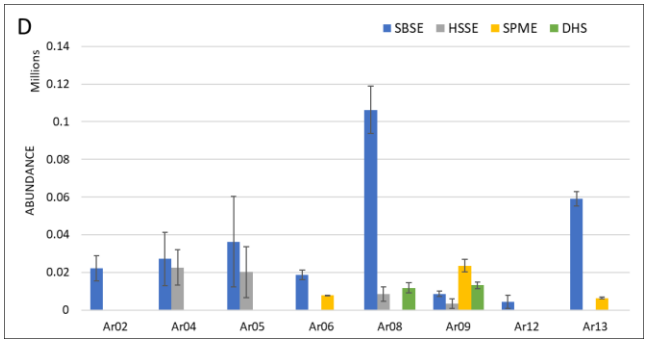
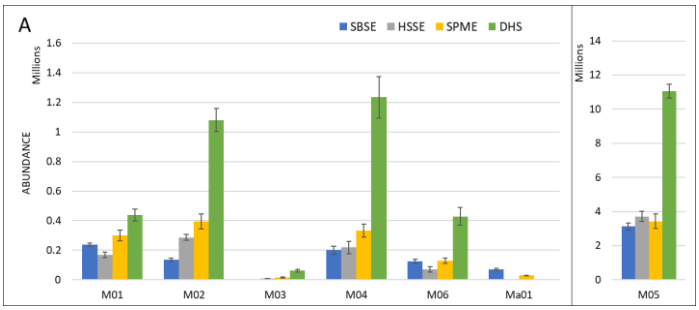
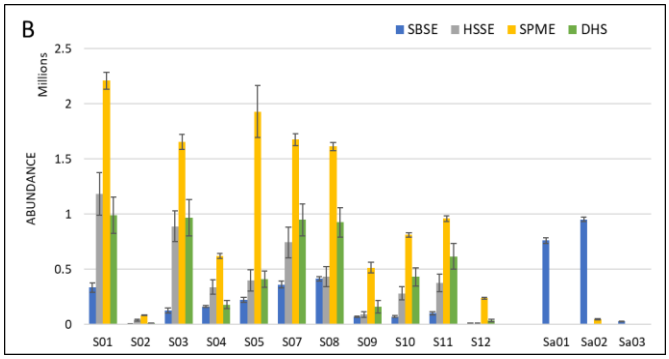


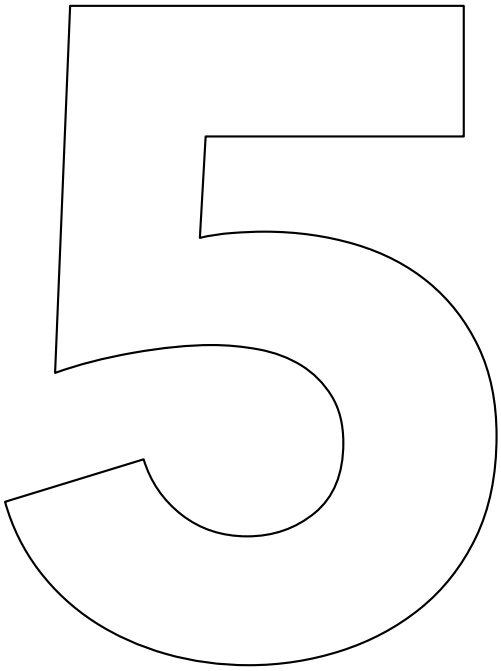
Figure S3. Abundance of monoterpenes **(A)** and sesquiterpenes **(B)** expressed in Total Ion Current (TIC) * 10^6 for the different trapping techniques in sample S99.

A	Monoterpenes	Molecular Formula	RI
M01	p-Cymene	C10H14	1035
M02	alpha-Pinene	C10H16	939
M03	Camphene	C10H16	959
M04	gamma-Terpinene	C10H16	1016
M05	D-Limonene	C10H16	1040
M06	alpha-Terpinolene	C10H16	1099
Ma01	alpha-Terpineol	C10H18O	1211



B	Sesquiterpenes	Molecular Formula	RI
S01	delta-Elemene	C15H24	1351
S02	alpha-Cubebene	C15H24	1363
S03	alpha-Copaene	C15H24	1395
S04	beta-Elemene	C15H24	1405
S05	beta-cis-Caryophyllene	C15H24	1426
S07	alpha-Caryophyllene	C15H24	1478
S08	alpha-Curcumene	C15H22	1490
S09	alpha-Zingiberene	C15H24	1503
S10	beta-Selinene	C15H24	1511
S11	alpha-Selinene	C15H24	1516
S12	cis-Calamenene	C15H22	1536
	Sesquiterpene alcohols		
Sa01	Caryophyllene oxide	C15H24O	1606
Sa02	Isospathulenol	C15H24O	1643
Sa03	Xanthorrhizol	C15H22O	1749





Stir bar sorptive extraction of aroma compounds in soy sauce: Revealing the chemical diversity

Carmen Diez-Simon*, Charlotte Eichelsheim*, Doris M. Jacobs, Roland Mumm, Robert D. Hall

*Equal first authors

Food Research International (2021)

DOI: 10.1016/j.foodres.2021.110348

Abstract

Fermented soy sauce is used worldwide to enhance the flavour of many dishes. Many types of soy sauce are on the market, and their differences are mostly related to the country of origin, the production process applied and the ratio of ingredients used. Consequently, several aromas, tastes, colours, and textures are obtained. Nowadays, soy sauce can also be produced without microorganisms making the process shorter and cheaper. However, flavour may be lost. We have carried out a comprehensive metabolomics analysis of volatile compounds using stir bar sorptive extraction (SBSE)-GC-MS to relate differences in volatile content to production history and origin. The results revealed major differences between fermented and non-fermented soy sauces, and a list of volatile compounds is reported as being characteristic for each type. This study was also able to relate volatiles to the production process using SBSE-GC-MS and to aroma characteristics using GC-O-MS.

Keywords: soy sauce; aroma; stir bar sorptive extraction (SBSE); gas chromatography-mass spectrometry (GC-MS); GC-olfactometry-MS (GC-O-MS); volatiles; fermentation

1. Introduction

Soy sauce is a very common seasoning used worldwide. It is characterized by its strong umami, salty and smoky flavour, which is used to enhance the overall taste and aroma experience of many types of dishes. Soy sauce originated in China over 2200 years ago during the Western Han dynasty and spread to Japan in the 7th century where it is still today known as *shoyu* (Hosking 1996). In Japan, soy sauce export was started in 1647 by the Dutch East India Company (Ohyama 2013). Today, several countries (and industries) are involved in producing soy sauce, with China and The Netherlands being the two largest exporters in the world ("FAOSTAT" 2017). Currently, the annual global consumption is estimated at 10 billion litres (Lee & Khor, 2015), which makes soy sauce a substantial component of the global food seasoning market.

Soy sauce is a black liquid seasoning made essentially from soybeans, wheat (optional) and brine. There are two methods for producing soy sauce; firstly, traditional fermentation (TF) which uses starter cultures of microorganisms (e.g. *Aspergillus oryzae*) and secondly, acid hydrolysis (AH), which is an artificial process using concentrated acid and high temperatures to break down the raw ingredients and generate various flavour compounds. Although AH is less time-consuming (days compared to months – years (Diez-Simon et al. 2020b)), the product lacks many typical fermentation-derived flavour compounds. For this reason, AH soy sauce is sometimes mixed with TF soy sauce to create a product referred to as 'mixed' or 'blended' soy sauce. Recently, both production procedures have been explained and reviewed in depth, in relation both to the metabolite content as well as their sensory characteristics (Diez-Simon et al. 2020b).

Soy sauces can also be classified according to their geographic origin or to the ratios of raw materials used. For instance, the Japanese type is made using equal amounts of soybeans and wheat, whereas Chinese type soy sauce is produced with predominantly soybeans (80%) (Wanakhachornkrai and Lertsiri 2003). Tamari is a speciality Japanese type soy sauce which is prepared without any or only a small amount of wheat (Luh 1995). As volatiles are determinant of aroma characteristics, various studies have investigated the volatile profiles of different types of soy sauce (Feng, Cai, et al. 2014; Lee et al. 2013; Lee et al. 2019; Sun et al. 2010; Zheng et al. 2013). So far, nearly 300 aroma-related volatile compounds have been identified (Devanthi and Gkatzionis 2019). Soy sauce fermentation is a complex process which depends on several variables and chemical reactions to produce a wide range of volatile compounds. Both enzymatic, and non-enzymatic reactions lead to the formation of volatiles such as alcohols, aldehydes, acids, esters, furan(one)s, ketones, phenols, pyrazines and sulphur-containing compounds (Nunomura et al. 1976), many of which have specific aroma qualities and intensities.

To date, only a single study has analysed and compared TF and AH soy sauces in relation to their volatile differences and aroma properties (Lee et al. 2006). The authors performed volatile analysis in TF and AH soy sauces using solvent extraction and solid-phase microextraction (SPME) coupled to gas chromatography-mass spectrometry (GC-MS). The TF soy sauces analysed were characterized mainly by alcohols and esters which are formed from microbial activity. In contrast, pyrazines, furans and acids were more abundant in the AH soy sauces, and these are known to be formed during heating processes.

The most common and robust technique to analyse (and identify) volatile aroma compounds is gas chromatography-mass spectrometry (GC-MS). Sorptive-based techniques in combination with GC-MS are effectively used to extract (semi-)volatiles from the complex (food-)matrices such as soy sauce. These have proven to provide a suitable alternative for solvent extraction techniques since for example, they are fast, less labour intensive and more cost-effective (Diez-Simon et al. 2020a; Nogueira 2015). In soy sauce, solid-phase microextraction (SPME)-GC-MS is now the most widely used approach (Chen et al. 2015). However, in recent years, other techniques, such as stir bar sorptive extraction (SBSE), have gained some popularity and have been optimized to cover a broader range of (semi-)volatiles in liquid food materials (David et al. 2019). Interestingly, only one study describes the application of SBSE in soy sauce (Lee et al. 2019). The authors investigated the changes in volatile compounds in a Korean-type soy sauce during the fermentation period. They analysed the volatiles using both SPME and SBSE, and concluded that during long-term fermentation, the levels of most esters, some phenols, benzene and benzene derivatives, lactones, and pyrroles increased, while some alcohols, except for ethanol, and ketones decreased amid fermentation time. However, little is described about the technical comparison of both extraction techniques

The main goal of this study was to characterize comprehensively the volatile profiles of TF and AH soy sauces using an untargeted metabolomics approach. The aim was to relate compositional differences in the volatile profiles to the origin and production history of the samples and hence allow us to make hypotheses on the possible mechanisms behind the formation of volatile compounds and aroma. In order to link aroma to individual volatiles and hence gain insight into their specific contributions to sensory impact, we also initiated the application of GC-Olfactometry-MS to enable the determination of the individual aroma characteristics of key soy sauce volatiles. In the present study, both TF and AH soy sauces were included in order to make a detailed, direct chemical and sensory comparison. A broad range of commercially available products was selected which had originated from different countries and which had been made using different ratios of raw ingredients, processing temperatures and storage histories (aging). The sample metadata is summarized in **Table 1**.

2. Materials and methods

2.1 Soy sauce materials and chemicals

Twenty contrasting commercial soy sauces were purchased from local stores in the Netherlands, or via online outlets. The different brands and types were selected based on the processing method (TF or AH), country of origin, salt content, wheat content and other ingredients or additives (**Table 1**). Physico-chemical properties of the soy sauces (such as density and pH) were measured and values were constant around 1.14 g/mL and pH 4.6. Samples were kept closed in their own containers at 4 °C, and aliquoted shortly before analysis.

AH soy sauces are often not clearly distinguished as “acid-hydrolysed”. Instead, the term soy protein or soy extract is often stated in the ingredients list, or the use of “brewed” and/or “fermented” in the label is avoided. Many times, they also come in cheaper plastic bottles, and are manufactured in countries such as The Philippines or Thailand where the use of AH is still allowed (China has banned AH soy sauces).

Therefore, we hypothesized that soy sauces on the market without any clear labelling of a fermentation process and, instead, stated soy protein were AH soy sauces. In the case of mixed soy sauces (SJS and SJG), this was clearly stated on the product label. In the case of unclassified samples (CYL and MYL, **Table 1B**), no fermentation process was stated nor the use of soy protein.

Sodium chloride (NaCl) was purchased from Honeywell-Fluka (Seelze, Germany). An n-Alkane (C_8-C_{22}) series for calculating retention indices (RI) was prepared (Sigma Aldrich, St. Louis, MO, USA).

Table 1A. Characteristics of the ten contrasting soy sauces selected for the first experiment.

Soy sauce	Type	Abbrev.	Country of origin	Salt content (g/100 mL)	Soybean (content)	Wheat (content)	Other ingredients	Additives	Production details (if known)
Kikkoman natural	TF	**KNA**	Japan	16.9	Soybeans (n.d,)	Wheat (n.d,)	-	-	Traditional fermentation
Kikkoman low salt	TF	**KLS**	Japan	9.1	Soybeans (n.d,)	Wheat (n.d,)	Spirit vinegar, alcohol, sugar	-	Traditional fermentation, 43% salt is removed in a special process
Tamari gluten-free (Kikkoman)	TF	**TAM**	Japan	16.4	Soybeans (n.d,)	-	Spirit vinegar (Brandy)	-	Traditional fermentation using Aspergillus tamarii
Lee Kum Kee Premium Light	TF	**LPL**	China	17.2	Soybeans (11%)	Wheat flour (n.d,)	Sugar	E631, E627	Traditional fermentation
Lee Kum Kee Double Deluxe	TF	**LDD**	China	16.7	Soybeans (11%)	Wheat flour (n.d,)	Sugar	-	Traditional fermentation, double fermenting process
Haday superior light	TF	**HSL**	China	17.3	Soybeans (21.5%)	Wheat (n.d,)	Yeast extract	E631, E627	Traditional fermentation
Healthy Boy Shoyu	TF	**HBS**	Thailand (Japanese style)	10.2	Soybeans (n.d,)	Wheat flour (n.d,)	Sugar, high fructose syrup, mirin (glucose syrup, alcohol, glutinous rice, alanine)	E631, E627, E640, E211, E150c	Traditional fermentation (> 6 months)
Sempio Jin Gold F3	Mixed	**SJG**	Korea	13.5	Defatted soybean (n.d.)	Wheat (n.d,)	High fructose corn syrup, spirits, liquorice extract, sucralose, yeast extract	E211	Combination of traditional fermented and acid hydrolysed soy sauces
Sempio Jin S	Mixed	**SJS**	Korea	13.4	Defatted soybeans (n.d,)	Wheat (n.d,)	High fructose corn syrup, spirits, sucralose, yeast extract	E211	Combination of traditional fermented and acid hydrolysed soy sauces
Data Puti	AH	**DAP**	Philippines	5.2	Soybean extract (35%)	n.d.	-	E150c, E211	Acid hydrolysed soy sauce using soy sauce protein

n.d.: Not defined on the product label or other reliable sources
ND: Not defined, unclassified soy sauce

Table 1B. Characteristics of the fifteen contrasting soy sauces selected for the second experiment.

Soy sauce	Type	Repeated in exp. 1	Abbrev.	Country of origin	Salt content (g/100 mL)	Soybean (content)	Wheat (content)	Other ingredients	Additives	Production details (if known)
Kikkoman natural	TF	**Yes**	**KNA**	Japan	16.9	Soybeans (n.d,)	Wheat (n.d,)	-	-	Traditional fermentation
Tamari gluten-free (Kikkoman)	TF	**Yes**	**TAM**	Japan	16.4	Soybeans (n.d,)	-	Spirit vinegar (Brandy)	-	Traditional fermentation using Aspergillus tamarii
Lee Kum Kee Double Deluxe	TF	**Yes**	**LDD**	China	16.7	Soybeans (11%)	Wheat flour (n.d,)	Sugar	-	Traditional fermentation, double fermenting process
Pearl River Bridge	TF	**No**	**PDS**	China	21.2	Soybeans (20%)	Wheat flour (n.d,)	Sugar	-	Traditional fermentation, non-GMO, no artificial flavouring added
Pearl River Bridge	TF	**No**	**PLS**	China	17.5	Soybeans (29%)	Wheat flour (n.d,)		E202	Traditional fermentation, non-GMO, no artificial flavouring added
Yamasa soy sauce	TF	**No**	**YAM**	Japan	16.5	Soybeans (n.d,)	Wheat (n.d,)	Alcohol	-	Brewed
Tai Hua Soy Sauce	TF	**No**	**THD**	Singapore	17.4	Soybeans (n.d,)	Wheat flour (n.d,)	Sugar, caramel	E211	Naturally brewed
Tamari megachef	TF	**No**	**TOB**	Thailand	15.1	Soybeans (21%)	-	Rice, glucose fructose syrup	E202	Traditionally brewed. Gluten-free. No artificial colour added. No MSG added
Wan Ja Shan (less salt)	TF	**No**	**WJS**	Taiwan	0.98	Soybeans (12%)	Wheat (n.d,)	Sugar, alcohol, lactic acid, yeast extract	-	100% naturally brewed. No preservatives
Inproba Bio-organic Soja sauce	TF	**No**	**IBO**	The Netherlands	19	Soybeans (24%)	Wheat (n.d,)	Alcohol	-	100% natural ingredients. No artificial additives
Chan's soy sauce Yellow Label	ND	**No**	**CYL**	The Netherlands	18.9	Soybeans (n.d.)	Wheat flour (n.d,)	Sugar	E150a	n.d.
Maekrua soy sauce Yellow Label	ND	**No**	**MYL**	Thailand	18.4	Soybeans (63%)	Wheat flour (20%)	Sugar (3.9%)	E211	n.d.
Sempio Jin S	Mixed	**Yes**	**SJS**	Korea	13.4	Defatted soybeans (n.d,)	Wheat (n.d,)	High fructose corn syrup, spirits, sucralose, yeast extract	E211	Combination of traditional fermented and acid hydrolysed soy sauces
Silver Swan soy sauce	AH	**No**	**SSW**	Philippines	17	Soybean protein (20%)	Wheat flour (n.d,)	-	E202, E150c	n.d.
Data Puti	AH	**Yes**	**DAP**	Philippines	5.2	Soybean extract (35%)	n.d.	-	E150c, E211	Acid hydrolysed soy sauce using soy sauce protein

n.d.: Not defined on the product label or other reliable sources
ND: Not defined, unclassified soy sauce

2.2 Sorptive-based techniques

For the determination of the volatile and semi-volatile profiles, two different trapping approaches (SPME and SBSE) were tested (Supplementary **Figure S1**). After extracting the analytes, these were then thermally desorbed and introduced into the GC-MS. The same GC-MS instrument and settings were employed for both trapping techniques, and both were performed using a multi-purpose autosampler (MP-2, Gerstel, Mülheim an der Ruhr, Germany), operated using Gerstel MAESTRO software (version 3.2). Below, the trapping and GC-MS conditions are summarized based on a previous method (Diez-Simon et al., 2020a), with slight optimizations for the soy sauce matrix.

2.2.1 Solid-phase micro extraction (SPME)

A 5 mL aliquot of soy sauce was placed in a 10 mL crimp cap vial and 1.75g of NaCl (to generate saturated conditions) was added. The vials were incubated at 60 °C for 10 min with agitation (250 rpm) to drive volatiles out of the liquid. Subsequently, volatiles were trapped by exposing the SPME fiber to the headspace above the liquid for 20 min at 60 °C without agitation (Diez-Simon et al., 2020a). The fiber was then thermally desorbed onto the GC column via the cooled injection system (CIS, Gerstel, Mülheim an der Ruhr, Germany) containing a glass liner with a helium flow of 1 mL/min at 280 °C for 2 min, in splitless mode.

2.2.2 Stir bar sorptive extraction (SBSE)

A volume of 9 mL of soy sauce in a 10 mL screw-cap glass vial was used to trap volatiles. Subsequently, a stir bar (Twister® coated with 24 µL PDMS) was immersed in the sample and with continuous stirring at 450 rpm for 60 min. After extraction, the bars were removed from the samples, rinsed for 2-3 seconds with distilled water, dried with a lint-free tissue and placed inside clean glass liners for thermal desorption. Analyses were started immediately after sampling to prevent any loss of volatiles. Stir bars were desorbed in a thermal desorption unit (TDU, Gerstel, Mülheim an der Ruhr, Germany) in splitless mode using a helium flow of 1 mL/min. The initial temperature was 30 °C (0.5 min hold) and was then heated at a rate of 120 °C/min to a final temperature of 175 °C (5 min hold). Volatile compounds were transferred to the CIS containing a packed sorbent liner (Tenax TA), which was heated from -10 °C at a rate of 720 °C/min to a final temperature of 250 °C (5 min hold) to transfer the volatiles to the analytical column. During this, a split of 1:5 was used.

2.3 Gas Chromatography-Mass Spectrometry (GC-MS) conditions

Analyses were conducted on an Agilent GC7890A coupled to a 5975C quadrupole mass spectrometer. The column used was a Zebron ZB-5Msplus with dimensions 30 m x 0.25 mm x 1.00 µm (Phenomenex). The column oven was programmed starting at 45 °C for 2 min, then increased at a rate of 5 °C/min to 280 °C and then maintained at 280 °C for 5 min. The column effluent was ionised by electron impact at 70 eV, in the scan range m/z 33–500. The interface temperature was 280 °C. Compound identification was based on the principles explained in Section 2.5.

2.4 GC-Olfactometry-MS analyses (GC-O-MS)

In order to relate the specific aroma characteristics with the individual compounds, a double detector (olfactory detection port, ODP, and MS) was used, by splitting the GC column outlet. Volatile extracts for GC-O-MS analysis were obtained using SPME, employing the sampling procedure described in Section 2.2.1. Analysis was performed on the same GC-MS instrument, but now connected with an ODP2 (Gerstel). A four-port splitter stand was located in the GC oven; two ports were connected to the column outlet and an auxiliary gas outlet; the two remaining ports were connected to the MS and the ODP. The column outlet pressure was 20 kPa. The capillary column (0.1 mm) of the transfer line was kept under constant temperature of 240 °C. An auxiliary gas (helium) flow of 5 mL/min was maintained constantly during analyses.

The analytical conditions were similar to those for GC-MS analysis, with a few modifications. The oven temperature was programmed as follows: 45 °C (2 min), 5 °C/min to 150 °C and then 15 °C/min to 250 °C (5 min). The sniffing procedure for each soy sauce was carried out by at least three individual panellists. A compound was deemed to be aromatically active if it was perceived by two or more panellists at the same retention time and when it was described using similar odour qualities. Occasionally, aroma was also considered when only perceived by one panellist, but only if a similar attribute was identified at the same retention time in more than one soy sauce. The sensory attributes perceived by panellists at the ODP were linked to specific compounds by matching the retention times and by validating these compounds with aroma-active compounds previously reported in soy sauce. In order to relate these observations to (semi)quantitative differences between the soy sauces, the relative abundances (expressed in Total Ion Current, TIC) of the corresponding compounds after combined data processing were assessed.

2.5 Experimental setup

In the first experiment, a series of 10 soy sauces were analysed in a randomized way (**Table 1A**). In the second analysis, the sample set included 10 other soy sauces, along with five samples from experiment 1, but which had been re-purchased so that they could be considered as true biological samples (**Table 1B**).

For both experiments an empty glass vial and a blank (6M NaCl in water) sample were measured at the beginning of each sequence. Quality control samples (QCs), which were a mix of all soy sauces, were repeatedly analysed along the sequence to test the performance stability. An n-Alkane (C_8-C_{22}) series was analysed at the end of the sequence.

The raw GC-MS data were processed using a nontargeted metabolomics approach. Baseline correction and alignment of all mass signals (with a signal to noise ratio s/n ≥ 3) were performed using MetAlign software (Lommen 2009). Signal redundancy was reduced to single representative variables and mass spectra were reconstructed using MSClust (Tikunov et al. 2012). Metabolites were putatively identified by matching the obtained mass spectra and retention indices (RI) to those in the NIST17 Mass Spectral library (v.2.3), following the criteria for metabolite identification proposed by Sumner et al. (2007). Compounds that did

not fit the criteria, were annotated as being non-identified. Retention indices were calculated based on a series of n-alkanes (C_8-C_{22}) using a third order polynomial function.

For statistical analysis, we compared and visualized the main tendencies of the generated data by principal components analysis (PCA) after log 10 transformation and unit variance (UV) scaling of the samples using SIMCA 15.0.2. software (Sartorius Stedim Data Analytics AB, Umeå, Sweden). Hierarchical cluster analysis (HCA) and partial least squares discriminant analysis (PLS-DA) were carried out using SIMCA 15.0.2. software. Graphs were produced using Microsoft Excel 365 and Rstudio.

3. Results and discussion

3.1 Comparison between SPME- and SBSE-GC-MS volatile profiles of soy sauce

Soy sauce aroma is characterized by a wide range of volatile compounds which enhance different sensory attributes of the product. Considering the large number of different volatiles present, it is important to develop a technique that allows us to analyse the broadest possible spectrum of volatiles, when we want to relate the chemical profile to aroma attributes and sample origin. We tested two volatile extraction techniques (SPME and SBSE, Supplementary **Figure S1**) to determine which one was able to deliver the broader range of compounds, in a repeatable manner. A nontargeted, MS-based metabolomics workflow revealed a total of 246 volatiles using SPME as compared to 542 for SBSE (Supplementary **Table S1**), demonstrating that the overall analyte coverage was considerably higher for SBSE. Similar results had been obtained in a previous investigation of other food flavouring additives (Diez-Simon et al., 2020a). We putatively identified 87 and 114 compounds for SPME and SBSE, respectively (Supplementary **Table S1**). The group of compounds trapped by both techniques was the largest (78). On the other hand, an additional 9 and 36 compounds were uniquely detected by SPME and SBSE, respectively. When comparing the volatile GC-MS profiles of the same soy sauce extracted by SPME and SBSE (e.g. Lee Kum Kee Premium Light, LPL **Figure 1**), clear differences with respect to selectivity become apparent. A considerable number of semi-volatile compounds appearing at higher retention times was particularly trapped by SBSE, whereas SPME trapped more volatile compounds eluting earlier. For example, SPME exclusively trapped some highly volatile compounds such as 2-butenal and 2-pentanone, with vapour pressures of 30.0 and 35.4 mmHg at 25 °C respectively ("CompTox Chemicals Dashboard"). SBSE trapped more, less-volatile compounds, including phenylethyl acetate and ethyl cinnamate, which have vapour pressures of 0.03 and 0.003 mmHg at 25 °C respectively ("CompTox Chemicals Dashboard"). Interestingly, other volatiles relevant for flavour, such as quinoxalines (see Section 3.3), were only trapped by SBSE.

We evaluated the repeatability of both SPME and SBSE by analysing 10 technical replicates of the same soy sauce (Kikkoman natural, KNA). Both techniques showed a good repeatability with an average relative standard deviation (RSD) of compound intensities of 11% and 13% for SPME and SBSE, respectively.

These findings revealed that SBSE performed well for the analysis of volatile compounds, while extracting an additional number of compounds as compared to SPME. By this comparison, we are confident that SBSE is a robust method that offers more coverage of volatile compounds compared to SPME. Moreover, the

targeted list of volatiles obtained by both SPME and SBSE had significant similarities to what was previously observed for a single Korean-type soy sauce (Lee et al., 2019).

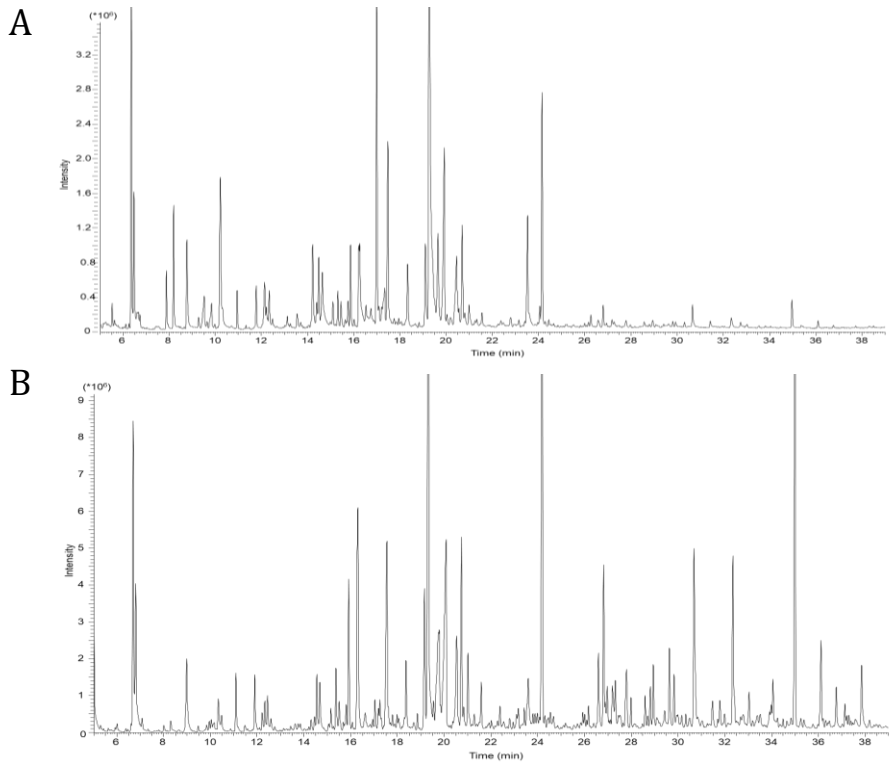


Figure 1. Examples of representative GC-MS Total Ion Current (TIC) chromatograms of Lee Kum Kee Premium Light (LPL) soy sauce volatiles analysed by SPME **(A)** and SBSE **(B)**. SBSE reveals more compounds later in the chromatogram, whereas SPME traps more compounds eluting at an earlier stage.

3.2 Relation between volatile profile and production procedure

In the first trial, a set of ten contrasting soy sauces of diverse origin which are routinely available via local and online outlets was used (**Table 1A**) and subsequently, a second experiment was then performed to validate the findings of the first but using an expanded set of samples including some new samples as well as a number of the same types but from different production batches (**Table 1B**).

Multivariate PCA based on 542 volatile compounds detected in samples of the first trial showed the distribution of the soy sauces according to their volatile profiles (**Figure 2A**). The technical variation in the volatile profiles as derived from the quality control samples and the replicate measurements was small compared to the variation between the different soy sauces (**Figure 2A**). The biggest differences were related to the production process, where PC1 (R^2X: 27.6%) separates the TF soy sauces on the right of the score plot from the AH (DAP) and mixed soy sauces (SJG and SJS) on the left. The first two PCs explained ~40% of the total variance, while no clear explanation for the distribution along PC 2 was found. Where PCA only depicts the grouping of samples – here in **Figure 2A** and **B** based on 40% of the total variance, hierarchical cluster analysis (HCA) depicts associations based on total variance. HCA identified the largest difference between the samples to be related to the production method: chemical hydrolysis (AH, cluster I)

versus biological fermentation (TF, cluster II) (**Figure 2C**). However, in the HCA we can also differentiate subgroups within the TF soy sauces. Subgroup Iia contained the gluten-free replicates and the two other subgroups were composed of Japanese-type (Iib) and Chinese-type soy sauce samples (Iic). On the basis of this first analysis, it was concluded that the production method and the ingredients used result in a clear separation of soy sauce samples based on their volatile profiles.

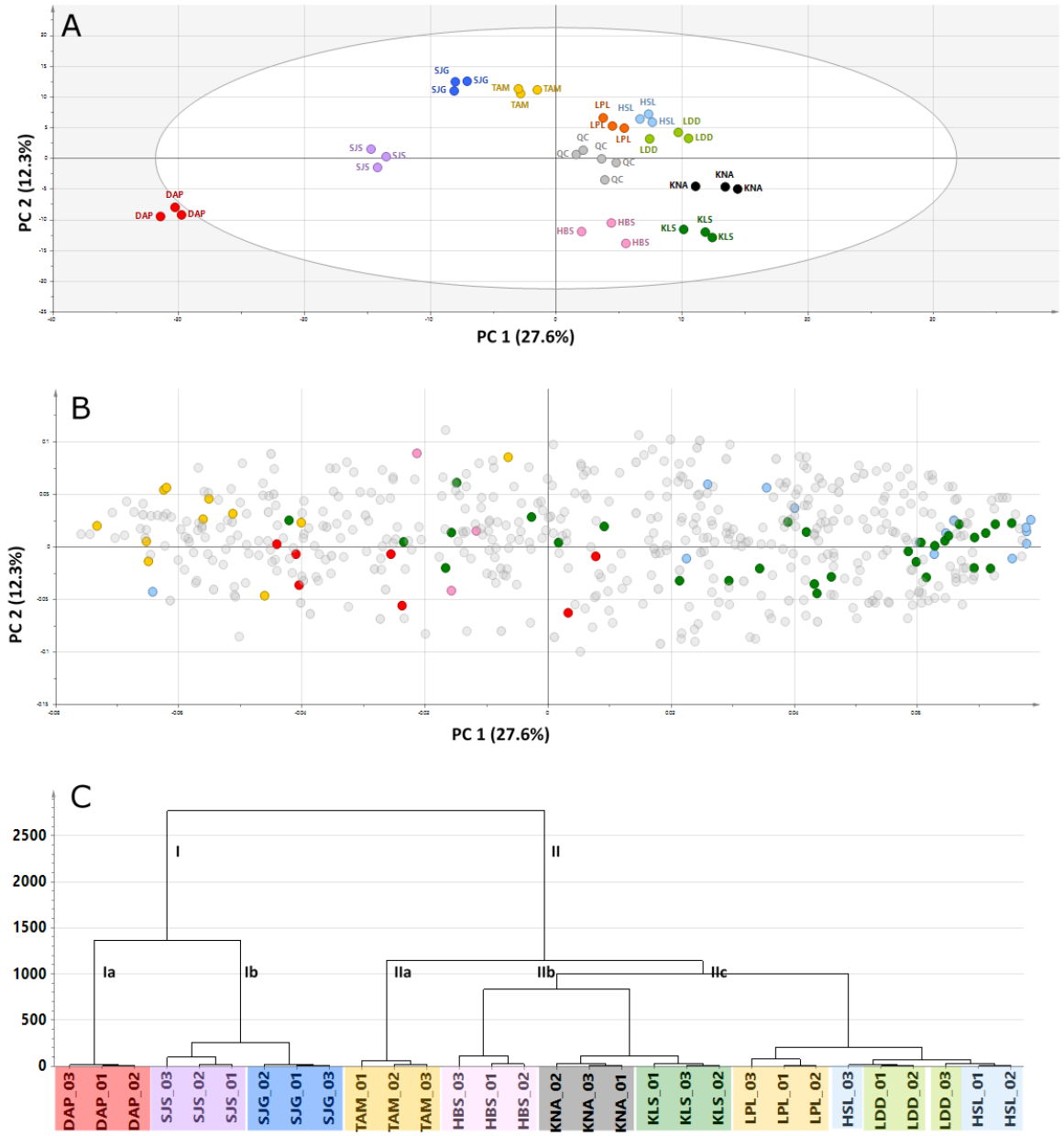


Figure 2. Principal components analysis (PCA) score plot **(A)** and loading plot **(B)** of the 542 volatile compounds of 10 contrasting soy sauces using SBSE-GC-MS in Experiment 1. Three technical replicates are presented for each sample and five quality control samples (QC) are presented in grey. The first and second PC explain the corresponding percentage of variation shown on each axis. Putatively identified volatiles belonging to relevant chemical groups are coloured: acids (red), alcohols (light blue), esters (dark green), pyr(an)ones (pink) and pyrazines (yellow). Unidentified metabolites are coloured grey. **(C)** Hierarchical cluster analysis (HCA) dendrogram of the volatile profiles of the contrasting soy sauces using SBSE-GC-MS. Groups I and II contain (mixed) acid hydrolysed and fermented soy sauces respectively. Groups Ia and Ib contain exclusively and mixed acid hydrolysed soy sauces respectively. Groups IIa, IIb

and IIc contain Tamari (TAM), Japanese type and Chinese type soy sauces respectively. Data Puti (DAP, red); Haday superior light (HSL, light blue); Healthy Boy Shoyu (HBS, pink); Kikkoman less salt (KLS, dark green); Kikkoman natural (KNA, black); Lee Kum Kee Double Deluxe (LDD, lime); Lee Kum Kee Premium Light (LPL, orange); Sempio Jin S (SJS, lavender); Sempio Jin Gold (SJG, dark blue) and Tamari Gluten-Free (TAM, yellow). For a full list of sample codes please see **Table 1**.

Looking deeper into the chemical differences, which are potentially causal to the distribution patterns observed in the PCA, revealed certain trends. As seen in **Figure 2B**, where the identified volatiles are coloured based on chemical class, several identified volatiles, such as pyrazines (yellow), are predominantly co-localizing with the AH sample. On the other hand, esters (dark green) and alcohols (light blue) co-localise on the right side of the PCA with the TF samples. In a previous study using solvent extraction and SPME GC-MS, a high presence of esters and alcohols was detected in TF soy sauces which was attributed to microbial fermentation (Lee et al., 2006). Pyrazines were more dominant in AH soy sauces and were linked to the long heating procedures used for acid hydrolysis, triggering the increased formation of Maillard reaction products. As our analyses are based on an untargeted metabolomics approach, we can now also detect many other still unidentified compounds which also appear to be specific for the different types of soy sauces. A number of these (in grey) are at the extremes of the axes of the PCA loading plot (**Figure 2B**). Further focus to identify these compounds should follow as these can also be potentially relevant for defining the specific chemical nature of different soy sauces and be of relevance to aroma differences.

3.3 Validation of the contrasts between samples

To corroborate the findings and differences described in the first analysis, a second experiment was designed. Here, additional soy sauce brands of similar production procedures, complementary to those previously used, were analysed alongside a number of brands repeated from the first experiment (**Table 1B**). Soy sauce brands that had unclassified product labels were also selected in order to investigate their volatile distribution and gain more insight into their likely production background. PCA was again performed on the GC-MS data generated (Supplementary **Figure S2**). The results showed good complementarity with those previously obtained, since, once again, the samples with comparable production procedures showed the same distribution pattern.

To compare directly the results of the two experiments in one dataset, a heatmap of the merged data is shown in **Figure 3**. This heatmap shows a two-dimensional hierarchical clustering, combined with a spatial heatmap representing the variation of metabolite intensities (log transformed data). The samples are divided into three main clusters. Group I comprised the AH soy sauces DAP and SSW (Ia), and the mixed Sempio soy sauces (Ib, SJS and SJG). Interestingly, unclassified soy sauces (CYL and MYL, in blue) grouped together with a number of other soy sauces labelled as being “naturally brewed” or “traditionally fermented” (PDS, PLS, THD and TOB), thus forming group II. That these soy sauces clustered separately from the other TF (group III) suggests differences in their production procedure and perhaps e.g. the use of higher temperatures. High temperature treatments could trigger the degradation of some of the metabolites common in yeast fermentation and, simultaneously, may trigger the formation of volatiles that have a similar chemistry of those in AH/mixed soy sauces. Moreover, PDS, PLS and TOB had a common description

in their label relating to the non-use of artificial flavourings. It may therefore be that these soy sauces may have had a common (natural) supplement. Lastly, the third cluster (group III) comprised TF soy sauces used. Within this cluster, subgroups are formed, which are again related to the country of origin / ratios of ingredients used. Subgroup IIIa consists of the gluten-free alternative (TAM), as was also seen in the first experiment. The absence of wheat in tamari soy sauce is known to result in a low abundance of alcohols, since these are formed from the carbohydrate degradation during yeast fermentation (Harada et al., 2018; Lee et al., 2013). However, as seen in Supplementary **Figure S3A**, alcohols such as 2-methyl-1-propanol are still present in low amounts in tamari, perhaps due to the carbohydrates present in the soybeans (~30%), and the addition of spirit vinegar lowers the pH, which is favourable for yeast activity. From the heatmap, specific volatiles that are highly correlated with Tamari soy sauce are highlighted in box number 1 (**Figure 3**). Most of these compounds remain unidentified (See Supplementary **Figure S3B** for an example) but can nevertheless be of great interest as potential markers for tamari soy sauce, although the other tamari-type (TOB) lacked several of these volatiles. Subgroup IIIb consisted of the low-salt soy sauces KLS, WJS and HBS. Both KLS and WJS are the low-salt alternatives from the brand Kikkoman and Wan Ja Shan, and show high degrees of similarity in their volatile profiles. Interestingly, HBS soy sauce was not labelled as being "low-salt soy sauce", even though it has the lowest salt content of all TF soy sauces (**Table 1**). Moreover, all three had alcohol and sugar added as additional ingredients (**Table 1**). This demonstrates that salt content in soy sauce changes the volatile profile. The profiles of these brands were characterised by a number of compounds as highlighted in Box 2 (**Figure 3**). The most dominant class of compounds from this selection was the acetals. Acetals have been previously detected in TF soy sauces at low amounts (Feng et al. 2017), but here, in the low-salt soy sauces some of the acetals are increased (Supplementary **Figures S3C**, **D** and **E**). The soy sauces KLS, WJS and HBS contained added ethanol, resulting in a higher ethanol percentage as compared to the regular TF ones. For instance, Kikkoman soy sauce (KNA) has 2.5% ethanol whereas the low-salt version of the same brand (KLS) has 3.5% (Shurtleff and Aoyagi 2012). Interestingly, comparable studies on Chinese soy sauce aroma type liquors containing 40-55% ethanol, revealed acetal presence at high concentrations (Fan et al. 2011; Fan and Qian 2006). This suggests that ethanol content may be a major factor influencing acetal formation in soy sauce.

The next subgroup (IIIc) is comprised of LDD, LPL and HSL soy sauces. These three were produced in China, where they use a ratio of soybeans to wheat of 80:20. Lastly, subgroup IIId, contained the soy sauces KNA, YAM and IBO which are characterized as being Japanese-type. IBO, however, was manufactured in The Netherlands but was expected to be a Japanese-type soy sauce, using an equal ratio of soybeans and wheat. As seen in **Figure 3**, Chinese- and Japanese-types contained various qualitative and quantitative differences in their volatile profiles. However this data is not further discussed here as both types have been studied and extensively compared before for their aroma profiles (Feng, Su, et al. 2015). These authors compared twenty-seven commercial soy sauces from both origins and found similar volatiles occurring in both although intensities differed. In addition, Sun et al. (2010) studied the volatile composition of twelve Chinese soy sauces and observed major differences between them. This was attributed to the raw materials collected from local areas and the different microorganisms employed by manufacturers.

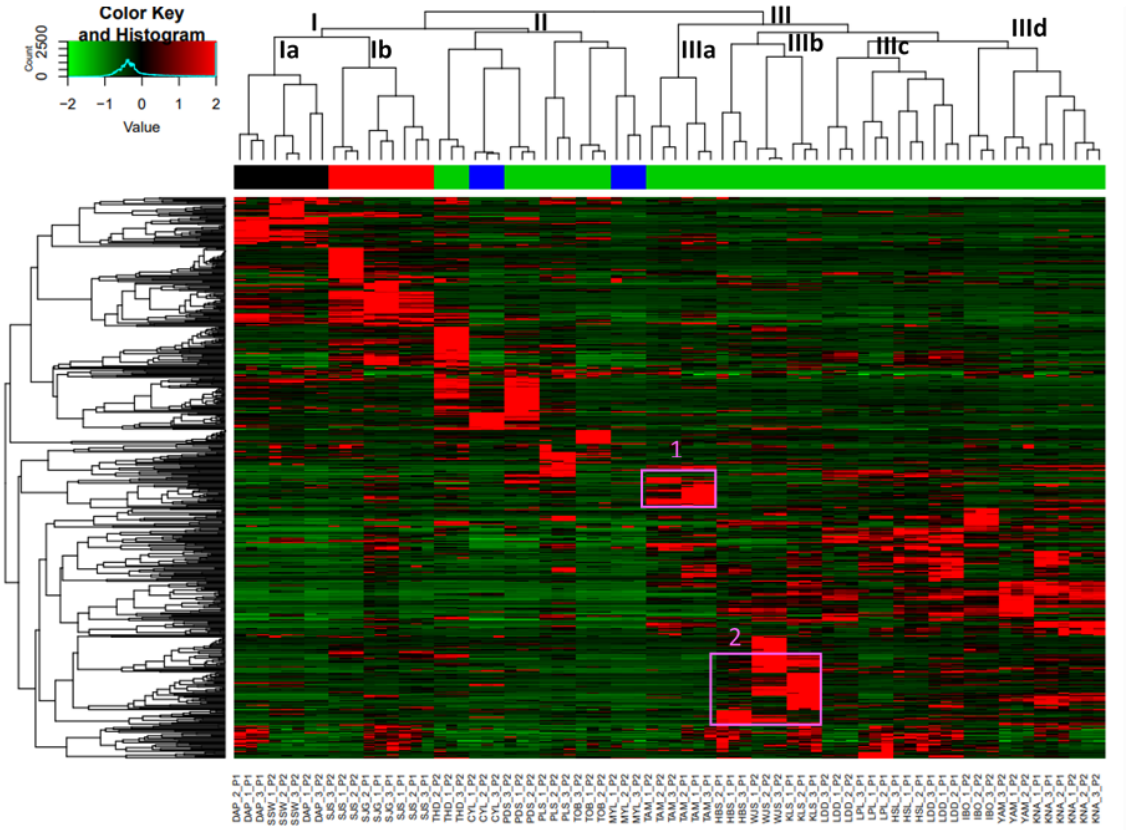


Figure 3. Heatmap based on HCA of twenty soy sauces analysed by SBSE-GC-MS (x-axis) showing the distribution of the different 594 volatile and semi-volatile compounds on the y-axis. Upper bar colour codes: AH soy sauces (black); mixed soy sauces (red); unclear labelled soy sauces (blue); and TF soy sauces (green). Compounds that are strongly correlated/associated with groups IIIa and IIIb are highlighted in Boxes 1 and 2 respectively. For a full list of sample codes please see **Table 1**.

In this study, the biggest differences in volatile profiles were found between TF and AH soy sauces. In order to find potential markers for the two main groups, a PLS-DA analysis was performed by including AH soy sauces in one class, and TF in the other class (thus – excluding soy sauces from group II). The resulting metabolites, including both identified and non-identified, with the highest regression coefficients were selected as being the best potential predictors to differentiate between the two classes. Volatiles such as esters, and ketones were highly (or only) present in TF soy sauces (**Table 2**)- for example, butanedioic acid diethyl ester (Supplementary **Figure S4A**), was found predominantly in TF soy sauces. Again, it was previously reported albeit from a single study that the presence of esters is related to fermentation (Lee et al., 2006). However, we also detected compounds that have not been described before in TF soy sauce. For instance, the two ketones 5(or 4)-methyl-2-hexanone were considerably more abundant in TF soy sauces (Supplementary **Figure S4B**). These ketones were also present in mixed soy sauces (SJS and SJG). Their presence in one of the soy sauces in group II (THD) suggests that this sample was also at least partly of TF origin. Moreover, many non-identified compounds had also a strong correlation with TF soy sauces (**Table 2**, and Supplementary **Figure S4C** and **D** as examples). However, further identification should follow to understand how these compounds appear in TF soy sauces.

Table 2. List of compounds selected, from a PLS-DA analysis, as being characteristic of [A]: TF soy and [B]: AH soy sauces, including both identified and non-identified compounds.

A: Selected markers for traditional fermented (TF) soy sauces									
Nr	**R.T.[a] min**	**RI[b] exp**	**Name**	**Molecular Formula**	**Chem. Group**	**CAS**	**Samp. Name[c]**	**Max. TIC value**	**Regression coefficient**
1	9.03	813	acetic acid, butyl ester	C6H12O2	Ester	123-86-4	HBS	3028721	-0.00393
2	10.16	846	4-methyl-2-hexanone	C7H14O	Ketone	105-42-0	LDD	1429707	-0.00461
3	10.49	856	5-methyl-2-hexanone	C7H14O	Ketone	110-12-3	THD	2147159	-0.00434
4	13.25	937	2-methyl-2-butenoic acid, ethyl ester	C7H12O2	Ester	5837-78-5	YAM	274795	-0.00426
5	18.70	1102	benzoic acid, methyl ester	C8H8O2	Ester	93-58-3	THD	2667757	-0.00457
6	21.01	1176	butanedioic acid, diethyl ester	C8H14O4	Ester	123-25-1	YAM	15190975	-0.00404
7	21.80	1201	methylbutanedioic acid, diethyl ester	C9H16O4	Ester	4676-51-1	YAM	52922	-0.00447
8	22.39	1221	3-pyridinecarboxylic acid, ethyl ester	C8H9NO2	Ester	614-18-6	KNA	6499150	-0.00442
9	24.80	1303	difurfuryl ether	C10H10O3	Furan(one)	4437-22-3	TOB	1866908	-0.00519
10	27.52	1401	N-acetylleucine, ethyl ester	C10H19NO3	Ester	4071-36-7	YAM	465680	-0.00539
11	28.95	1455	dimethyl phthalate	C10H10O4	Other	131-11-3	HBS	1710618	-0.00495
12	32.36	1591	4-hydroxy-3-methoxybenzoic acid, ethyl ester	C10H12O4	Ester	617-05-0	TAM	12171194	-0.00403
13	14.80	983	N.I.1				KLS	89193	-0.00494
14	15.92	1016	N.I.2				HSL	9380108	-0.00490
15	16.27	1027	N.I.3				HSL	6691928	-0.00525
16	17.18	1055	N.I.4				HSL	81901	-0.00514
17	19.55	1129	N.I.5				HSL	1348279	-0.00563
18	22.95	1240	N.I.6				SJG	685631	-0.00554
19	23.56	1260	N.I.7				WJS	380569	-0.00498
20	27.12	1386	N.I.8				YAM	348751	-0.00487
21	27.75	1409	N.I.9				YAM	683476	-0.00528
22	27.80	1411	N.I.10				LDD	3124467	-0.00518
23	30.06	1498	N.I.11				TAM	124975	-0.00515
24	32.33	1590	N.I.12				LDD	587077	-0.00486
25	34.50	1683	N.I.13				HBS	391675	-0.00583
26	35.00	1705	N.I.14				PLS	13428900	-0.00534
27	36.13	1757	N.I.15				LPL	2146211	-0.00505
B: Selected markers for acid hydrolysed (AH) soy sauces									
Nr	**R.T.[a] min**	**RI[b] (exp)**	**Name**	**Molecular Formula**	**Chem. Group**	**CAS**	**Samp. Name[c]**	**Max. TIC value**	**Regression coefficient**
1	6.94	752	2-methyl-2-butenal	C5H8O	Aldehyde	1115-11-3	DAP	3972751	0.00559
2	7.88	779	3-methylthiophene	C5H6S	Sulphur compound	616-44-4	PLS	683256	0.00552
3	9.48	826	methylpyrazine	C5H6N2	Pyrazine	109-08-0	SSW	9494707	0.00537
4	12.26	907	methional	C4H8OS	Sulphur compound	3268-49-3	SSW	428404	0.00539
5	12.45	913	2,6-dimethylpyrazine	C6H8N2	Pyrazine	108-50-9	PDS	12548582	0.00543
6	12.59	917	ethylpyrazine	C6H8N2	Pyrazine	13925-00-3	DAP	1455178	0.00532
7	12.65	919	2,3-dimethylpyrazine	C6H8N2	Pyrazine	5910-89-4	SSW	1316767	0.00599
8	15.51	1004	trimethylpyrazine	C7H10N2	Pyrazine	14667-55-1	DAP	3255142	0.00490
9	18.19	1086	2-ethyl-3,5-dimethylpyrazine	C8H12N2	Pyrazine	13925-07-0	DAP	517411	0.00533
10	21.38	1188	2,5-dimethylbenzaldehyde	C9H10O	Aldehyde	5779-94-2	DAP	10879260	0.00554
11	22.59	1228	quinoxaline	C8H6N2	Other	91-19-0	DAP	256354	0.00560
12	25.21	1317	2-methylquinoxaline	C9H8N2	Other	7251-61-8	SJG	281558	0.00552
13	26.06	1348	5-methylquinoxaline	C9H8N2	Other	13708-12-8	SSW	104147	0.00532
14	29.07	1459	N-(4-amino-3-furazanyl)acetamide	C4H6N4O2	Other	140706-47-4	SJS	163748	0.00570
15	12.33	910	N.I.1				SSW	2359853	0.00478
16	13.22	910	N.I.2				DAP	82860	0.00518

17	13.31	936	N.I.3	PDS	58539	0.00486
18	14.84	938	N.I.4	TOB	219753	0.00561
19	18.63	984	N.I.5	SJS	817662	0.00511
20	22.80	1100	N.I.6	SJG	181093	0.00484
21	25.01	1235	N.I.7	SSW	772663	0.00510
22	25.52	1311	N.I.8	SJG	141437	0.00507
23	26.53	1328	N.I.9	THD	5475170	0.00496
24	26.63	1364	N.I.10	SJS	28316367	0.00502
25	29.58	1368	N.I.11	DAP	899356	0.00560
26	35.58	1624	N.I.12	DAP	109617	0.00538
27	40.43	1731	N.I.13	DAP	1121931	0.00485

[a] Retention time (min)
[b] Retention Index (experimental)
[c] Sample code (see **Table 1**) which gave the maximum TIC value for each metabolite.

Volatiles such as pyrazine derivatives, some aldehydes and quinoxaline derivatives were highly (or only) present in AH soy sauces (**Table 2**). Pyrazines in soy sauce can be related to the heating procedures used as they can be formed through Maillard reactions at relatively high temperatures (Lee et al., 2006). Quinoxaline derivatives were present in the AH soy sauces, as well as in the mixed and some of the poorly defined soy sauces from group II (Supplementary **Figure S4E**). Quinoxaline derivatives contain a bicyclic heterocycle formed from a benzene ring fused to a pyrazine ring. 5-Methylquinoxaline, for instance, has been detected in coffee and related products and it is characterized as a burnt, coffee, and corn tasting compound (www.foodb.ca). Quinoxalines have not yet been characterized before in soy sauce. The origins and formation of these compounds are diverse, however, they may be related to Maillard reactions and/or the reduction of amino acids (Mamedov 2016). Lastly, methional had also a high PLS regression coefficient for AH/mixed soy sauces. Its intensity was observed to be higher in DAP and SSW (Supplementary **Figure S4**), but was however, also present in smaller amounts in some of the fermented soy sauces, including the low-salt TFs. The sulfurous compound methional (cooked potato-like odour) has been recently reported as a key aroma compound in Chinese soy sauces (Wang et al. 2020) and is well-known to be generated from methionine by Strecker degradation (Pripis-Nicolau et al. 2000). It has also been detected in a Korean soy sauce after long-term fermentation (Lee et al. 2019). Using acid hydrolysis, this reaction may become more prominent and thus a higher abundance is found in AH soy sauces.

Overall, soy sauces from different production methods could readily be characterized using SBSE-GC-MS and volatiles annotated as being significantly present in both TF and/or AH types, as well as subgroups within. This augured well for the further determination of the possibility to link these chemical differences with sensory relevance and impact. Consequently, the following section focuses on identifying compounds responsible for the aroma characteristics.

3.4 Odour attributes linked to chemical compounds and their potential contribution to overall flavour

The typical aroma of soy sauce is associated with the large variety of volatiles found in appropriate configurations (Nunomura et al. 1976). Next to concentration, the odour threshold per compound is also important in determining sensory impact. Those compounds with a low odour threshold therefore still frequently dominate the overall odour despite perhaps only being present in trace amounts. Recently, the first flavour wheel of soy sauce has been proposed, which compiles the most important taste and aroma

attributes described in the literature, together with the chemical compounds linked to each attribute (Diez-Simon et al., 2020b). However, little is known about the differences in flavour characteristics between soy sauces produced by the contrasting TF and AH methods.

To relate the chemical similarities and differences between TF and AH soy sauces to sensory attributes and hence get a better view of their sensory relevance, gas chromatography-olfactometry-mass spectrometry (GC-O-MS) analysis was performed using four contrasting soy sauces. KNA, LDD, SJG and DAP were selected, based on their differential volatile profiles and contrasting origins. Chemical structural information obtained using GC-MS has been combined with the aroma characteristics as perceived by panellists using an Olfactory Detection Port (ODP). To trap and concentrate volatiles, SPME was chosen since it is more representative of the real-life situation of odour perception.

The list of aroma-active compounds recognized by the panellists using SPME-GC-O-MS is presented in **Table 3**. Most of the aroma-active compounds identified were related to sweet, floral, savoury and smoky descriptors. Benzeneacetaldehyde (floral) and guaiacol (smoky, burnt) were perceived in all samples as 'intense' odorants. These compounds had also been previously described in soy sauce as being "key odorants" with low odour thresholds (Feng, Cai, et al. 2015; Kaneko et al. 2013; Steinhaus and Schieberle 2007; Wang et al. 2020). Phenylethyl alcohol, methional and 3-methylbutanal, characterized by having floral/sweet, cooked potato and malty notes respectively, have also been frequently reported as being key aroma compounds in soy sauce (Devanthi and Gkatzionis 2019; Wang et al. 2020). Moreover, compounds that were perceived by panellists, but which are not yet identified, are also reported here as these represent a new set of metabolites of clear importance to sensory impact. Future attention is needed to ascertain their identities. For example, a certain popcorn aroma was perceived at 16.45 min, which was linked to the non-identified compound number 21, was only detected here in DAP soy sauce (AH) (**Table 3**). This aroma has not been linked to an identified compound in soy sauce yet. However, it has been characterized using GC-O-MS in TF Chinese soy sauces as being an important attribute (Wang et al. 2020). Being only detected here in DAP suggests this compound has higher concentrations in AH soy sauces. Another interesting example is compound #25, characterized by an intense liquorice/candy aroma at an R.T. of 18.90 min (**Table 3**). Liquorice is a unique aroma attribute that has to our best knowledge not yet been characterized in soy sauce. It was perceived in both the TF soy sauces (KNA and LDD) and again, more detailed study is needed to identify its chemical structure. Other compounds have been characterized by having a liquorice-like odour in fermented fish and soy products, such as estragol (aniseed-like, liquorice-like) and 2-pentyl furan (beany, grassy, liquorice-like) (Czerny et al. 2008; Giri et al. 2010) which could be potential candidate compounds.

Table 3. Aroma-active compounds detected using SPME-GC-O-MS in four selected soy sauces. Panellists' evaluations (experimental) are compared to the characteristics reported in the Food database (www.fooddb.ca). A compound was associated with each aroma quality.

No.	R.T min	Aroma quality (experimental)	Aroma quality (FoodDB)	Compound name	O. T. [a] (µg/kg)		Detected in/number of panellists			
							DAP	SJG	LDD	KNA
1	4.46	**candy, sweet**		2-butenal	-	-	x 2			
2	4.77	**malty, savoury**	malty, fatty	3-methylbutanal	1.2	low				x 3
3	9.28	**cheese, pungent, chemical**		Unknown						x 3
4	9.45	**fresh, floral**		Unknown						x 2
5	9.85	**unpleasant, sweaty, cheese**	sweaty, rancid	3-methylbutanoic acid	1200	high		x 2	x 3	x 3
6	10.19	**sweet, liquorice, winegum**		N.L.					x 2	
7	10.60	**fruity, sweet**		1,1-diethoxy-2-methylpropane	-	-				x 2
8	11.20	**medicine, sweet**		methylbutyl acetate	-	-		x 2		x 2
9	12.15	**chemical, acid**		N.L.					x 3	x 3
10	12.47	**cooked potato**	cooked potato	Methional	1.4	low	x 2		x 1*	x 3
11	12.58	**rancid, cheese**		Unknown					x 3	
12	12.74	**sweet**	cocoa, roastbeef, roastednut	2,6-dimethylpyrazine	157.6	medium	x 3			
13	12.99	**popcorn, cooked rice**		Unknown					x 2	x 3
14	13.03	**sweaty, rancid**		N.L.					x 2	
15	14.19	**sweet, pungent**		N.L.					x 3	x 3
16	14.40	**sweet**	almond, burnt sugar	Benzaldehyde	751	medium		x 3		
17	14.90	**mushroom, earthy**		Unknown				x 2	x 3	
18	14.88	**rancid, sulphur**	cabbage, sulphur, cooked	dimethyl trisulfide	0.016	very low	x 3			x 2
19	15.60	**sweet, floral**	n.a.	2-ethyl-6-methylpyrazine	-	-	x 2			
20	15.97	**mushroom, fungal**		Unknown					x 2	
21	16.45	**popcorn, intense**		Unknown			x 2			
22	17.24	**floral**		benzeneacetaldehyde	4	low	x 1*	x 3	x 3	x 3
23	17.40	**caramel, sweet, candyfloss**		Unknown					x 3	x 3
24	18.55	**smoky, burnt**	smoky, medicinal	Guaiacol	1.6	low		x 2		
25	18.90	**liquorice candy**		Unknown					x 3	x 2
26	19.30	**caramel, honey**		Maltol	35000	high			x 3	
27	19.47	**floral, sweet**		phenylethyl alcohol	390	medium			x 1*	x 1*
28	19.60	**burnt, fireplace, carbon**		Unknown			x 2			
29	19.80	**caramel, sweet, candyfloss**		Unknown				x 2	x 3	x 3
30	21.34	**sour**		Unknown						x 3
31	21.39	**savory, mushroom**		Unknown					x 3	
32	22.98	**sweet, unpleasant**	n.a.	3-phenylfuran	5.9	low				x 2
33	23.10	**sweet**		Unknown				x 3		x 3
34	23.31	**floral, caramel**	floral, sweet, anise	ethyl phenylacetate	155.6	medium			x 2	x 2
35	24.15	**cinnamon, spicy**	spice, clove, bacon	4-ethyl guaiacol	16	medium			x 3	
36	25.50	**smoky, spicy**		2,6-dimethoxyphenol	-	-		x 3		
37	26.19	**sweet**		Unknown						x 3

[a] Odour threshold values were collected from the same source (Czerny et al. 2008)

* Only perceived by one panellist (less reliable)

N.L.: attribute not linked to any visible chromatographic peak

In general, most odorants were detected by GC-MS in each of the four soy sauces. However, not all were also perceived by GC-O, as illustrated in **Table 3**. Furthermore, some compounds were detected by GC-O and not by GC-MS (labelled N.L.). This suggests that both abundance and odour threshold play a role in determining whether or not a compound is perceived in a particular soy sauce and hence, has sensory impact or not. Some compound concentrations were too low to reach human detection level in certain samples. This is visible for example when the semiquantitative differences in the GC-MS data are compared with GC-O observations across samples. For example, the attribute smoky, spicy was associated with 2,6-dimethoxyphenol (**Table 3**). 2,6-Dimethoxyphenol was detected by GC-MS in all 4 samples but had a four-fold higher abundance in SJG soy sauce compared to the others (Supplementary **Figure S5A**). However, the characteristic smoky/spicy aroma was only perceived by panellists in sample SJG. This indicates that odour threshold was only reached in the single SJG sample with the highest concentration. Another attribute, characterized by an intense cinnamon and spice-like odour, was associated with 4-ethyl guaiacol and was only observed in LDD where it was also found to have an 18-fold higher abundance compared to the other three samples (Supplementary **Figure S5B**). This compound has previously been reported as a key odorant in Chinese soy sauces (Wang et al. 2020). Lastly, maltol (caramel-like) is a key odorant in soy sauce and it is characterized by having a high odour threshold value (3500 μg/kg). In our study, maltol was detected by GC-MS in all four samples, however was only perceived using GC-O in LDD (Supplementary **Figure S5C**). Equally important, a number of odours were conversely clearly perceived by panellists but were not visible as a peak in the GC-MS trace (annotated as N.L. in **Table 3**). This suggests that the causal compounds are at sub-detection limits but yet are still at levels above their odour threshold. This implies that the human receptors in our nose are more sensitive to these compounds than the Mass Spectrometer used and indicates we cannot fully rely on instrumentation to define sensory experience.

In conclusion, 37 aroma attributes were detected by SPME-GC-O-MS and associated with volatile compounds, of which 19 could be identified and 18 are still unknown. These volatiles also displayed a contrasting aroma pattern between the four tested soy sauces and the presence of some odorants appears to be linked to production origin. This suggests that the aroma of a soy sauce, and hereby potentially also consumer preference, can be linked to the production procedure.

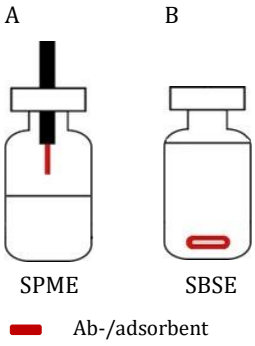
4. Conclusions

This study reports a comprehensive metabolomics analysis of volatiles from soy sauces of different origins, production methods, or using different ratios of ingredients. Under the optimized experimental conditions, SBSE proved to be the extraction method which covered a large number of volatiles detected by GC-MS, in a reproducible manner. The nontargeted approach revealed that the largest difference in the volatile profiles between the twenty commercial soy sauces was related to the production procedure applied: Traditional Fermentation versus Acid Hydrolysis. The AH group was strongly associated with pyrazines, quinoxalines and sulphur compounds, whereas TF soy sauces were more affiliated with esters and ketones. The use of high temperatures during acid hydrolysis likely resulted in higher amounts of Maillard reaction products, while microbial activity in traditional fermentation gave the highest diversity in volatiles. A large

number of non-identified compounds also appeared to be characteristic/unique for either of the two types of soy sauces, which delivers additional potential compounds for further identification.

Characterization of some aroma-active compounds by SPME-GC-O-MS, revealed a contrasting distribution of the detected odorants across the samples related to both technical and human detection thresholds. Further investigation is needed to establish how these differences affect the overall aromas of the soy sauces.

Supplementary material



	SPME (A) Solid-phase microextraction	SBSE (B) Stir bar sorptive extraction
Sampling mode	Headspace	Immersion
Ab-/adsorbent phase	PDMS/DVB/CAR	PDMS
Volume sorbent phase (µL)	0.5	24
Possibility of automating extraction	Fully automated	Requires manual actions
Level of labour intensiveness	Low	Intermediate
Selectivity	A variety of fiber coatings is available (~10) More selective for apolar, highly volatile compounds	Mainly two robust stir bar coatings are available (PDMS and EG-Silicone) More selective for polar, semi-volatile compounds

Figure S1. A schematic representation, together with some important features of the two volatile trapping techniques used in this study. SBSE: stir bar sorptive extraction **(A)**; SPME: solid-phase microextraction **(B)**. The ab-/adsorbent coating is highlighted in red.

Table S1. Results of the trapping techniques obtained from measuring analytical features of experiment 1.

Parameter	SPME	SBSE
Analyte coverage	246	542
Number of non-identified metabolites	159	428
Number of identified metabolites	87	114
Number of common identified metabolites	78	78
Number of unique identified metabolites	9	36

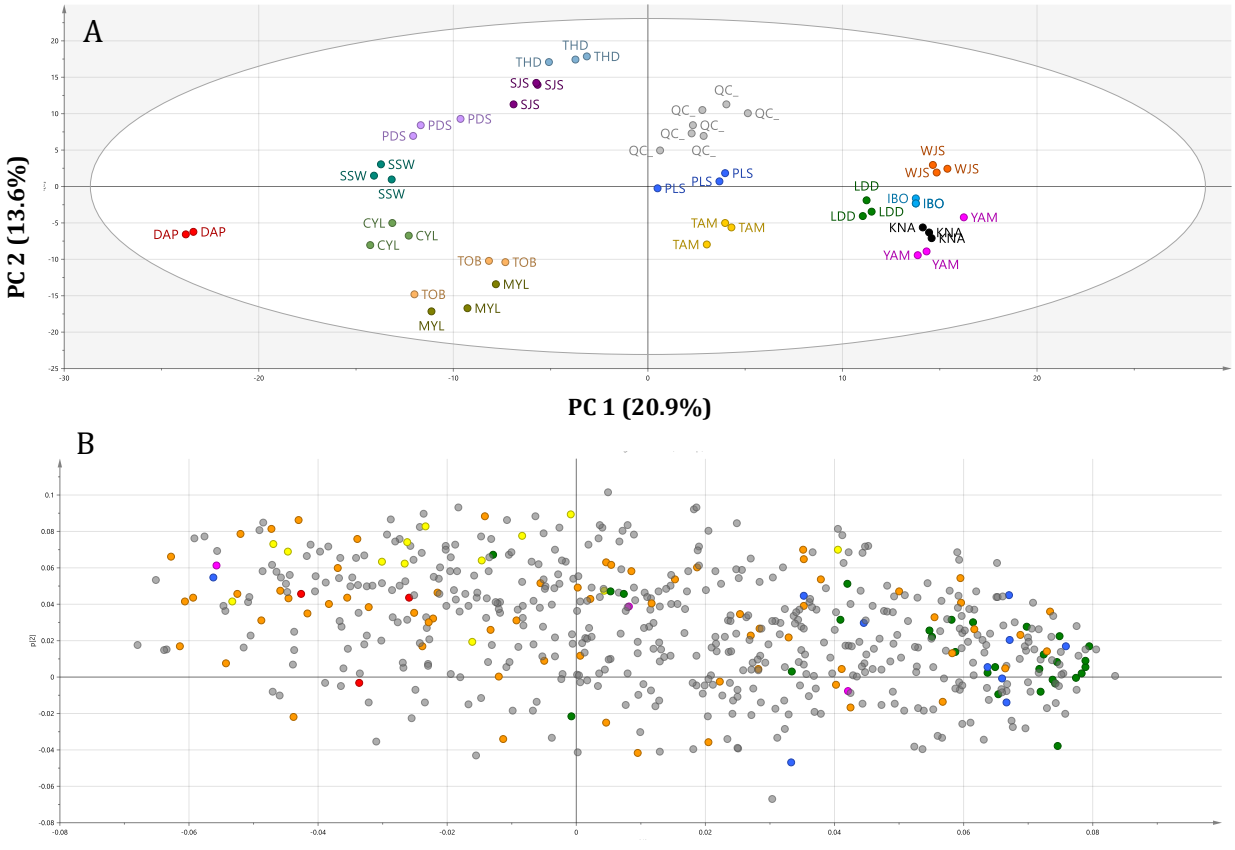


Figure S2. (A) Principal components analysis (PCA) score plot of the 591 volatile compounds of the soy sauces from Experiment 2: Data Puti (DAP, red); Silver Swan (SSW, sea green); Chan's Yellow Label (CYL, lime); Maekrua Yellow Label (MYL, brown); Tai Hua Dark (THD, sea blue); Pearl River Bridge Dark superior (PDS, lavender); Pearl River Bridge Light superior (PLS, blue); Kikkoman natural (KNA, black); Lee Kum Kee Double Deluxe (LDD, green); Wan Ja Shan less salt (WJS, orange); Yamasa (YAM, pink); Inproba Bio-organic (IBO, light blue); Sempio Jin S (SJS, purple); Tamari Gluten-Free (TAM, yellow); and Tamari megachef (TOB, light orange). Three technical replicates are presented for each sample and seven quality control samples (QC) are presented in grey. The first and second PC explain the corresponding percentage of variation shown on each axis. **(B)** Principal components analysis (PCA) loading plot of the volatile profiles of the soy sauces from A using SBSE-GC-MS. Putatively identified volatiles belonging to relevant chemical groups are coloured: acids (red), alcohols (light blue), esters (dark green), pyr(an)ones (pink), pyrazines (yellow), others (orange) and non-identified (grey).

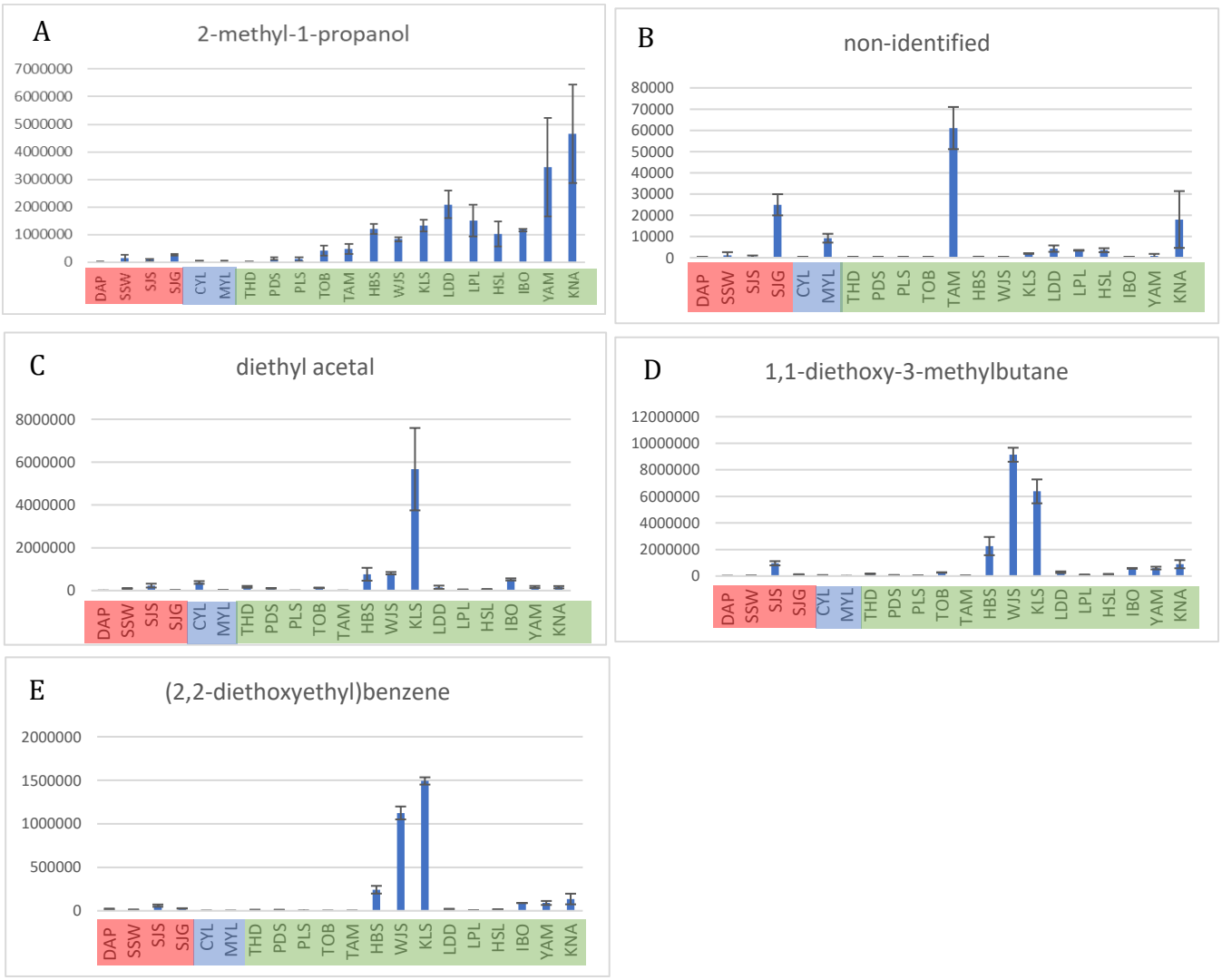


Figure S3. Bar graphs showing examples of volatiles highlighted for Tamari (A and B) and low-salt soy sauces (C, D and E). 2-Methyl-1-propanol **(A)**; non-identified **(B)**; diethyl acetal **(C)**; 1,1-diethoxy-3-methylbutane **(D)**; and (2,2-diethoxyethyl)benzene **I**. The Y-axis shows the TIC intensity of the mean calculated from the technical and biological replicates measured (n=3 or n=6). Red-coloured samples are AH/mixed soy sauces, green-coloured samples are TF soy sauces, and blue-coloured samples are unclassified soy sauces. For a full list of sample codes please see **Table 1**.

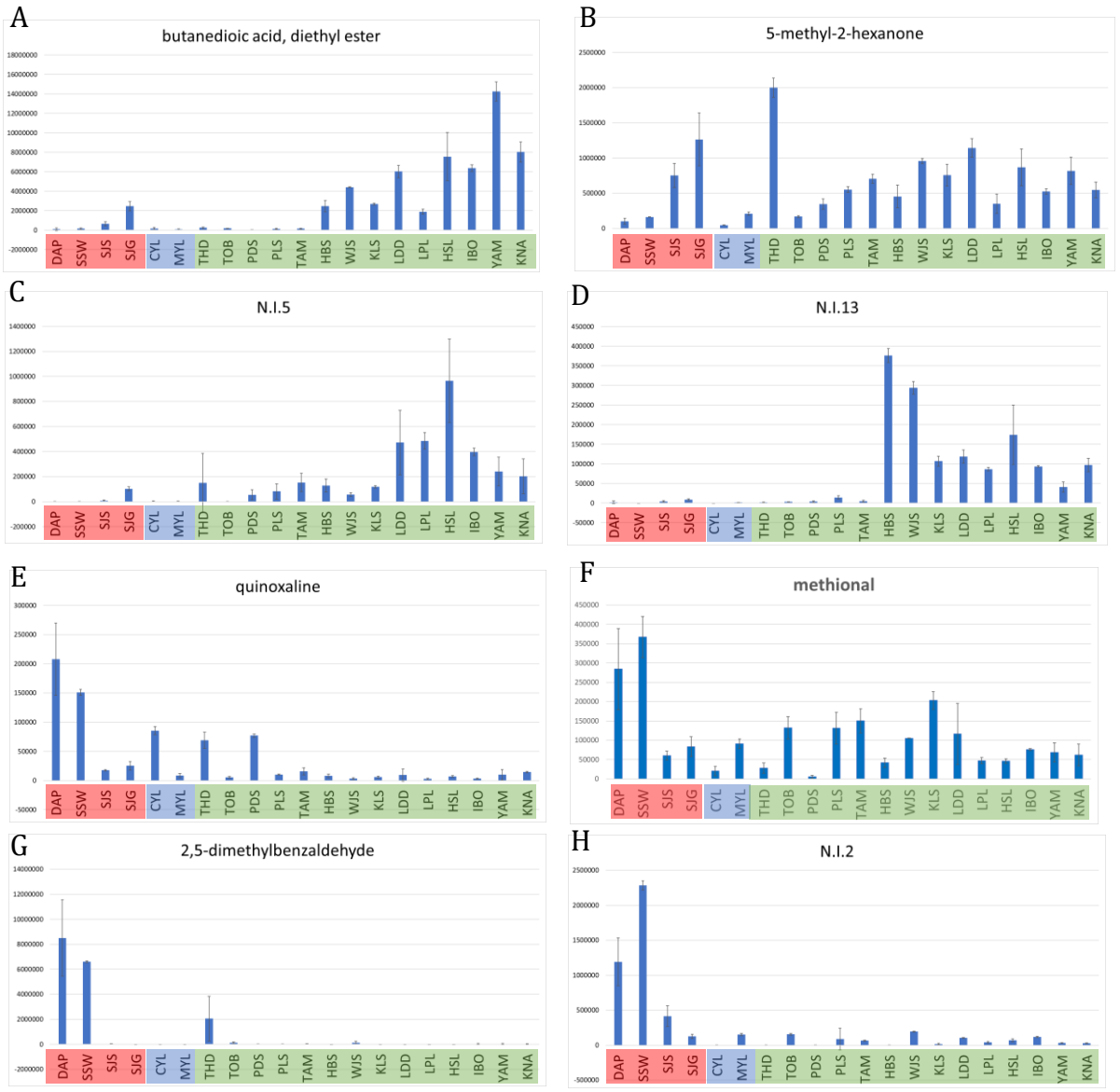


Figure S4. Bar graphs showing examples of potential metabolites being markers for TF (A – D) and AH (E – H) soy sauces. These compounds showed the highest regression coefficient from the compounds selected in **Table 2**. Diethyl ester butanedioic acid **(A)**; 5-methyl-2-hexanone **(B)**; non-identified nr.5 **(C)**; non-identified nr.13 **(D)**; quinoxaline **I**; methional **(F)**; 2,5-dimethylbenzaldehyde **(G)**; and non-identified nr.2 **(H).** The Y-axis shows the TIC intensity of the mean calculated from the technical and biological replicates measured. Red-coloured samples are AH/mixed soy sauces, green-coloured samples are TF soy sauces, and blue-coloured samples are unclassified soy sauces. For a full list of sample codes please see **Table 1**.

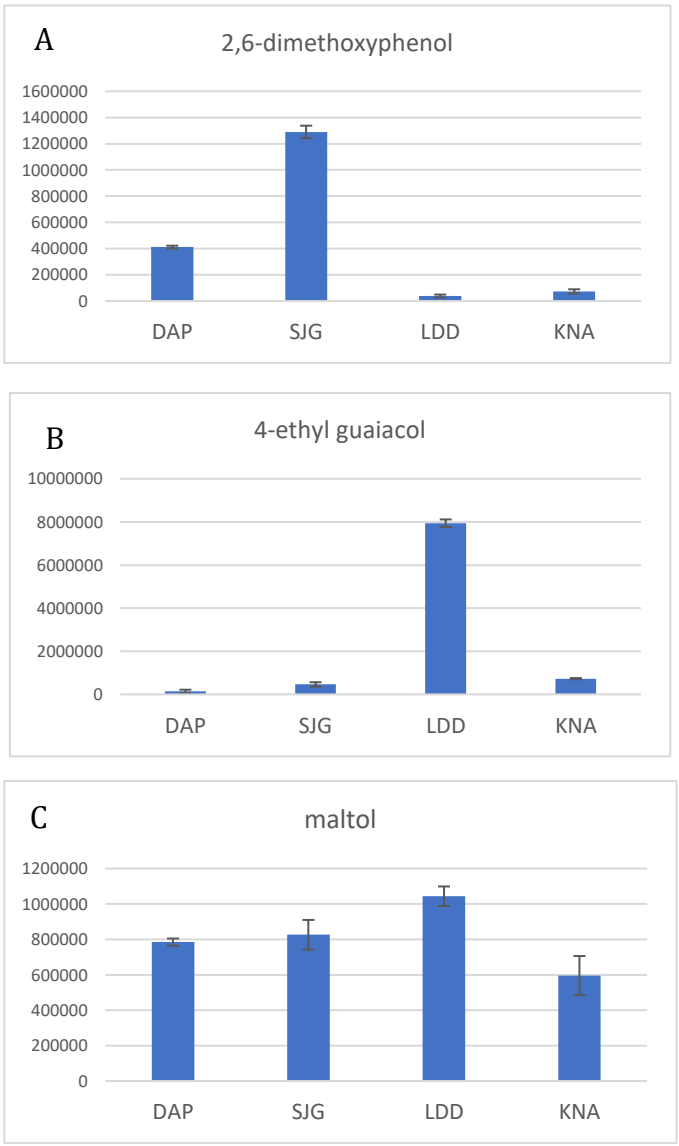
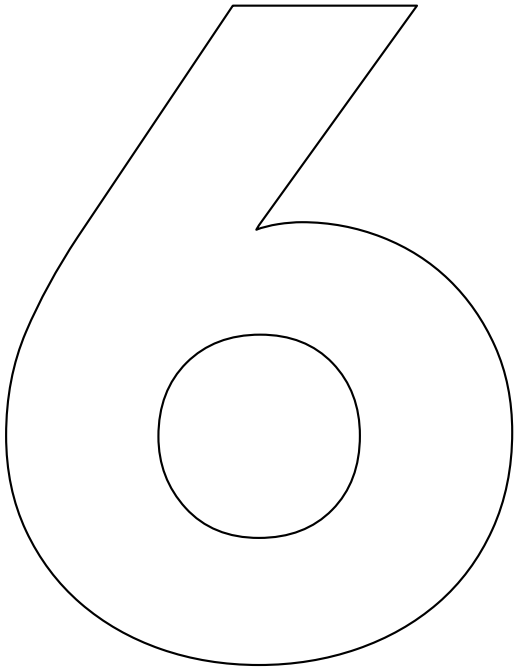


Figure S5. Bar graphs showing examples of key aroma volatiles detected by SPME-GC-(O)-MS in four contrasting soy sauces. 2,6-dimethoxyphenol **(A)**; 4-ethylguaiacol **(B)**; and maltol **(C)**. The Y-axis shows the TIC intensity of the mean calculated from the technical and biological replicates measured. These compounds showed a link between GC-O data and GC-MS data. Data Puti (DAP) is an acid hydrolysed soy sauce; Sempio Jin Gold (SJG) is a mixed soy sauce; Lee Kum Kee Double Deluxe (LDD) is a Chinese-style fermented soy sauce; Kikkoman natural (KNA) is a Japanese-style fermented soy sauce.



Systematic selection of competing metabolomics methods in a metabolite-sensory relationship study

Carmen Diez-Simon, Naser Davarzani, Justus L. Großmann, Doris M. Jacobs, Rudi van Doorn, Marco A. van den Berg, Age K. Smilde, Roland Mumm, Robert D. Hall, Johan A. Westerhuis

This chapter is a modified version of a manuscript published in *Metabolomics* (2021) DOI:

Abstract

Introduction: The relationship between the chemical composition of food products and their sensory profile is a complex association confronting many challenges. However, new untargeted methodologies are helping correlate metabolites with sensory characteristics in a simpler manner. Nevertheless, in the pilot phase of a project, where only a small set of products are used to explore the relationships, choices have to be made about the most appropriate untargeted metabolomics methodology.

Objective: To provide a framework for selecting a metabolite-sensory methodology based on: the quality of measurements, the relevance of the detected metabolites in terms of distinguishing between products or in terms of whether they can be related to the sensory attributes of the products.

Methods: In this paper we introduce a systematic approach to explore all these different aspects driving the choice for the most appropriate metabolomics method.

Results: As an example we have used a tomato soup project where the choice between two sampling methods (SPME and SBSE) had to be made. The results are not always consistently pointing to the same method as being the best. SPME was able to detect metabolites with a better precision, SBSE seemed to be able to provide a better distinction between the soups.

Conclusion: The three levels of comparison provide information on how the methods could perform in a follow up study and will help the researcher to make a final selection for the most appropriate method based on their strengths and weaknesses.

Keywords: volatiles; stir bar sorptive extraction (SBSE); solid-phase micro extraction (SPME); sensory analysis; tomato soup; yeast product; chemometrics

1. Introduction

One of the new trends in food science is to relate the chemical composition of a food product to its sensory attributes at the level of individual metabolites or groups of metabolites (Seisonen et al. 2016). These relationships, that will improve our understanding of flavour formation due to processing and the supplementation of ingredients, are of great relevance to food design/formulation studies. Examples can be found for various products such as wine (Malherbe et al. 2013), olive oil (Procida et al. 2016), apples (Corollaro et al. 2014) and dairy products (Croissant et al. 2011).

The multivariate sensory profile of a food product can contain information on its appearance, flavour (aroma and taste), mouthfeel and texture which are usually assessed by an expert panel. In an experimental setup, expert panellists "measure" the products by scoring them for a number of pre-defined contrasting sensory attributes. The chemical content of the products can be measured by various analytical methods. Metabolomics is a common term used for the comprehensive characterization of the molecules present in a biological sample (Cubero-Leon et al. 2014). Most importantly, metabolomic analyses can help to compare accurately the metabolite profile between groups of contrasting samples and thus to identify discriminatory compounds and subsequently enabling us to correlate chemotype determined by the metabolic space and phenotype as defined in the sensory space. A large number of analytical techniques are suitable for measuring the broad range of metabolites. Mass spectrometry platforms are commonly used, coupled with gas chromatography (GC-MS) or liquid chromatography (LC-MS) as separation techniques. Within these two hyphenated approaches there is still a large variety of possibilities with respect to the choice for e.g. sample preparation and sample extraction (Thomsen et al. 2016). Furthermore, methodologies can be chosen which focus on a specific group of metabolites (targeted approach) or are aimed to detect as many metabolites as possible (untargeted approach) without any prior knowledge on the identity of the molecules (Alonso et al. 2015). The latter allows us to get a more comprehensive metabolite profile which can make associations with sensory data stronger.

The relationships between chemical compounds and sensory impact may entail various challenges which play an important role in gaining a complete mechanistic understanding. These challenges are for example related to masking effects regarding individual compounds, nonlinearities or saturation effects and the presence of low abundant compounds which nevertheless have low odour thresholds. Such compounds play an overly-influential role in determining the aroma experience and are essential for the overall sensory characteristics (Diez-Simon et al. 2019; Yang et al. 2015). Moreover, current prediction models in the relationship between chemical composition and the sensory profile (Aprea et al. 2012; Calingacion et al. 2017; Esslinger et al. 2014) should be expanded to cover a larger number of compounds (untargeted approach) and multiple sensory attributes, as well as combining complementary analytical platforms. Knowledge of such approaches will help industry enormously in developing new product formulations with a predefined sensory profile. Because of the large number of possible metabolomics methods to choose from in a chemical sensory relationship study, it is often difficult to decide which method to use to explore optimally the existing metabolite sensory relationships. In a pilot study, one could assess multiple methods to decide which one best suits the project at hand, particularly if the methods provide partly overlapping

metabolite profiles. However, there are several issues that need to be taken into account while selecting the optimal metabolomics method in a given project.

In this paper we introduce a methodology to select between potential metabolomics methods for a given metabolite-sensory relationship project using untargeted approaches. We use the term methods very generally – it could mean two different metabolomics platforms, but it could also mean two different separation columns in an LC-MS, or two different sample extraction approaches. Such a decision between methods often has to be made in an early phase of the study based on only pilot data, where the consequence of this decision will become apparent in the larger study, for which similar products are expected. Here, the study presented is an example of such a pilot study. In the larger follow-up study, new products will be tested, which are, although related, different from the pilot study samples. Therefore, it is not sufficient to only focus on how well the sensory attributes could be predicted from the metabolite levels using only the pilot products. Instead, the potential application of the two metabolomics methods for new products has to be assessed.

To introduce the methodology, we will use a tomato soup flavour study, where two different sample extraction techniques for volatile and semi-volatile aroma compounds were used in combination with GC-MS: solid-phase micro extraction (SPME) and stir bar sorptive extraction (SBSE). These two approaches detect partly the same volatile compounds, but also one may be able to detect compounds that cannot be detected by the other method. Twenty-seven different tomato soup samples were prepared with different recipes according to a well-defined experimental design (**Table 1**). The soups were assessed for their aroma and taste profile by a trained sensory panel and analysed by SPME- and SBSE-GC-MS. In the theory section we will lay down the methodology to select the most useful of the two analytical methods that best characterises the variation applied to the tomato soup flavour. The methodology comprises three parts that compare 1) the methods based on their analytical figures of merit, 2) the variation induced in the metabolite levels by the different products and, finally, 3) the quality of the relationship between the metabolite levels and the sensory attributes. In the results section we will present exemplary results of the comparison steps outlined above. Note that our aim is not to select the best method for this specific tomato soup project, but to outline a strategy that aids the decision-making process.

2. Theory

The comparison between the analytical platforms is carried out on three levels: (1) comparison of analytical figures of merit; (2) comparison of how well the measured metabolite levels are able to distinguish between the different products and finally, (3) how well the variation in the metabolite levels over the products are able to model the variation in the sensory attributes. The information obtained from each of these levels can then be combined to make a final decision with respect to the most suitable method in the study.

Table 1. Compositional factor levels of the tomato soups used in this study.

Soup nr.	Tomato dosage	Oil type	Oil dosage	Yeast derived product	Yeast dosage	Included in QCs
1	low	olive	high	-	-	
2	low	olive	high	Maxarome	low	
3	low	olive	high	Maxarome	high	
4	low	olive	high	Maxavor	low	
5	low	olive	high	Maxavor	high	
6	low	olive	high	Maxagusto S-99	low	
7	low	olive	high	Maxagusto S-99	low	Included
8	low	olive	high	Maxagusto S-99	high	
9	low	olive	high	Maxagusto S-99	high	
10	low	olive	high	Maxagusto O-31	low	
11	low	olive	high	Maxagusto O-31	low	Included
12	low	olive	high	Maxagusto O-31	high	
13	low	olive	high	Maxagusto O-31	high	
14	low	olive	high	Maxagusto G-28	low	
15	low	olive	high	Maxagusto G-28	high	
16	high	olive	low	-	-	
17	high	olive	low	Maxagusto S-99	high	Included
18	high	olive	low	Maxagusto O-31	high	Included
19	high	olive	low	Maxagusto G-28	high	Included
20	low	corn	high	-	-	
21	low	corn	high	Maxagusto S-99	high	
22	low	corn	high	Maxagusto O-31	high	
23	low	corn	high	Maxagusto G-28	high	
24	high	corn	low	-	-	
25	high	corn	low	Maxagusto S-99	high	Included
26	high	corn	low	Maxagusto O-31	high	Included
27	high	corn	low	Maxagusto G-28	high	Included

2.1 Comparison of analytical figures of merit (Level 1)

Traditionally, the quality of analytical platforms is assessed using analytical figures of merit such as repeatability, linear range, sensitivity, selectivity, etc. All these figures of merit are defined for single specific compounds. However, current analytical platforms provide levels of many metabolites simultaneously. The comparison of figures of merit between different methods can be performed in two ways: *metabolite-independent* and *metabolite-dependent*. The *metabolite-independent* approach calculates a specific figure of merit for all metabolites detected by the method. As an example, the repeatability can be selected as the figure of merit of interest. Then, for each method, a density plot or a histogram gives information in general about the repeatability of the whole set of metabolites in the two methods. Such a figure could provide information on the difference between platforms on how many metabolites can be measured with a specific repeatability. In contrast, the *metabolite-dependent* approach considers only those metabolites that are measured in both methods (these are called common metabolites). For each of these metabolites, the figure of merit value is directly compared between the methods. If e.g. the repeatability for most of these common metabolites is better in SPME, then that would be a good reason to select SPME over SBSE as the trapping method. Note that it is not necessary to know the identity of the metabolite, only that they are identical.

For the *metabolite-dependent* approach, the common metabolites must be defined first. A comparison of common metabolites detected in both methods will focus on their figures of merit: which method is able to detect these metabolites with the highest accuracy. The comparison of the metabolites that are unique will rather focus on which of the unique metabolites give a broader view of the differences between the products and whether they are able to improve the prediction of sensory properties of the products.

To find the common metabolites, we use the spectral information available for each metabolite, i.e., mass to charge ratios *m/z* (parent ions and ion fragments) and their corresponding relative intensities, together with its retention time index. In the following, we show for two mass spectrometry platforms how to find the common metabolites. The approach works as follows:

1. For a given metabolite in SPME data, select those metabolites in the SBSE dataset for which their retention indices have a smaller difference than a user-defined threshold value. This is done to make a preselection of potential candidates that could be the same metabolite. In our example project, this value was set to two units, but can be adjusted depending on the methods used. If the methods are rather different, then this step could be omitted.

2. Suppose that for a given SPME metabolite (A) with N_A masses and intensities $I_{A1}, \dots, I_{A_{NA}}$ an SBSE candidate metabolite (B) with N_B masses and intensities $I_{B1}, \dots, I_{B_{NB}}$ was found in the first step. The masses that are present for both the SPME compound and the SBSE candidate are called common masses. Let's assume K common masses exist when comparing metabolites A and B. The similarity S_{AB} is defined as:

$$S_{AB} = \sum_{k=1}^{K} I_{Ak} * I_{Bk} - \sum_{i_A=K+1}^{N_A} I_{Ai_A}^2 - \sum_{i_B=K+1}^{N_B} I_{Bi_B}^2 \qquad [1]$$

The first summation in S_{AB} is over the K masses of A and B that are in common, while the second and third summations are over the masses that are unique in A and in B, respectively.

3. Select the candidate SBSE metabolite with the highest similarity S_{AB} to the given SPME metabolite. As a threshold for the similarity we used the value 0. Thus, the candidate B with the highest similarity with metabolite A is considered to be the same metabolite if the similarity is larger than 0.

4. Repeat the analysis starting with each SBSE metabolite and use the same procedure to find the optimal SPME match. Only if this reverse analysis finds the same pair of metabolites, they are considered common metabolites.

2.2 Comparison of ability to distinguish between product preparation differences (Level 2)

In projects that aim to build relationships between sensory and metabolic profiles of a product, usually a set of contrasting "pilot products" is developed based on an experimental design strategy in which certain ingredients are added or certain process steps have been taken. The formulations of the pilot products should be defined with the aim of ensuring sufficient variation in their metabolic as well as in their sensory profile. In this second level we will analyse the relationships between the metabolite levels of the products

and the compositional factors, i.e. the different ingredients and processing steps that are expected to influence the molecular composition of the products. This is a discrimination problem between the products that are made in a specific manner and the other products. Such a discrimination problem can be explored in a univariate manner, where for each metabolite we test its performance in discriminating between the groups, as well as in a multivariate manner where also the correlation between the metabolite levels are considered in the discrimination model. Multivariate discrimination methods that can be used to quantify how well the specific platform is able to discriminate between the two groups of products are for example PLS-DA, SVM and random forest classifiers (Lee et al. 2018; Liu et al. 2013).

2.3 Comparison on ability to predict sensory attributes (Level 3)

In the final level of comparison we explore how well the metabolite levels of the different platforms are able to predict the sensory attributes of the different products. Here, a predictive multivariate model is built between metabolite levels and sensory profiles to see which platform is best able to predict the sensory profile as well as which of the metabolites has important roles in these models. For this step, a multivariate regression model has to be used as the sensory attribute is usually a quantitative feature. Examples for such models are SVR (Sugimoto et al. 2010), PLS (Grabež et al. 2019), penalized multivariate regression such as the LASSO or Elastic Net Regression, or Random Forest (Welzenbach et al. 2016).

Note that one could assume that only level 3 is important as this aspect reflects precisely the goal of the project, namely the metabolite-sensory relationship. However, the comparison is done on a pilot scale, and should provide more information to select new samples for the larger study. Therefore, the check whether there is sufficient variability between the samples that can be observed with the platform (level 2) is important similarly whether the platform is able to quantify the measured metabolites with sufficient quality (level 1).

3. Experimental

The specific case study we will use in the paper to demonstrate the methodology is from a project in which the volatile metabolites in 27 different tomato soup products had to be related to various sensory attributes. The tomato powders (Unilever R&D, Wageningen, The Netherlands) consisted of: tomato powder, sucrose, roux, starch, oil, salt, lemon juice, pepper and yeast-derived flavour products (DSM, Delft, The Netherlands) (or nothing, in case of blanks) (**Table 1**). The main difference between the powders were the type of oil used (corn vs. olive), the tomato dosage (high vs. low), the type of yeast product and its dosage (high vs. low). The soups were prepared by stirring 70 - 99 g of the tomato mix powder into 1L of boiling water. The exact amounts depended on the formulation. The soups were gently simmered for 5 minutes and occasionally stirred.

3.1 Sensory analysis description

Investigating the effects of odour, flavour, mouthfeel, aftertaste, and after-feel requires an extended evaluation of the products and this was done by Quantitative Descriptive Analysis (QDA). During the QDA

measurement, the intensities of the attributes were obtained by EyeQuestion (Logic8), using unstructured line scales ranging from 0 – 100. A very experienced (>10 years) group of panellists (n=14) assessed the 27 different products divided over 4 sessions. For each session, all 27 products were freshly prepared, and offered one-by-one to the panellists according to an incomplete, balanced design that was specifically developed to assess all products in each session. The products were stored in a holding cabinet at 60°C prior to sensory analysis. Fifty mL of each tomato soup was provided to the panellists in a white polystyrene cup. Overlap between the sessions was ensured by having replicate samples across the different sessions which were tasted by several panellists.

The variation in assessment of the different products varied greatly between the different assessors. To correct for difference between assessors in terms of level effect and scaling effect, a standardization of the sensory data was used. The intensity assessed for attribute *a* by panellist *p* for soup *s* (y_{aps}) is standardized to $\tilde{y}_{aps}$ as follows:

$$\tilde{y}_{aps} = \frac{y_{aps} - m_{ap}}{s_{ap}} \qquad [2]$$

where m_{ap} is the mean intensity for attribute *a* and panellist *p* of all soups and s_{ap} is the standard deviation for attribute *a* and panellist *p* of all soups (Romano et al. 2008). This standardization corrects for panellists that, on average, give higher or lower intensity values for all products or use different ranges of the rating scale than the average panellist (e.g. from 40 to 55 instead of from 20 to 70).

3.2 Metabolomics platforms

For the determination of the metabolite levels, two different approaches (SPME and SBSE) were used in which the extraction of the metabolites from the samples differed. After extraction, analytes were thermally desorbed and analysed by GC-MS. The same GC-MS instrument and settings were employed for both extraction techniques. Extraction and analysis of both techniques followed the same procedure as described before by Diez-Simon et al. (2020a). The only variation from the previous study was the temperature used, in both SPME and SBSE, to desorb the analytes into the GC. In this study 280 °C, instead of 250 °C, was applied in order to extract the less volatile, higher molecular weight compounds which are present in the complex tomato soup matrices.

3.3 Experimental setup

For both techniques, a set of 27 tomato soups were analysed in a randomized way. An empty glass vial and a blank (water) sample were measured at the beginning of each sequence. Quality control samples (QCs), which were a mix of a few selected tomato soup samples (See **Table 1**), were repeatedly analysed along the sequence to test the performance of the analysis. Five QCs were analysed in SPME series, once every six samples and four QCs for SBSE, after every ninth sample. The raw GC-MS data were processed using an untargeted metabolomics approach, which has also been described in detail before (Diez-Simon et al. 2020a).

4. Results and Discussion

For the two trapping techniques the number of detected metabolites after processing of the data was different. With SPME a total 331 metabolites were detected, while 482 metabolites were detected with SBSE. For each of the detected metabolites a mass spectrum is available. By matching the obtained mass spectra and retention indices to either our authentic reference standards or those in the NIST17 Mass Spectral library (v.2.3), 100 and 110 metabolites were annotated in the SPME and SBSE datasets, respectively. Before analysis on how well the metabolomics data is able to distinguish between products and to relate to the sensory attributes, a square root transformation was applied to correct for heteroscedasticity.

In this paper we will focus on two sensory attributes, the odour intensity (intensity_OD) and umami flavour (umami_FL). The former is an attribute highly affected by the volatile metabolites, while the latter has been described to be mainly related to non-volatile glutamic acid and some nucleotides, compounds that were not targeted with the current analysis protocols. Statistical testing of only the sensory data yielded F-test values for differences between products of $3.5E^{-13}$ for odour intensity and 0.03 for umami flavour. Thus, there is sufficient variation in odour intensity, but limited variation between umami levels of the 27 soup products. Additionally, as an illustration of application, the predictive performance of the garlic flavour attribute (Garlic FL) was evaluated and volatiles that correlated with garlic flavour were selected.

4.1 Level 1

For the first level of the comparison, we selected to analyse the repeatability using the relative standard deviation of the processed metabolite data. The repeatability was obtained from the QC samples that were measured throughout the batches. Furthermore, we used the concentration range (defined as the difference between maximum and minimum intensity value) found for the common metabolites in the soup product samples. A larger range across these soup product samples relates to a higher sensitivity for those compounds as the product samples in both methods are the same.

4.1.1 *Metabolite-independent* comparison

For the metabolite independent approach, we calculated the relative standard deviation (RSD) which is the ratio of the standard deviation over the mean value obtained from multiple measurements of the QC sample in both methods. For the SPME method, five QC measurements were available while for the SBSE method four QC samples were available.

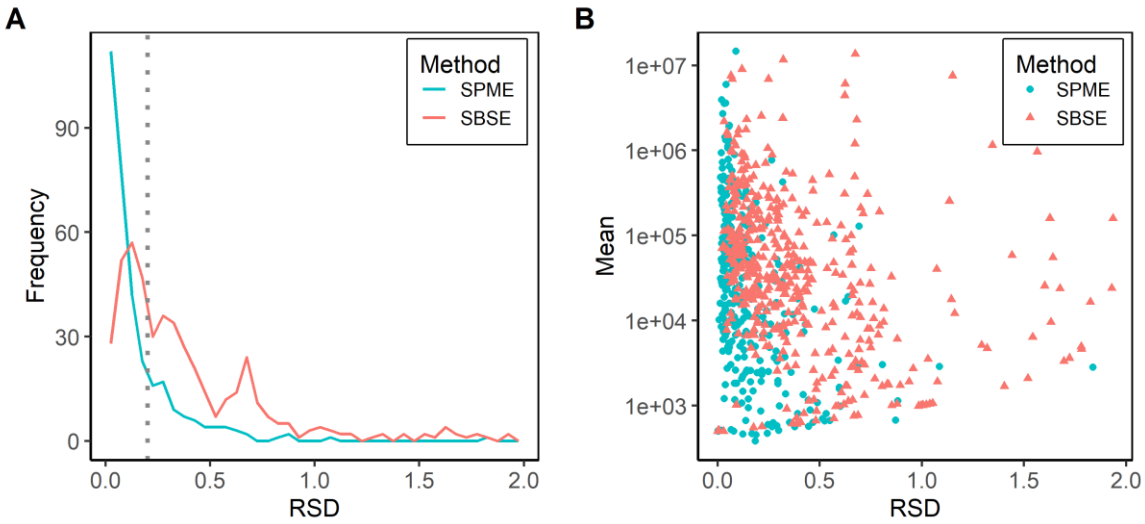


Figure 1. *Metabolite-independent* analyses on QC measurements. **(A)** Histogram of relative standard deviations for metabolites measured using SPME and SBSE. **(B)** Mean intensity over QC samples of the metabolites in logarithmic scale as a function of their RSD.

Histograms of the RSD in SBSE and SPME are shown in **Figure 1A**. The dotted line in the density plot indicates an RSD value of 0.20, which is sometimes used as a threshold for whether metabolites will be reported or not (Siskos et al. 2017). The fraction of metabolites that were measured with a low relative standard deviation (RSD) is much higher for SPME than for SBSE. A large number of metabolites detected using SPME has an RSD value below 0.20, pointing to many metabolites that can be quantified rather precisely, whereas for SBSE, a much larger group of metabolites were measured with an RSD > 0.2. Moreover, many of the SPME metabolites with small RSD have relatively high mean values (**Figure 1B**), and thus could be very useful for the distinction between the soups.

4.1.2 *Metabolite-dependent* comparison

In the second part of level 1 comparison we focused on the common metabolites found using both methods. To find metabolites that are present using both approaches, both the mass spectral information as well as their retention information were used, as was discussed in Section 2.1.

In the first step, for each compound in SPME we preselected those compounds in SBSE with a retention index deviation of up to 2 units. For this preselected set of compounds, the similarity index of the respective mass spectra (see eq 1) was calculated, and the compound having the maximum similarity was selected given that the similarity value is > 0. Subsequently, the approach was also applied vice versa starting with compounds detected in SBSE and matched with those in SPME.

The procedure selecting common metabolites in the two extraction methods was validated using the annotated metabolites in the SPME and SBSE datasets. Out of the 61 annotated compounds that were present in both datasets, 59 were correctly detected to be common whereas zero out of 88 non-common compounds were falsely detected to be common. The two compounds that were not found to be common by our approach were rejected because of a too large difference in retention index due to transfer issues for these two compounds which eluted at the beginning of the chromatogram.

In addition to the 61 annotated common metabolites, application of the matching procedure followed by manual verification yielded another 67 metabolites that are common but not annotated. Using these 128 common metabolites, the RSD and the range comparison for the SPME and SBSE methods were performed. The RSD scatterplot (**Figure 2A**) shows more dots above the diagonal line, indicating that many of the common metabolites have higher RSD values for SBSE than for SPME. The range of the common metabolites is quite similar for both SPME and SBSE methods (**Figure 2B**). Thus, although the repeatability seems better for the SPME method, the sensitivity for the metabolites that are in common is similar for both methods.

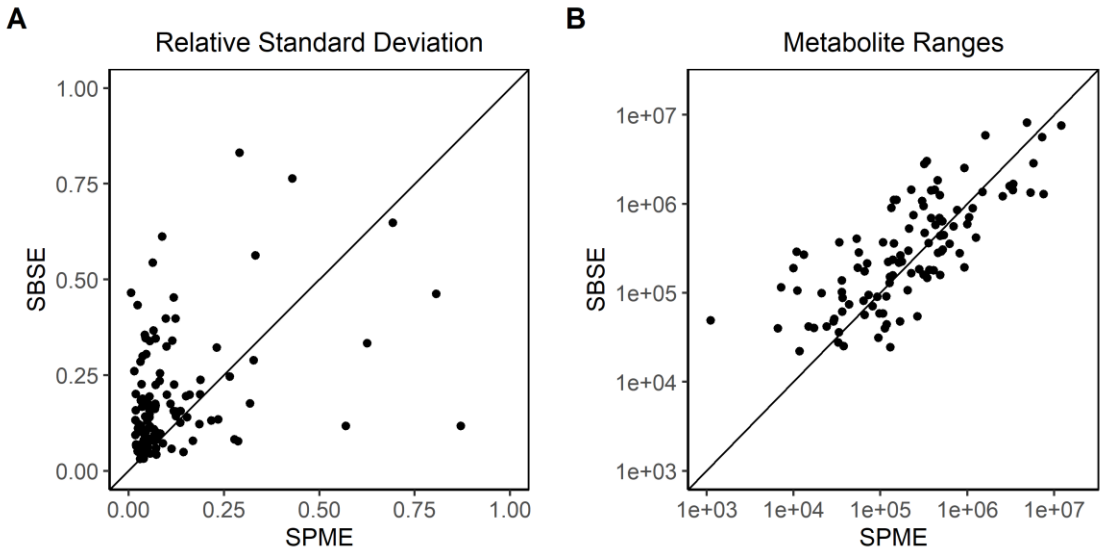


Figure 2: *Metabolite-dependent* analyses on common metabolites. **(A)** Relative standard deviations of common metabolites between QC measurements, compared between SPME and SBSE. **(B)** Ranges of common metabolites compared between SPME and SBSE represented in log scale.

4.2 Level 2

For the second level of comparison we used discrimination methods to investigate how well the metabolites were able to distinguish between groups of products that were produced using different recipes. This can be applied using univariate approaches that present how many of the metabolites are able to distinguish between the groups by themselves, or in a multivariate manner where combination of metabolite levels can be used to discriminate between the groups of products. For the univariate approach we calculated the Point-biserial correlation (Sheskin 2003) between the group indicator variable and the metabolite levels. The Point-biserial correlation is comparable to the Pearson correlation coefficient, but corrects for the fact that the group indicator variables have a dichotomous structure (Sheskin 2003). For multivariate discrimination, we applied Partial Least Squares-Discriminant Analysis (PLS-DA). This method uses a latent variable model (using many correlated metabolites simultaneously) to discriminate the two groups in the data.

To show how well the two different platforms detect metabolites that can discriminate soups made with different recipes we considered two exemplary compositional factors. In the first example we compared soups containing the PF O31 vs all other soup types, and in the second example, soups produced using olive

oil were compared to soups made using corn oil. Note that the metabolomics data was square root transformed.

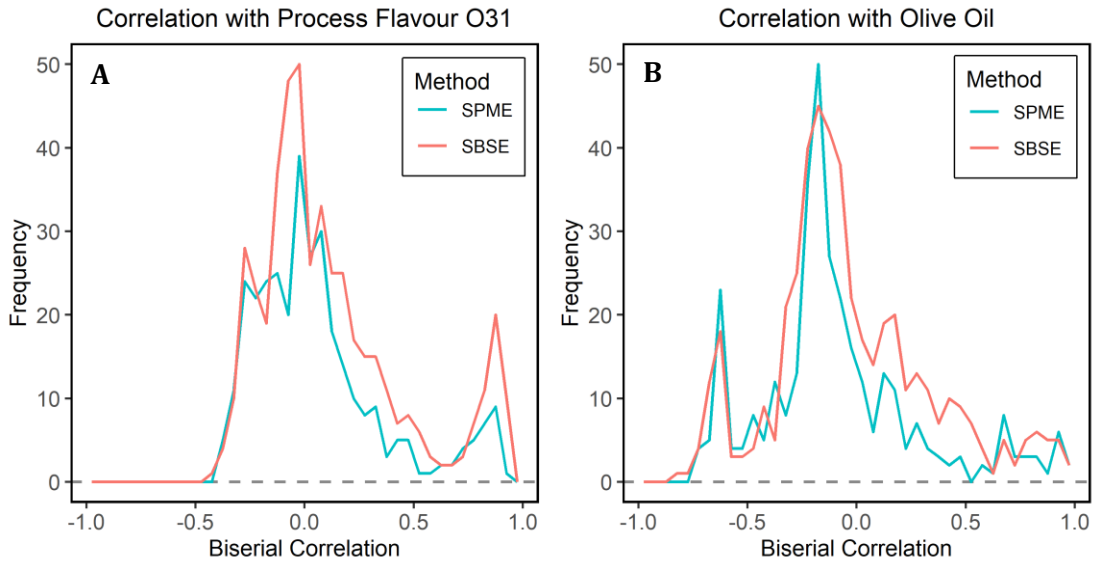


Figure 3. Univariate analysis of the relationships between metabolite levels and compositional factor levels. Histograms of biserial correlation between the metabolite levels of the SBSE and SPME platforms and the group indicator of O31 vs the other soups **(A)** and between soups made with olive oil vs corn oil **(B)**. Here positive correlation means the peak is higher for olive oil samples, while negative correlation means the peak is higher for corn oil samples.

In the biserial correlation coefficient histogram plots for the O31 vs others discrimination (**Figure 3A**) and the olive oil vs corn oil discrimination (**Figure 3B**), the highest peaks were found near a correlation of zero, indicating that most metabolites in both platforms are not correlated to the O31 status nor to the oil status. However, in the O31 correlation histogram plot we see a small peak around a correlation of 0.9 for both platforms, indicating some metabolites with a high positive correlation with O31. For SBSE this peak is much higher than for SPME, suggesting that the SBSE method data contains more metabolites indicative for O31 than the SPME method. For the oil status, the number of metabolites with a high absolute correlation value is much lower, and virtually no difference can be observed between SPME and SBSE. The histogram also reveals a distinct group of compounds with a correlation of around -0.65 that were only present in samples with high tomato concentration in combination with corn oil. A possible explanation may be that these volatiles are only present in corn oil and more released in samples with low oil concentrations. Previous studies have shown the release of fat-soluble flavour compounds is dependent on the oil content (Patana-Anake 2015). Further studies are needed to explore such interaction effects.

For the multivariate PLS-DA models, the O31 status and also the olive oil status were predicted from the metabolite levels of SBSE or SPME platforms. To quantify the model performance, we use the Balanced Error Rate (BER), which is the average of the False Positive Rate and the False Negative Rate (Rohart et al. 2017), and should be as small as possible. The BER is useful when the groups in the discrimination model are of unequal size, which is the case for our examples. The BER is not a relative measure, thus it can be used to compare the SBSE and SPME models for the same response (e.g. O31) but not for comparing models for

different responses. To indicate metabolite importance in the PLS-DA model we used the selectivity ratio (Rajalahti et al. 2009).

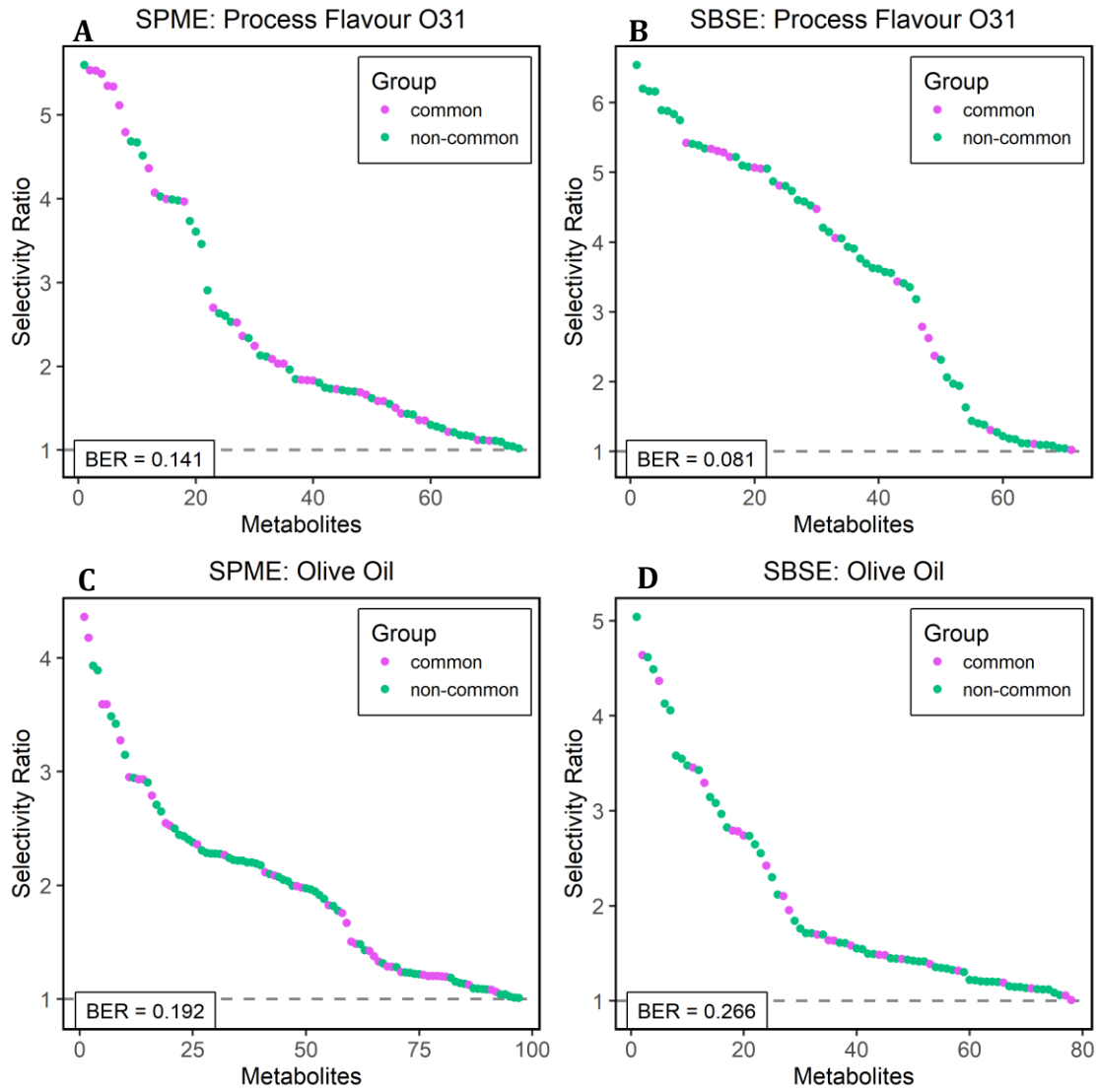


Figure 4. Multivariate analysis of the relationships between metabolite levels and compositional factor levels using PLS-DA. The plots show selectivity ratios and balanced error rates (BER) of the PLS-DA models for process flavour O31 using the SPME **(A)** and the SBSE **(B)** data, and olive oil status using the SPME **(C)** and the SBSE **(D)** metabolite data. The selectivity ratio (only variables with a selectivity ratio above 1 are shown) reflects the discriminatory power of each variable, whereas the BER is a measure for the overall discriminatory performance of the models.

Selectivity ratios for metabolites in the SPME and SBSE data, as calculated from the PLS-DA models, and the balanced error rate (BER) values are provided in **Figure 4**. The pink coloured dots represent metabolites which are in common in both methods, while the green ones are unique per method. It can be observed that the O31 status is predicted better using SBSE than using SPME, as indicated by the lower BER value. For the olive oil prediction, the SPME model performs slightly better. Inspection of the metabolite selectivity ratios reveals that for the SBSE models many of the highly predictive metabolites are unique to the SBSE dataset.

Thus, the use of SBSE enables quantification of a number of metabolites that are discriminative for the O31-containing soups as well as the olive oil containing soups, but do not appear in the SPME dataset.

4.3 Level 3

The third level of comparison between the two platforms is to assess which set of metabolites is better able to predict the sensory properties of the different soups. A multitude of sensory attributes were assessed by the expert panel. We selected two different attributes (intensity_OD and umami) to see how well these can be predicted using metabolite level data from the SPME or SBSE platforms. Odour intensity is expected to relate strongly with volatile metabolites while umami flavour is not.

As an example, we used an elastic net regression model approach to make predictive models for the selected attributes using the square root transformed metabolite data from SPME and SBSE. The elastic net model is a penalized regression model that selects only metabolites that have a sufficiently large effect on the prediction of the attribute. The tuning parameters for the elastic net models were optimized for each platform using cross validation. The mean squared error (MSE) of the prediction as a measure of model performance is obtained using cross validation. The MSE is not a relative measure as it depends on the size of the response value. Therefore, MSE values of models for the same response can be compared, but MSE values for different responses cannot.

The observed values for odour intensity could be well predicted using SBSE, while the prediction performance using SPME is limited (**Figure 5A-B**). Metabolites most important to predict odour intensity using SBSE were 2,6-diethyl-pyrazine (Koehler et al. 1971), methylbutanal and eugenol (Patana-Anake 2015). For umami flavour, as expected, both models had difficulty making good predictions (**Figure 5C-D**).

One more attribute was included to check for the model performance and to identify the selected variables for garlic flavour. Garlic flavour is an important attribute in processed products, such as tomato soups, and its inclusion has also been linked to ingredients that enhance sensory richness and lingeringness (Inoue et al. 2016). In this case, the observed values could be well predicted using both SBSE and SPME (**Figure 5E-F**). Both platforms performed well for the garlic flavour attribute. The predictive variables are shown in **Table 2**. As might be expected, (poly)sulfide compounds, including isocaryophyllene as well as many other non-identified compounds, were found here to be associated with garlic flavour. These allyl (poly)sulfides have already been reported as having various garlic-related aromas (e.g. roasted garlic) and have also been previously identified in garlic samples (Lee et al. 2003). This helps validate our expectations that SPME and SBSE were able to predict garlic-related volatiles from tomato soup mixtures. Moreover, the concentration of these volatiles was highly associated with the presence of one specific yeast product G28 (**Figure 6**). Thus, from this initial study we can hypothesise that the addition of G28 should therefore enhance the garlic odour of tomato soup or indeed other products. Once this is confirmed, further research is justified into determining the identification of the non-annotated compounds as they are also important contributors to the garlic flavour of the tomato soups and knowledge of their chemical structure should give us deeper insights into the potential chemical basis of this sensory attribute.

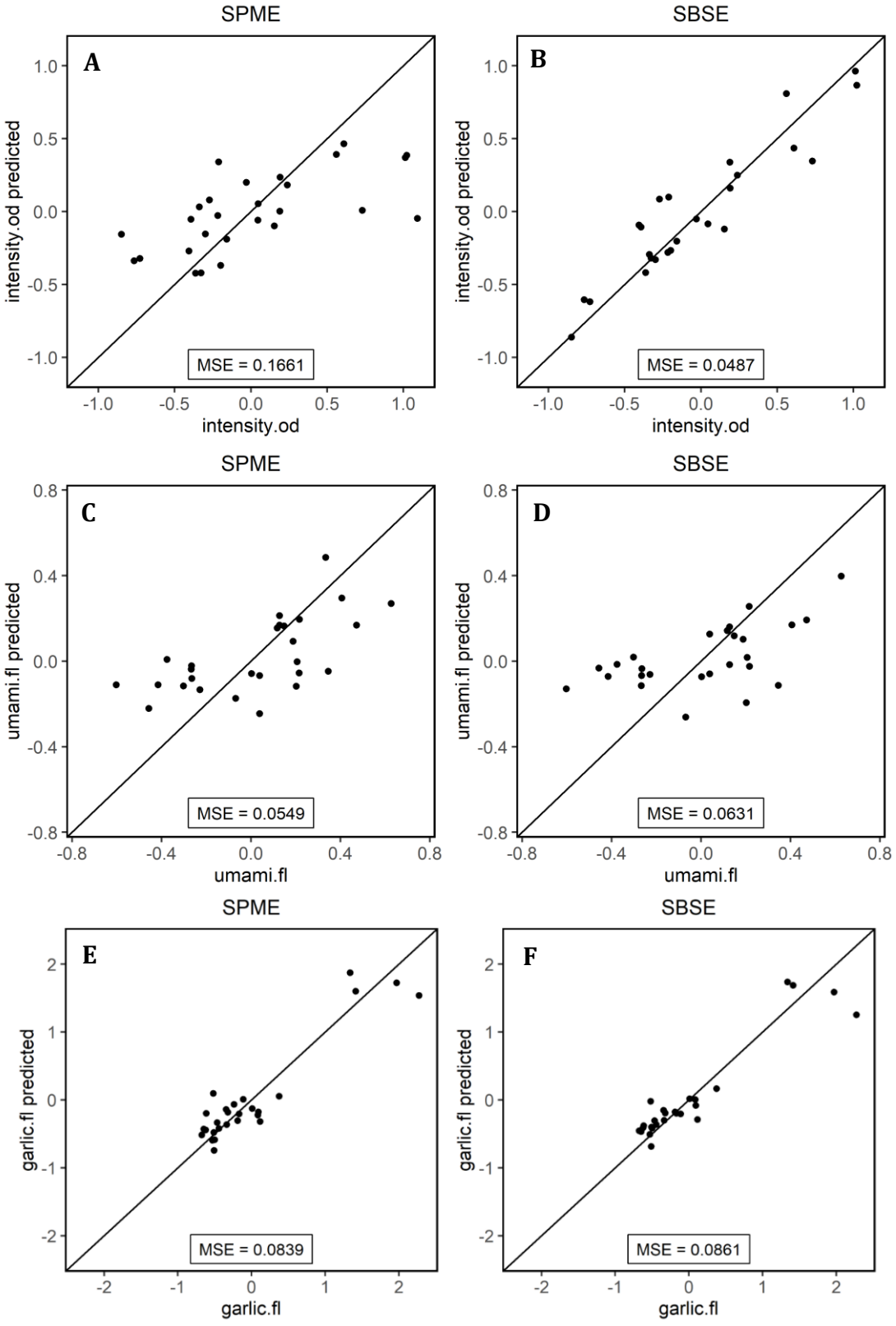


Figure 5. Performance of Elastic Net models in predicting odour intensity from SPME **(A)** and SBSE **(B)**; umami flavour from SPME **(C)** and SBSE **(D)**; and garlic flavour from SPME **(E)** and SBSE **(F)** data. The diagonal lines represent the predicted = observed line. The mean squared error (MSE) is a measure for the quality of prediction.

Table 2. Selected volatiles for the prediction of garlic flavour in tomato soups analysed by SBSE-GC-MS.

Number	Name	Chemical group	Molecular formula
1	allyl methyl disulfide	(Poly)sulfide	C4H8S2
2	allyl methyl trisulfide	(Poly)sulfide	C4H8S3
3	diallyl sulfide	(Poly)sulfide	C6H10S
4	diallyl disulfide	(Poly)sulfide	C6H10S2
5	diallyl trisulfide	(Poly)sulfide	C6H10S3
6	isocaryophyllene	Sesquiterpene	C15H24
7	non-identified		
8	non-identified		
9	non-identified		
10	non-identified		
11	non-identified		
12	non-identified		
13	non-identified		
14	non-identified		
15	non-identified		
16	non-identified		
17	non-identified		
18	non-identified		
19	non-identified		
20	non-identified		
21	non-identified		
22	non-identified		
23	non-identified		

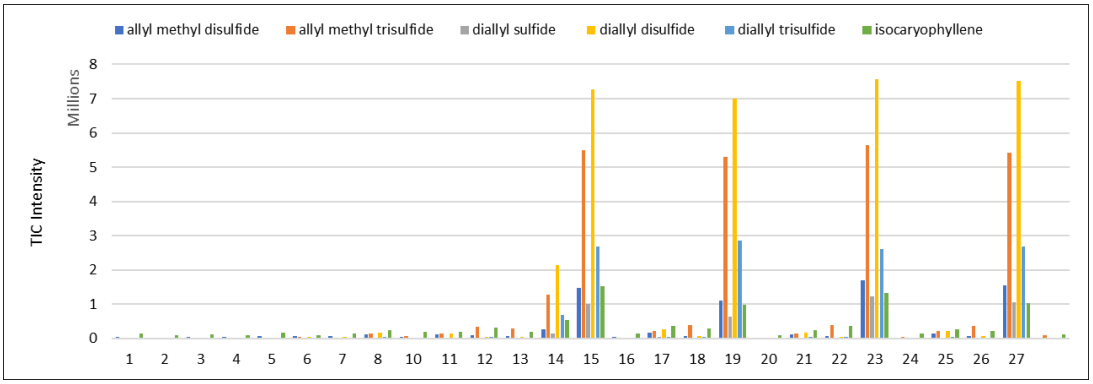


Figure 6. Bar graphs showing the TIC intensity of the selected identified volatiles for garlic flavour (**Table 2**) present in the tomato soup samples numbered as in **Table 1**. Allyl methyl disulfide (blue), allyl methyl trisulfide (orange), diallyl sulfide (grey), diallyl disulfide (yellow), diallyl trisulfide (light blue) and isocaryophyllene (green).

5. Conclusions

In this paper we have provided a methodology to decide between two analytical methods as to which one would be most useful in a metabolite-sensory relationship study. Such a decision often has to be made in an early phase of a study where only pilot data are available while decisions have to be made for a larger study.

In the larger study, new products might have to be tested which are different from the products included in the pilot study, and therefore it is not sufficient to only focus on how well the sensory attribute could be predicted from the metabolite levels using only the pilot products. The potential application of the two metabolomics methods for new products has been assessed.

We have tackled this challenge by analysing the pilot results on three levels.

> **Level 1:** How well are the two methods able to measure the metabolites they can detect? For this we looked at the precision of the methods using QC samples. We have shown this can be done in an overall manner using a metabolite independent approach, but also in a direct comparison using only the metabolites that were measured in both metabolomics methods (common metabolites).
>
> **Level 2:** How well are the two methods able to detect differences between the products? Here we assumed that the products in the pilot study were made using an experimental strategy, such that the experimental factors could be tested for differences between their levels, e.g. compare all products made using olive oil vs all products made using corn oil. It is expected that in the follow-up study, similar experimental strategies will be used, but that the specific products made using those strategies will be different. Therefore, it makes more sense to focus on how well the different strategies can be distinguished using the metabolomics methods.
>
> **Level 3:** How well are the two methods able to provide metabolite levels that are able to predict sensory attributes of the pilot products? Of course, for the larger study this is the relevant question, and we would like to choose the method best able to make this prediction. However, we have to make a decision based on a small set of pilot samples.

In the example study of the 27 tomato soups reported in this paper, the results are not always consistently pointing to the same method as being the best. SPME was able to detect metabolites with a better precision, SBSE seemed to be able to provide a better distinction between the 27 soups and was also better in predicting some of the attributes. This could be due to the larger number of metabolites measured with SBSE that were not detected using SPME, and which also seemed to be relevant for the metabolite sensory relationship in the tomato soup project.

Depending on the final goal of the larger follow-up study, a different conclusion could be obtained from the results provided. If the study aims at precise quantification e.g. when a comparison with a product of a competitor is necessary, the many metabolites with low RSD could be useful and favour SPME. When new soup products have to be developed with specific sensory properties, the SBSE method could perform better.

Concluding, there are many aspects to consider when designing a study relating sensory characteristics to metabolomics data. We have provided a framework for designing such studies that addresses three major aspects of such a study: a) the repeatability of the possible platforms to use, b) the diversity of the platforms with regard to the products to be tested and, finally, c) initial analysis to confirm whether the platform has

predictive relevance for the sensory characteristics. We recommend using this framework for future sensory studies since it provides a systematic strategy for analysing screening experiments which may be used to design a full sensory study.

7

A data-driven approach to link volatiles with key sensory attributes of chicken bouillons

Carmen Diez-Simon, Justus L. Groβmann, Johan A. Westerhuis, Age K. Smilde, Doris M. Jacobs, Rudi van Doorn, Brenda Ammerlaan, Marco A. van den Berg, Roland Mumm, Robert D. Hall

This chapter is in preparation for publication after the incorporation of non-volatile (LC-MS) data

Abstract

There is an increased demand to create new, superior sensory food experiences. In industry, yeast extracts are often used as ingredients in processed foods (bouillon, instant soups, snacks, etc.) to create new aromas and taste qualities that are appreciated by consumers. The aim of the present study was to characterize chicken bouillon samples containing contrasting yeast derived products (YP) using combined chemical and sensory approaches, which were associated by means of statistical multivariate analyses. Untargeted volatile profiles were obtained using stir bar sorptive extraction (SBSE) coupled to gas chromatography-mass spectrometry (GC-MS), and the sensory evaluation was performed on identical samples by a quantitative descriptive analysis (QDA) conducted by trained panellists. In total, 290 volatiles were detected, including 'unknowns', from chemical groups of predominantly aldehydes, esters, pyrazines, and ketones. Thirty-four sensory attributes were assessed, in the categories of odour, flavour, mouthfeel, aftertaste and after-feel. Random Forest revealed volatiles associated with roast odour (2-ethyl-5-methyl pyrazine, 2,3,5-trimethyl-6-isopentyl pyrazine and a number of 'unknowns') and chicken odour (2,4-nonadienal, 2,4-decadienal, 2-acetyl furan and a number of 'unknowns') which were mostly present in the yeast derived process flavours (PF). This study describes a straight-forward data-driven approach to improve, in an effective manner, the flavour of processed foods by design. It also highlights the limitations and preconditions for a good prediction model.

Keywords: volatiles; stir bar sorptive extraction-gas chromatography-mass spectrometry (SBSE-GC-MS); sensory analysis; roast odour; savoury processed ingredients; chicken bouillon

1. Introduction

In food studies it is often of particular interest to understand better the relationship between the perceived aroma and the chemical composition of volatile compounds in food products. Associating (groups of) volatile compounds with odour qualities is important with respect to food product design and formulation. If certain volatile compound(s) have a strong correlation with a desired odour attribute, the composition and processing procedures of a food product can be more efficiently improved through use of this knowledge. For industry, formulating new products with superior sensory qualities is of great relevance, especially when it comes to savoury ingredients. Such ingredients are widely used as constituents of many processed food products. For example, yeast-based ingredients are often added to ready-to-eat soups and meals, bouillons and snacks. For many years, researchers and food designers have been trying to relate sensory data and chemical data of ingredients in processed food to evaluate, monitor and improve food production strategies. Food processing is a complex procedure which involves many chemical reactions that lead to the formation of a vast array of molecules capable of impacting the overall flavour of a product (Jaeger et al. 2010). The most common reactions occurring during food processing are lipid oxidation and Maillard reactions (Diez-Simon et al. 2019). The latter occur between amino acids and reducing sugars, triggering the formation of many small flavour molecules including volatile and non-volatile compounds. Monitoring food processing reactions should help industry and scientists gain a more mechanistic understanding of the steps involved and allow us to select processes which contribute best to the formation of desirable flavour attributes, while importantly also avoiding the formation of off-flavours. There have been regular attempts to associate certain odour properties to specific volatile compounds (Chambers IV and Koppel 2013). For instance, wine is one of the most common food products often studied to link desired aromas with the presence of certain volatile aroma compounds. Unwanted smoke taint attributes ('ashiness', 'greenness' and 'burnt rubber') for example, were found to be associated with the presence of volatile phenols (McKay et al. 2019). However, the same study also found an increased taint perception with high levels of guaiacol, and low levels of phenols. These authors suggested that multiple molecules (volatiles) can be associated with a sensory attribute, showing that correlations are often difficult to interpret and are frequently attributed to the generally complex nature of flavour (Chambers IV and Koppel 2013).

In the study described here, we focused on the most-used savoury ingredients in industry: yeast *(Saccharomyces cerevisiae)* derived products (YP). YP used in this study belong to the categories of yeast extracts (YE) and process flavours (PF). Using YE and PF in food has many advantages, first and foremost, they can replace the need for artificial flavourings while delivering greater flavour enhancement and balance (Halasz and Lasztity 1991). During the production of YP, combinations of fermentation processes and Maillard reactions, typical of food processing, create an extended diversity of chemical compounds and flavours that makes food more palatable. However, it may also bring off-flavours (yeasty odours) that are not desired in most food applications. Recently, one study characterised aroma-active compounds using both sensory and instrumental data of yeast extracts (Wang et al. 2020). These authors analysed four contrasting YE, and concluded that p-cresol and indole were significantly linked to the 'phenolic' and

'animal' notes which caused a distinct off-flavour. Other authors reported 4-methylphenol and 3-methylpyridine, among others, as being the volatiles responsible for off-flavours in other commercial YE (Zhang et al. 2017). On the positive side, meat flavour is one of the most desired attributes for processed food and can also be obtained via specific yeast derived products. YP can positively enhance meaty aromas substituting, in this way, meat extracts or even artificial additives. This has been the topic of much research and results have shown for example, that especially for chicken and pork flavours, chemical (Maillard) reactions between cysteine and reducing sugars were correlated with sensory impact (Jayasena et al. 2013). Volatile compounds including 2-methyl-3-furanthiol, 2-furfurylthiol, methionol, 2,4,5-trimethyl-thiazole, nonanol and 2-nonenal have been proposed as being important for the specific flavour of chicken (Aliani and Farmer 2005; Jayasena et al. 2013). In addition, alkylpyrazines and trithiolanes have been associated with fried chicken and roasted chicken, but not with chicken broth (Jayasena et al. 2013). This suggests that many volatile groups (or a combination of them) can contribute to the overall sensory perception. Despite the new developments in metabolomics studies, there is still a substantial knowledge gap in relation to yeast derived food ingredients. In this study, we focus on the application of metabolomics techniques and sensory evaluations to develop prediction models for chicken bouillons supplemented with contrasting YP (yeast extracts, YE; and process flavours, PF).

Modelling of intrinsic non-linear data between sensory and instrumental measurements has been the scope of many food studies. Regression analyses, such as partial least squares analysis (PLS), are the most used statistical strategies to find associations between sensory and instrumental data (Gromski et al. 2015). However, other tools like Random Forest and Elastic net have also been used (Shi et al. 2019). In an earlier related pilot study, where tomato soups with added YP were tested, PLS-based strategies for predicting certain aroma attributes like odour intensity or specific ingredients (oil type, YP type) were evaluated (Chapter 6 of this thesis). The goal of that pilot study was to lay down a strategy to help in the selection of one of the two volatile trapping techniques in metabolomic studies using GC-MS (SPME and SBSE) that had a better performance for the predictive models used to associate sensory and metabolite (volatiles) data. The pilot study revealed that when samples have little variability in the sensory space, making models with strong prediction power is a challenge. To continue the study of the sensory impact induced by YP, a new set of food samples were created comprising a broader range of ingredients and having more contrasting sensory perceptions and chemical compositions. A simpler matrix (chicken bouillon instead of tomato soup) was selected to better expose the supplement contribution. SBSE-GC-MS untargeted analysis was selected for this diverse set of samples since it was able to extract a wider range of volatile groups, including the less volatile and less apolar compounds (Chapter 6 of this thesis).

The goal of the present study was therefore to highlight relationships between volatile compounds and key sensory attributes by applying predictive models in a data-driven efficient workflow, purposive for food formulation. To do this, we combined the outcome of a comprehensive sensory evaluation by quantitative descriptive analysis (QDA) and the untargeted volatile profiles by SBSE-GC-MS of the chicken bouillon samples prepared. Using multivariate statistical analysis (Random Forest) we aimed to correlate specific volatile groups with individual odour attributes, and set up the fundamentals for obtaining good prediction

models that can be used for future formulation studies. The present study is part of a larger study, where LC-MS analyses of non-volatile compounds (which are more related to taste attributes) will also be included and combined with the sensory and metabolite data. These results will not be described in detail as the main topic of this thesis is based on the GC-MS analysis of volatile compounds.

2. Material and methods

2.1 Food materials and soup preparation

Samples used for the sensory and chemical analysis were prepared by using a base recipe of chicken bouillon powders with reduced salt and no MSG (provided by Unilever R&D, Wageningen, The Netherlands) which were supplemented with yeast derived products, YP (provided by DSM, Delft, The Netherlands). The final composition of the bouillon samples was varied by adding different YP and different levels of a commercial chicken flavour from Unilever (dosages, high: 80% of the normal dosage; medium: 40% of the normal dosage; no: no chicken flavour). Since the different YP contain different levels of salt, salt was separately added so that the final salt concentration in each bouillon sample was 1.44 g/L. Seventy-one bouillon samples were prepared and analysed, each of which contained one of the thirty-six different YP, each with two different chicken flavour dosages (For the experimental design see **Table 1**). YP were categorized into four types: process flavour (PF), process flavour blend (PB), yeast extract (YE) and yeast extract blend (YB). The bouillon soups were prepared by stirring the bouillon powders (2.7-2.8 g/L of chicken flavour and varied amounts of YP) into 1 L of hot water (90 ˚C). After that, the majority was subjected to a trained sensory panel for sensory evaluation, while simultaneously, small aliquots (20 mL) were transferred into glass vials which were stored in a freezer (at -80 ˚C) until further chemical analysis.

Table 1. Compositional factors of the chicken bouillon samples used in this study. Chicken bouillons varied in chicken flavour level (dosages: high, medium, no) and yeast derived products (YP). YP types are divided in four: process flavour (PF), process flavour blend (PB), yeast extract (YE) and yeast extract blend (YB). The samples were assessed by panellists in different sessions, and a few samples were duplicated in the SBSE-GC-MS analysis.

Chicken bouillon samples					Sensory part	Analytical part
Number	Name	Chicken flavour dosage	Yeast derived product (YP)	Type of YP	Panel session	Technical duplicates
1	CLPFA	no	PFA	process flavour	5	
2	CMPFA	medium	PFA	process flavour	3	
3	CLPFB	no	PFB	process flavour	6	
4	CMPFB	medium	PFB	process flavour	1	
5	CHPFC	high	PFC	process flavour	1	
6	CLPFC	no	PFC	process flavour	4	
7	CLPFD	no	PFD	process flavour	6	Yes
8	CMPFD	medium	PFD	process flavour	2, 4	Yes
9	CLPFE	no	PFE	process flavour	8	
10	CLPFF	no	PFF	process flavour	8	
11	CMPFF	medium	PFF	process flavour	4	
12	CHPFG	high	PFG	process flavour	8	
13	CLPFG	no	PFG	process flavour	2, 4	
14	CHPB1	high	PB1	process flavour blend	8	
15	CMPB1	medium	PB1	process flavour blend	7	
16	CHPB2	high	PB2	process flavour blend	2	
17	CMPB2	medium	PB2	process flavour blend	6	
18	CHPB3	high	PB3	process flavour blend	1	
19	CLPB3	no	PB3	process flavour blend	3, 5	
20	CHPB4	high	PB4	process flavour blend	8	Yes
21	CMPB4	medium	PB4	process flavour blend	1, 7	Yes
22	CHPB5	high	PB5	process flavour blend	4	
23	CMPB5	medium	PB5	process flavour blend	5	
24	CMPB6	medium	PB6	process flavour blend	6	
25	CLPB6	no	PB6	process flavour blend	3	
26	CHPB7	high	PB7	process flavour blend	5	
27	CLPB7	no	PB7	process flavour blend	1	
28	CMPB8	medium	PB8	process flavour blend	2	
29	CLPB8	no	PB8	process flavour blend	4	
30	CHPB9	high	PB9	process flavour blend	5	
31	CMPB9	medium	PB9	process flavour blend	8	
32	CHPB10	high	PB10	process flavour blend	3	
33	CMPB10	medium	PB10	process flavour blend	1	
34	CHPB11	high	PB11	process flavour blend	7	
35	CLPPB11	no	PB11	process flavour blend	3	
36	CMPB12	medium	PB12	process flavour blend	3	
37	CLPB12	no	PB12	process flavour blend	1	
38	CHPB13	high	PB13	process flavour blend	3	
39	CMPB13	medium	PB13	process flavour blend	4	
40	CMPB14	medium	PB14	process flavour blend	6	
41	CLPB14	no	PB14	process flavour blend	2	
42	CHYEA	high	YEA	yeast extract	6, 7	
43	CLYEA	no	YEA	yeast extract	2	
44	CHYEB	high	YEB	yeast extract	1	

45	CLYEB	no	YEB	yeast extract	7	
46	CHYEC	high	YEC	yeast extract	6	
47	CMYEC	medium	YEC	yeast extract	5	
48	CHYED	high	YED	yeast extract	2	
49	CMYED	medium	YED	yeast extract	3,5	
50	CMYEE	medium	YEE	yeast extract	8	
51	CLMYEE	no	YEE	yeast extract	1	
52	CMYEF	medium	YEF	yeast extract	5	
53	CLYEF	no	YEF	yeast extract	1, 8	
54	CHYEG	high	YEG	yeast extract	3, 5	Yes
55	CLYEG	no	YEG	yeast extract	4	Yes
56	CMYEH	medium	YEH	yeast extract	8	
57	CLYEH	no	YEH	yeast extract	6	
58	CHYEI	high	YEI	yeast extract	6	
59	CLYEI	no	YEI	yeast extract	7	
60	CHYB1	high	YB1	yeast extract blend	6, 8	Yes
61	CMYB1	medium	YB1	yeast extract blend	2	Yes
62	CHYB2	high	YB2	yeast extract blend	2	
63	CMYB2	medium	YB2	yeast extract blend	7	
64	CHYB3	high	YB3	yeast extract blend	4	
65	CLYB3	no	YB3	yeast extract blend	3	
66	CHYB4	high	YB4	yeast extract blend	4	
67	CLYB4	no	YB4	yeast extract blend	7	
68	CMYB5	medium	YB5	yeast extract blend	7	
69	CLYB5	no	YB5	yeast extract blend	5	
70	CHYB6	high	YB6	yeast extract blend	3	
71	CMYB6	medium	YB6	yeast extract blend	2	

2.2 Sensory evaluation

Sensory evaluations were performed at the sensory facilities of DSM Biotechnology Center (DBC, Delft, The Netherlands) using established standard protocols. An extended sensory evaluation of the samples was performed by applying Quantitative Descriptive Analysis (QDA). QDA investigates the odour (aroma), flavour, mouthfeel, aftertaste and after-feel attributes, such as odour intensity, yeast odour, umami flavour, chicken flavour, fatty mouthfeel, salt aftertaste and mouthcoating after-feel of a set of prepared chicken bouillon soups (See **Table 2** for an overview of the sensory attributes). In total, thirty-four attributes were assessed by a group of experienced, trained (>10 years) panellists (n=14). The sensory evaluation was divided into eight sessions. For each session, ten samples were tested by each panellist. All bouillon samples were kept in a holding cabinet at 60˚C and 50 mL aliquots of chicken soup were served one-by-one to the panellists in white polystyrene cups. During the QDA measurement, the intensities of the attributes were obtained using EyeQuestion (Logic8), using unstructured line scales ranging from 0-100. Eight bouillon soups were also tested across different sessions to check for "between-session" variation. In total, seventy-nine bouillon soups were tested by the sensory panel. The number of panellists per session varied from 10-14, as during the whole series of sensory evaluations panellists can turn ill or otherwise be prohibited from attending a particular session.

Once the sensory data were obtained, standardization of the data was performed to correct for different use of scales by the different panellists in terms of their level effect and scaling effect, following the approach

described by Romano et al. (2008). This standardization corrects for panellists who, on average, gave higher or lower intensity values for all products and/or used different ranges for the rating scale than the average panellists (e.g. some used 40 to 55 while others, 20 to 70). F-statistics were also applied to compare the different models. This is the same standardization procedure as used in the pilot study on tomato soups with added YP (Chapter 6 of this thesis). Pearson correlation matrix for the sensory data, as well as the combined heatmap clustering matrix, were performed using Rstudio (R 4.0). For the sensory heatmap, data was re-scaled to an interval of [0, 1], and clustered using the Ward criterion (R 4.0).

Table 2. Overview of the sensory attributes assessed by QDA in this study.

Sensory attributes					
Odour (od)	1	intensity-od	**Mouthfeel (mf)**	23	fullness-mf
	2	fatty-od		24	pungent-mf
	3	yeast-od		25	fatty-mf
	4	roast-od		26	astringent-mf
	5	musty-od	**After-taste (at)**	27	intensity-at
	6	chicken-od		28	sweet-at
	7	sulphury-od		29	salt-at
	8	off odor-od		30	bitter-at
Flavour (fl)	9	intensity-fl		31	sour-at
	10	sweet-fl		32	length-at
	11	sour-fl	**After-feel (af)**	33	mouthcoating-af
	12	bitter-fl		34	astringent-af
	13	salt-fl			
	14	umami-fl			
	15	yeast-fl			
	16	roast-fl			
	17	chicken-fl			
	18	herbs-fl			
	19	sulphury-fl			
	20	complex-fl			
	21	balance-fl			
	22	off flavor-fl			

2.3 Untargeted volatile analysis

For the determination of the volatile levels, stir bar sorptive extraction (SBSE) was used to trap the volatiles from the liquid matrix to a polymer-based coating stir bar (polydimethylsiloxane, PDMS). Analytes trapped by the stir bar were thermally desorbed (TDU, Gerstel, Mülheim, Germany) and analysed using GC-MS (Agilent Technologies, Ratingen, Germany). Extraction and analysis followed the same procedure as described before (Diez-Simon et al. 2020a), with slight modifications. In brief, bouillon soups were defrosted, sonicated for 10 min, and a 9 mL aliquot was pipetted into a 10 mL screw-cap vial. Immediately, a magnetic stir bar was immersed into the liquid sample. The vial was then incubated at 60˚C for 10 min and stirred at room temperature for another 80 minutes. After that, the stir bar was removed, rinsed with water, dried and placed into a glass desorption tube. Thermal desorption of the analytes onto the column was performed using a similar method as described before (Diez-Simon et al. 2020a), also the same GC-MS

conditions were used. The two deviations from the previous study were that samples were desorbed in splitless mode during the first two minutes of the CIS parameters, and the temperature used to desorb the analytes into the GC, as well as from the GC column to the MS was set to 270˚C (instead of 250˚C previously).

The analysis sequence was as follows: a set of 80 bouillon samples was analyzed in a randomized way. Duplicates of twelve of the samples were also prepared to check for repeatability. A quality control (QC) sample, which was a mix of all soups ('pooled' sample), was also repeatedly analyzed along the sequence after every ten samples to test the performance of the method. An empty glass tube, a clean stir bar, and a blank stir bar (water) were also measured at the beginning of the sequence to check for background peaks. An n-alkane series (C_8-C_{22}) was analyzed to calculate retention indices (RI). In total, 114 samples were analyzed in a single series, but the bouillons were prepared and extracted in a set of eight batches of 15 samples each before placing each batch in the GC rack for analysis. This was done to avoid long 'storage' times on the autosampler. The whole analysis sequence took around 5 days.

The raw GC-MS data were processed using an untargeted metabolomics approach as detailed before (Diez-Simon et al. 2020a). In brief, raw data was baseline-corrected and mass peaks of samples were aligned using MetAlign software and mass spectra were reconstructed using MSClust. Volatiles coming from the instrument (siloxanes) and blanks were removed. Metabolites were identified by matching the mass spectra and RI to authentic reference standards or those in the NIST17 Mass Spectral library (v.2.3), following the MSI criteria for metabolite identification as proposed by Sumner et al. (2007). Compounds that did not fit the criteria were annotated as being non-identified. Heatmaps combined with clustering analysis were used to visualize the chemical data. For the metabolite heatmap, data was transformed (square root), re-scaled to an interval of [0, 1], and clustered using the Ward criterion (R 4.0).

2.4 Prediction models

To explore the relationships between sensory attributes and chemical composition of the bouillon samples, the MUVR R package (Shi et al. 2019) was used with a random forest regressor (R 4.0). The random forest regression method is a machine learning algorithm that uses an ensemble of small regression trees, which are all based on a small subset of volatiles and samples in the data set, to predict a phenotype (Hastie et al. 2001). The final prediction is based on an average of a large number of trees. This makes the prediction robust against overfitting. Furthermore, metabolites have to show their importance in multiple trees before being selected as important.

MUVR uses a double cross-validation approach (Filzmoser et al. 2009) to obtain an unbiased estimate of prediction performance and an unbiased set of selected variables. In addition to a *minimal-optimal* set of variables, MUVR also provides an *all-relevant* variable set to enable a more complete view on which compounds are relevant for the respective sensory attribute. However, here only the *minimal-optimal* set was considered. For identification of relevant variables, MUVR random forest regression models were fitted (14 outer cross-validation folds, 100 repetitions) for each sensory attribute using the GC-MS data as

predictors (biological and technical replicates aggregated into means). Q^2 values, also known as cross-validated R^2 values (R^2_{CV}), were calculated as a measure for predictive performance of each model.

3. Results and Discussion

3.1 Sensory data

Sensory descriptive analyses were carried out to obtain a broad sensory profile of chicken bouillon samples, including attributes related to odour, taste, aftertaste, after-feel and mouthfeel. In total, thirty-four attributes were assessed by trained panellists based on an unstructured line scale from 0 to 100. The score of the panellists were averaged by applying standardization methods that are commonly used for this purpose to obtain a mean value that represents the difference between the bouillon products (Romano et al. 2008). In this study, sensory data were standardized as explained in Section 2.2 and the mean values are represented in **Figure 1A**. Large differences exist between individual panellists (coloured lines), not all assessed all products and the scale range used was different for each panellist. After standardization, the mean value was centred and the standard error between panellists decreased considerably (**Figure 1B**), which clearly increased the F-value of the ANOVA indicating that the products could be better discriminated from each other after standardizing the data (**Figure 1C**).

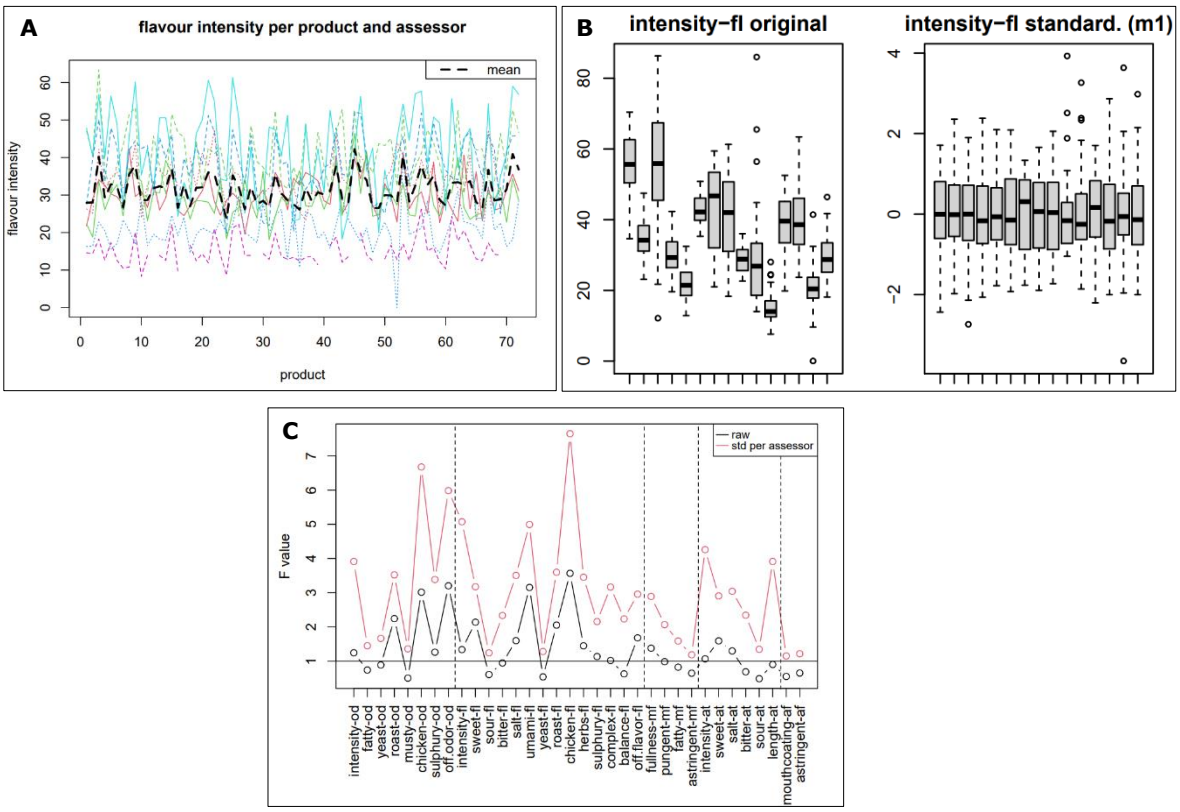


Figure 1. (A) Scores for the sensory attribute 'intensity flavour' per bouillon sample given by each panellist (coloured lines) and the calculated mean value (discontinuous black line). **(B)** Scores for 'intensity flavour' per assessor (x-axis) before (left) and after (right) applying standardization method. Standard deviations decrease and mean values become equal between assessors. **(C)** F-values of the ANOVA per sensory attribute before (black) and after (red) applying standardization method (Section 2.2).

A Pearson correlation matrix representing the transformed sensory mean values of the bouillon samples is shown in **Figure 2A**. As can be observed, certain attributes showed a correlation with others. In general, odours, flavours and aftertaste from the same attribute are positively correlated: for instance, salt-fl with salt-at (0.8-0.9) and roast-fl with roast-od (0.6-0.7). The same can be observed between different attribute classes and clear groupings can be noticed. Attributes such as balance, umami, salt, sweet, intensity, length, complexity and fullness are highly correlated (0.3-0.8), meaning that some bouillon samples scored high in all the seven sensory attributes. Interestingly, studies have shown that umami enhances saltiness sensations (Onuma et al. 2018), as well as sweetness as they share a common receptor subunit (Li et al. 2002; Shim et al. 2015). This can be observed here, as umami shows particularly its highest correlation with salt (0.7-0.8) as well as with sweet (0.6-0.7). In addition, a correlation between chicken and fatty odour was also observed (**Figure 2A,** 0.4-0.6), which is interesting when looking for highly desired chicken aroma and flavour characteristics for bouillons. There is also a weaker, but still positive, correlation visible between chicken and sulphury (0.3-0.5). The same can be seen for roast and bitter attributes (and for yeast, to a lesser extent), which also appear to correlate in the data (0.2-0.4). Many flavour researchers have already noticed this correlation between sensory attributes on taste enhancers, such as roast and bitter in coffee (Alstrup et al. 2020). Bitter, in this study, appeared to negatively correlate with chicken odour which can also be considered as undesired and desired attributes for chicken bouillons, respectively.

To visualize the complete sensory scores among the different chicken bouillon samples, a combined heatmap and hierarchical clustering is shown in **Figure 2B**. It can be observed that bouillon samples scored differently and sparsely. Two main sensory clusters are observed (Groups I and II, **Figure 2B**), where attributes driven by bitter, sour, herbal and roast notes (or an analogy) were clustered separately from the attributes driven by chicken, sweet and umami notes (group I and II, respectively). This indicates that a group of bouillon samples were characterized by the sensory attributes in group I, while a second group of bouillon samples had higher scores for the group II attributes. Moreover, a small group of chicken bouillon samples (block 1, **Figure 2B**) were perceived as having the highest roast and bitter flavour than the other products. Those samples contained the yeast derived products PB6 and PB12. High chicken odour and chicken flavour are also attributes of interest for this study, and the samples that scored highest of these attributes (arrowed, **Figure 2B**) contained the yeast derived product PB3 (**Table 1**). Hence, the addition of a certain combination of ingredients induced changes to the sensory characteristics. Here, the addition of YP from the category of process flavour blends (PB6, PB12 and PB3, **Table 1**) seemed to enhance the roasted and chicken sensations in the bouillons. Similarly, others have reported that the addition of the amino acid lysine and yeast extract increased the sensory perception (and acceptance) of low-salt salted meat and decreased the rancid aroma (Vidal et al. 2020). In the present study, combinations of differing chicken flavour levels and various YP enhanced the perception of contrasting sensory attributes. As a broad sensory space appears to be covered, the metabolite data will be further analysed in the following section.

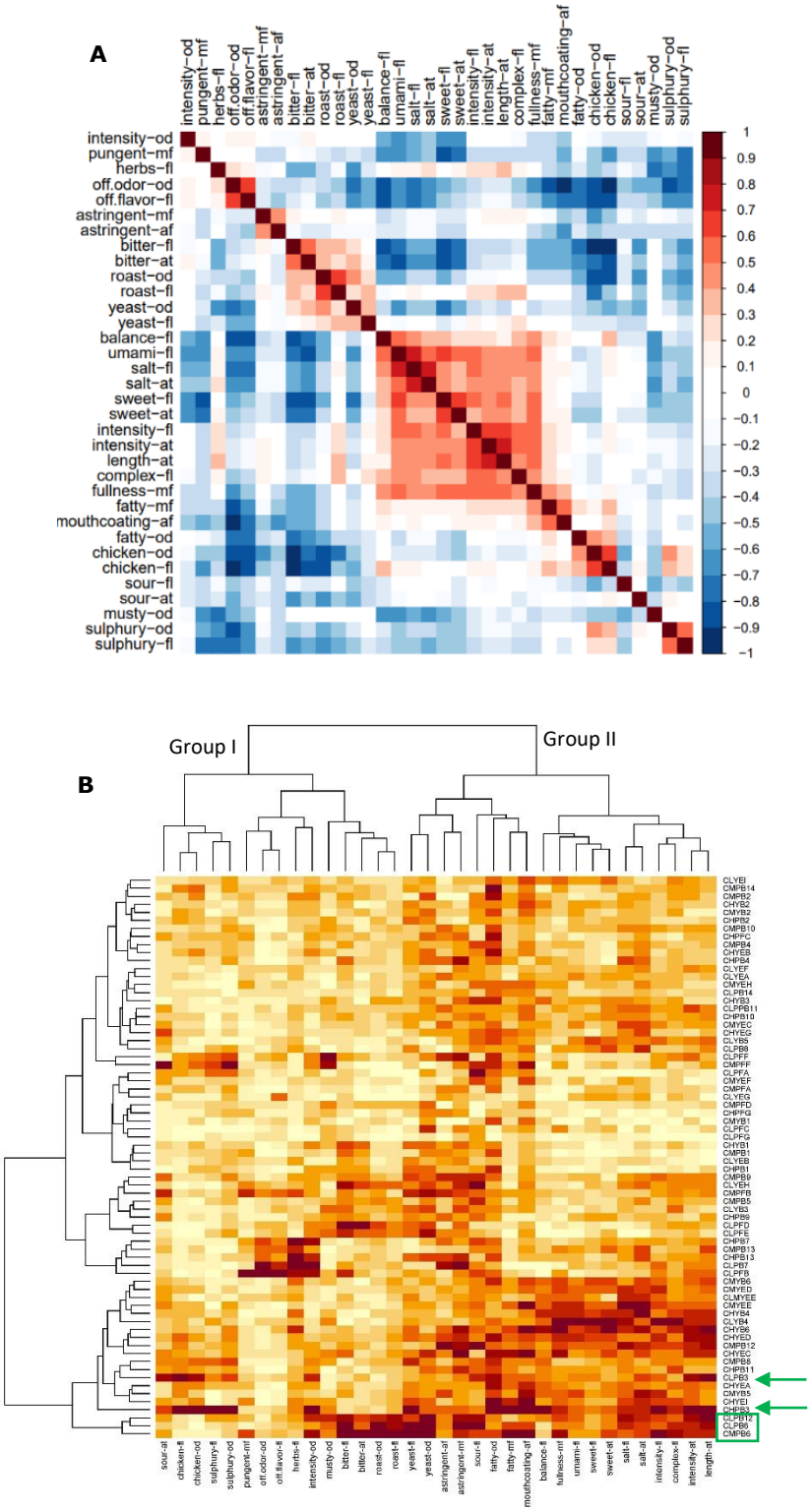


Figure 2. (A) Pearson correlation matrix of the sensory chicken bouillon data showing the correlation values of the different sensory attributes (high correlation: dark red, no correlation: white, negative correlation: dark blue). **(B)** A combined heatmap and hierarchical clustering analysis of the sensory data and chicken bouillon samples representing the high sensory scores in dark red, and low sensory scores in light yellow. Sensory attributes are clustered into two main groups (I and II). Box 1 and arrows highlight the samples scoring higher in roasted (box 1) and chicken flavour (arrows).

3.2 Metabolomics analysis of volatile compounds

Aliquots of the same samples that were used for the sensory evaluation, were immediately stored (-80˚C) and used later for volatile analyses. These were carried out using an untargeted SBSE-GC-MS approach and this resulted in a total of 290 volatile compounds being detected (after filtering instrumental and blank peaks). Multivariate heatmap and clustering analysis showed the distribution of the chicken bouillon samples according to their volatile profiles (**Figure 3**). The technical variation of the majority of volatile compounds in the QC samples was less than 20% (**Figure 4A**). However, compounds having higher RSD values were characterised by having low abundance (**Figure 4B**). This repeatability was thus in line with the previous pilot study using tomato soups containing the same or similar YP (Chapter 6 of this thesis). Of the 290 detected volatile compounds, 152 could be putatively identified (**Table 3**). The most common classes of volatiles identified were aldehydes (31), esters (18), pyrazines (18), ketones (15), followed by polysulfides (11), alcohols (11) and sesquiterpenoids (11), furan(one)s (9), monoterpenoids (8), and fatty acids (3). As can be seen in **Figure 3**, certain volatiles appear to be present in almost all bouillons, while others appear to be uniquely present or present in a sub-group of bouillon samples.

Chicken bouillons containing the process flavour PFB showed the most distinct volatile profiles compared to the other samples. This sample had a remarkably different volatile profile than the rest. This sample, with either medium or no chicken flavour, contained either unique and/or higher levels of particularly polysulfides and sesquiterpenes (box 1, **Figure 3**). The same volatiles were previously observed for this PF when analysed before in a different study (PF S99, Diez-Simon et al. (2020a)). Sesquiterpenes such as beta-caryophyllene and alpha-curcumene were found in PFB along with alkyl disulfides and trisulfides. This implies that a different production process and/or different ingredients were used to formulate PFB resulting in the formation of these specific volatiles. Thus, this bouillon sample contained components that were not detected (or were lower) in all other samples. Moreover, in the sensory panel, the bouillon samples containing PFB scored high in intensity and herb flavour, but also in off-odour, off-flavour and pungent mouthfeel. Off-flavours and off-odours are undesired attributes which could be related with the higher abundance of the compounds in box 1. The high abundance of sesquiterpenes could also be associated with the herbal flavour perceived by the panellists. A certain group of alkylpyrazines were also found to be more abundant in another sub-group of samples, which were also highly differentiated from the rest by the clustering analysis (box 2, **Figure 3**). This sub-group of samples comprised process flavours and blends named PFD, PFE, PB6 and PB14 (**Table 1**). This implies that the production of these specific process flavours triggered the formation of the alkylpyrazines.

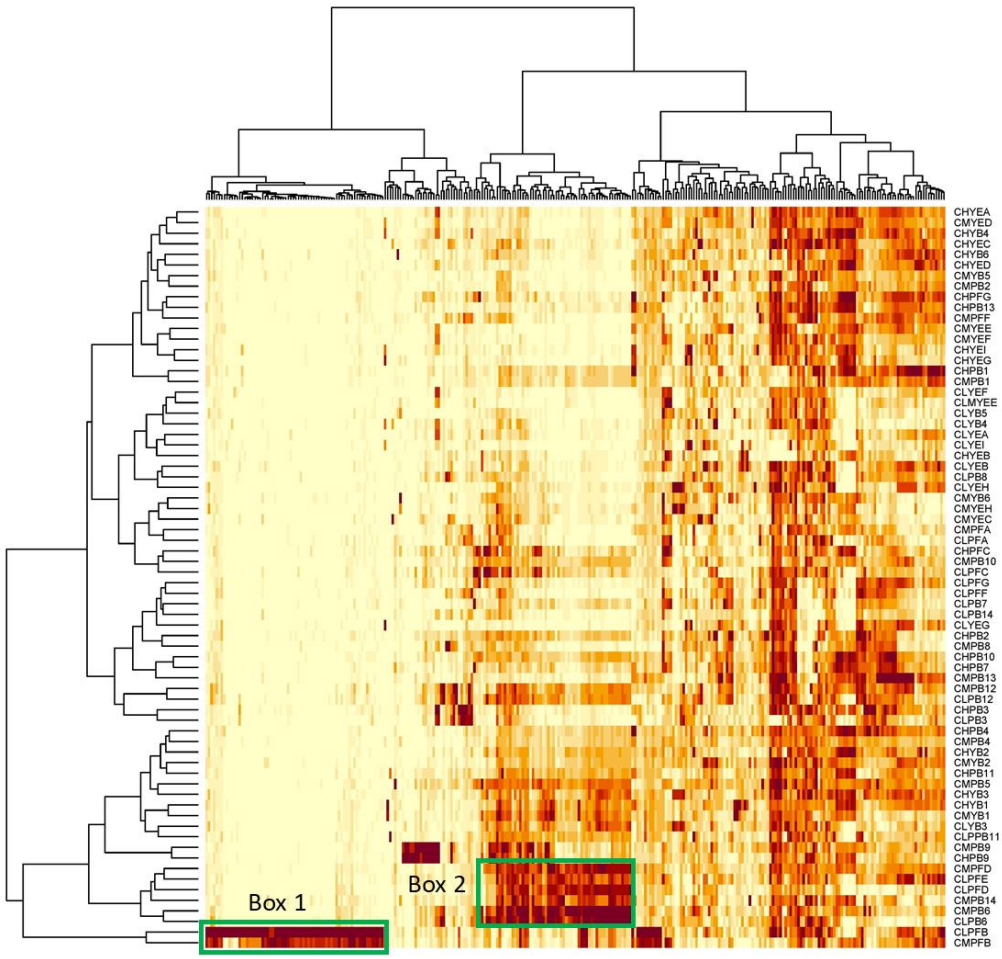


Figure 3. Heatmap based on a hierarchical clustering analysis of 290 volatile compounds detected in 71 chicken bouillon samples with added yeast derived products (YP). Samples are represented in the vertical axis while volatiles are visualized in the horizontal axis. Box 1 includes the volatiles highly (or uniquely) present in the chicken bouillon containing the PFB. Box 2 highlights a group of volatiles being characteristic of a subset of chicken bouillons. The high abundance is shown in dark red, while low abundance (or absence) is shown in light yellow.

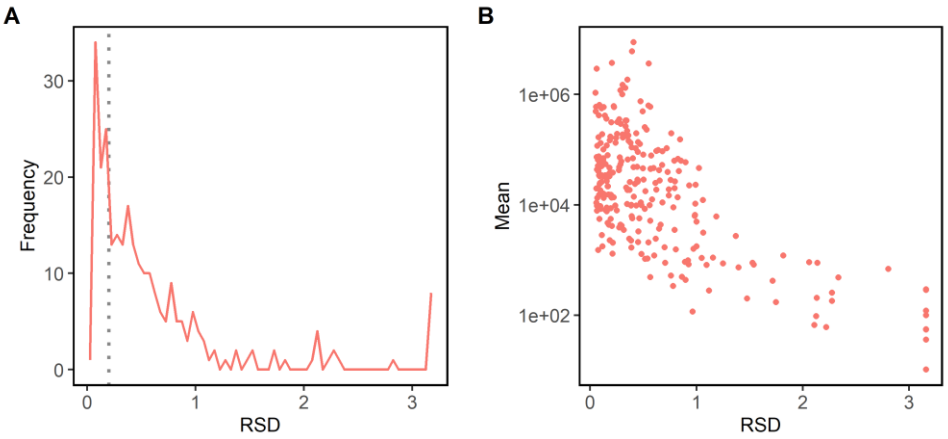


Figure 4. (A) Relative standard deviation (RSD) of the quality control (QC) samples for all 290 volatile compounds, expressed in frequency of occurrence and **(B)** in the abundance of each compound.

Table 3. List of volatile compounds analyzed by SBSE-GC-MS which were present in seventy-one chicken bouillon samples. Identified and non-identified volatiles following the criteria in Section 2.3 are shown, as well as the volatile classes found.

Total nr of metabolites	290
Nr of identified metabolites	152
Nr of aldehydes	31
Nr of esters	18
Nr of pyrazines	18
Nr of ketones	15
Nr of (poly)sulfides	11
Nr of alcohols	11
Nr of sesquiterpenoids	11
Nr of furan(one)s	9
Nr of monoterpenoids	8
Nr of fatty acids	3
Others (phenols, pyridine deriv. pyrroles, thiophenes, etc.)	17

Other volatiles, such as 2,4-decadienal, was present in the chicken bouillon samples containing medium to high levels of chicken flavour (Supplementary **Figure S1**), meaning that these compounds are likely to have come from the common ingredient (the chicken flavour mix). In fact, 2,4-decadienal, which is also known to be related to chicken aroma (Jayasena et al. 2013; Nishimura et al. 2016), was not detected in the samples without chicken flavour, although very low abundance can be seen for certain samples. This suggests that the YP could also form 2,4-decadienal, nonetheless, the abundance is so low that it is difficult to interpret, as it might also be contamination. In addition, the varied intensities among the samples containing high chicken flavour (green) indicate that the matrix effect may also play a role, as this compound appears only to come from the chicken flavour and not the YP. This aspect will be discussed later as this complexity may impact the modelling.

In conclusion, the volatile composition of the chicken bouillon samples resulted in a large diversity of molecules being detected. Many volatiles were only detected in a few products reflecting also the large diversity of different YP used to create the chicken bouillons. The relatively large distance in the volatile profiles between the differently composed chicken bouillons suggests that the samples covered a large area in the metabolic space. This makes, on the other hand, the search for molecular interactions which influence the sensory properties of the samples a challenging one because associations in the models are less strong as these will not explain a large proportion of the variance found in the data. Nevertheless, this does give us insights into the fundamentals of predictive models that will help improve the design and outcome of future experiments. To obtain a strong prediction model, both chemical and sensory space need to be well covered

by the samples. Furthermore, to be able to find molecular interactions relevant for sensory properties, it is important that these molecules are present in multiple samples. In other words, while data need to be variable, they should still also share enough common variables (volatiles). This will be further discussed in the following section.

3.3 The relationship between sensory and volatile profiles

The advanced development of metabolomics technologies and chemometric tools that are used to relate metabolites with flavour characteristics have expanded our knowledge of processing of food ingredients (Choi et al. 2017; De Toffoli et al. 2019; Kamani et al. 2019). Here, the sensory and chemical data obtained from a large set of bouillons with added YP have been shown to cover sufficient variability in order to permit an attempt to predict certain sensory characteristics from the detected volatile patterns. The prediction models applied here are based on multivariate regression models that have proven valuable in revealing reliable correlations between molecules and sensory attributes in several food products (Seisonen et al. 2016). This bouillon study was linked to a previous pilot study where PLS and Elastic net models were generated to examine the prediction power of the outcome data. However, here, it was decided to generate Random Forest models. The reason for this is related to the nature of the present data. Despite having similar food matrices in both pilot and the present studies (tomato with added YP and chicken bouillon with added YP), the chemical data obtained using bouillons resulted in more highly contrasting profiles. This was due to the variety of YP used being much larger (36 as compared to 4). Consequently, due to the metabolite variation in the samples, certain volatiles were found to be uniquely present in a small number of samples and only relatively few were found to be common. As a result, the PLS models were unable to explain most of the variance in the data, thus resulting in weak performance of the models. Therefore, Random Forest was selected instead to generate models that should better predict the chemical and sensory data of chicken bouillons. Random Forest does not require transformation of data and also avoids problems with sparse data (Gromski et al. 2015). **Figure 5A** shows an overview of the calculated Q^2 values as measure on how well all 34 sensory attributes could be predicted from the profiles of the 290 detected volatiles. As can be observed, some attributes could be well predicted including roast-od, chicken-od, sulphury-od off-odour, salt-fl, umami-fl, roast-fl, chicken-fl, herbs-fl, sulphury-fl and balance-fl (arrowed, **Figure 5A**) by the volatile profiles while others could not, such as yeast-od, musty-od, astringent-mf or astringent-af (**Figure 5A**). For example, for roast-od, a minimum of 8 variables (volatiles), which are required for good predictions, were needed to obtain a Q^2 value of 0.4 (**Figure 5B**). These selected variables are shown in **Table 4A**. Among the 8 selected volatiles being important to predict roast-od are 2 alkylpyrazines (2-ethyl-5-methyl pyrazine and 2,3,5-trimethyl-6-isopentyl pyrazine), 2 pyrazine types of volatile, and 4 non-identified compounds. 2-Ethyl-5-methyl pyrazine is known to be a sweet, bean and coffee tasting compound and has been found in a number of diverse food items which are often consumed after roasting including cereals and cereal products, coffee and tea (www.foodb.ca). Isomers of the above-mentioned pyrazines were also found in the volatile composition of fenugreek (Mebazaa et al. 2009). These and other pyrazines are compounds formed by Maillard reactions and Strecker degradation (Diez-Simon et al. 2019). Some can also be formed by heating hydroxyamino compounds such as threonine, serine, ethanolamine, and glucosamine (Shahidi et al.

1986). Alkylpyrazines have also been previously detected in roasted sunflower seeds (Guo et al. 2019), where the abundance of alkylpyrazines increased with an increase in roasting temperature and time. These pyrazines contributed most to a roast and nutty flavour of roasted sunflower seeds, but also in many other oil seed products: sesame oil, peanut oil, perilla seed oil (Cuicui and Lixia 2018). 2,3,5-Trimethyl-6-isopentyl pyrazine is not a common alkylpyrazine, but has been reported as being formed in a model system from the reaction between L-ascorbic acid and L-leucine at a basic pH (Yu et al. 2011). These studies matched with the presence of the pyrazine volatiles being selected in the roast odour model. Alkylpyrazines have also been associated with fried chicken and roasted chicken, but not with chicken broth (Jayasena et al. 2013), which gives these bouillon samples a higher complexity than that expected from a chicken broth.

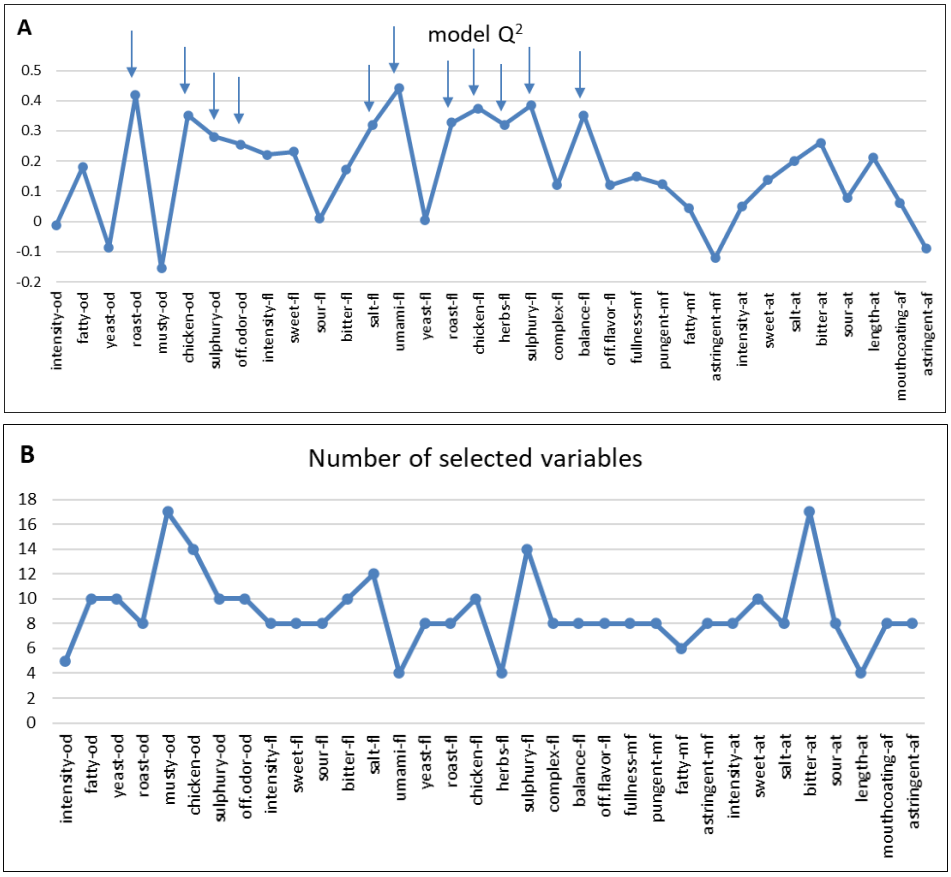


Figure 5. (A) Q^2 values of the predictive model (Random Forest) for each sensory attribute. Arrowed attributes are highlighted for their higher performance of the model. **(B)** Minimal optimal number set of selected variables from the model, for each sensory attribute, which are relevant for the sensory attributes.

Table 4A. Selected variables for roast odour, including identified and non-identified.

Number	Name	Chemical group	Molecular formula
1	**2-ethyl-5-methyl pyrazine**	Pyrazine	C7H10N2
2	**2,3,5-trimethyl-6-isopentyl pyrazine**	Pyrazine	C12H20N2
3	non-identified	Pyrazine	-
4	non-identified	Pyrazine	-
5	non-identified	-	-
6	non-identified	-	-
7	non-identified	-	-
8	non-identified	-	-

Table 4B. Selected variables for chicken odour, including identified and non-identified.

Number	Name	Chemical group	Molecular formula
1	**2-octenal**	Aldehyde	C8H14O
2	**2,4-nonadienal**	Aldehyde	C9H14O
3	**2,4-decadienal**	Aldehyde	C10H16O
4	**1-octen-3-ol**	Alcohol	C8H16O
5	**methyl propyl trisulfide**	(Poly)sulfide	C4H10S3
6	**dipropyl trisulfide**	(Poly)sulfide	C6H14S3
7	**4-hexen-3-one**	Ketone	C6H10O
8	**3,4-dimethyl thiophene**	Thiophene	C6H8S
9	**2-acetyl furan**	Furan(one)	C6H6O2
10	non-identified	-	-
11	non-identified	-	-
12	non-identified	-	-
13	non-identified	-	-

In order to relate the occurrence of the roast odour attribute to the different chicken bouillon samples, histograms were generated to distinguish which samples contained higher amounts of the selected compounds (**Figure 6**). Those samples containing higher amounts of the 2 pyrazines for roast odour are the process flavours PFB, PFD and PFE, and the process flavour blends PB5, PB6 and PB14 (**Figure 6A**). At reduced amounts, yet present are the pyrazines in the yeast extract blends YB1, YB2 and YB3 and some other process flavour blends. The addition of YP from the category of process flavours (most abundantly in PB6, PFB, PFD, and PB14) to chicken bouillon may therefore enhance the overall roasted odour of the product. Additionally, in the bouillon samples containing the process flavours PB6, PFD and PB12, roasted odour was highly scored by the panellists (**Figure 2B**), however, not for PFB and PB14. Likewise interestingly, the unknown compounds number 3 and 4 (pyrazine-related) had a similar distribution (**Figure 6B**) and were predominantly present in the process flavours/blends mentioned above. This supports the idea that the formation of these unknown compounds may be similar to the other alkylpyrazines, and may have related chemical structures. Their intensities are similar in the bouillon samples with different chicken flavour levels as compared to the identified pyrazines. This indicated that these compounds are originating from the added YP and not so much from the chicken flavour.

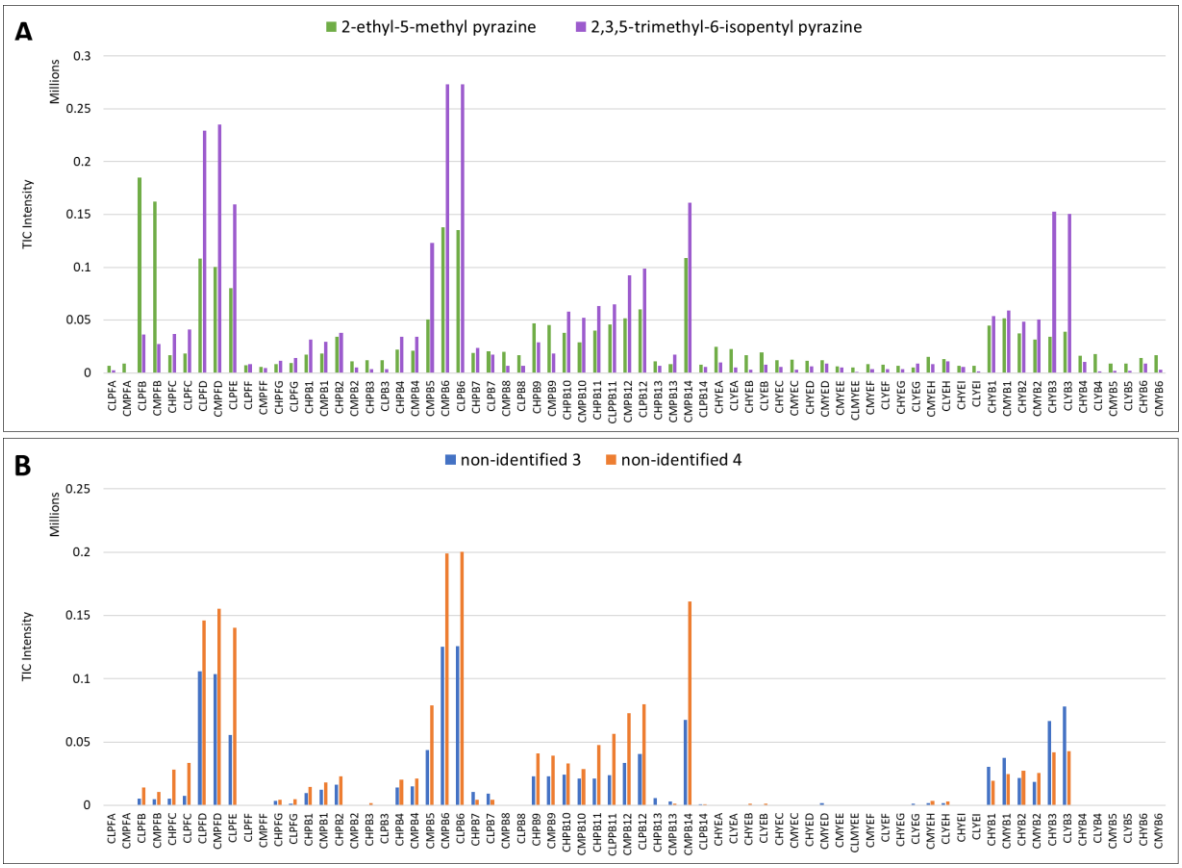


Figure 6. Bar graphs showing the TIC intensity of the selected volatiles for roast odour present in the chicken bouillon samples as ordered in **Table 1**. **(A)** 2-Ethyl-5-methyl pyrazine (green) and 2,3,5-trimethyl-6-isopentyl pyrazine (purple). **(B)** Non-identified compounds belonging to the pyrazine class from **Table 4**, number 3 (blue) and 4 (orange).

Other sensory attributes which have been modelled could also be well correlated to certain volatile components. For instance, chicken odour could well be predicted by a number of volatile compounds by the Random Forest models. The volatiles selected by the model belong to several chemical classes: aldehydes, alcohols, (poly)sulfides, ketones, thiophenes and furan(one) compounds (**Table 4B**). The aldehydes 2,4-nonadienal and 2,4-decadienal have been reported as important contributors to chicken flavour (Jayasena et al. 2013). Furanthiols have been associated with chicken flavour, however this was not the case here in the bouillon samples (**Table 4B**), instead 2-acetyl furan was detected. 2-Acetyl furan has been characterized in various foods such as orange bell pepper, brassicas and fruits as having odour properties like almond, balsamic, beef, caramel and cocoa (www.foodb.ca). The absence of furanthiols in the present data suggests that the SBSE-GC-MS technique could not detect these specific volatiles and/or that the chicken bouillon samples made with YP did not contain these meaty compounds. Although, chicken and sulphurous attributes are related (Jayasena et al. 2013) (**Figure 2A**), the selected trisulfides (**Table 4B**) are described to give cooked-like, cabbage-like aromas, but they have also been associated with chicken aroma (Aliani and Farmer 2005; Feng et al. 2018; Nishimura et al. 2016; Sun et al. 2014).

To corroborate the enhancement of chicken odour when the abundance of these compounds is higher in the samples, the distribution of these volatiles in the chicken bouillon samples was examined. Data for 2,4-

Decadienal can be seen in Supplementary **Figure S1**, and four additional examples of selected volatiles are shown in Supplementary **Figure S2**. Interestingly, samples containing a medium or high dosage of chicken flavour mix contained higher abundances of the aldehydes 2,4-decadienal and 2,4-nonadienal and also the ketone 4-hexen-3-one (results not shown for the last two). This suggests that the selected metabolites are coming from the chicken flavour component rather than from the YP. On the other hand, the remaining selected volatiles seem to be present at higher amounts in the process flavours PFA, PFB, PFF and PB3, without large differences appearing between chicken flavour levels (Supplementary **Figure S2**). The association with the chicken odour in the samples was not the chicken flavour ingredient itself but rather the combination of different chicken flavour levels and the YP from the category of process flavours (PFA, PFB, PFF and the blend PB3). Interestingly, in the bouillon samples containing the process flavour blend PB3, chicken odour was highly scored by the panellists, however, not for the other process flavours mentioned. This suggests that additional attributes (i.e. chicken-fl, sulphury-fl, intensity-fl, etc.) and, thus other volatile compounds, may still influence the overall complex perception of the chicken aroma. Certain volatiles from YP can clearly enhance certain base attributes like chicken, making those volatiles extremely interesting for future food formulation studies. Therefore, more detailed analysis on the overall models for all attributes needs to follow.

4. Conclusions

This study has helped develop a comprehensive sensory and chemical data-driven approach in a large set of chicken bouillon samples, which have been prepared using contrasting YP to enhance desired sensory characteristics. Although establishing a model describing the effect of YP to the chicken bouillon sensory attributes proved to be difficult, the present work confirms that the contribution of certain YP influences the sensory perceptions of chicken bouillon and provides leads how volatile compounds play a role in this. Predictive models were presented for roast odour and chicken odour. Predictive volatiles for roast odour were two pyrazines as well as other non-identified compounds which appear to have pyrazine-related nature. This makes the presence of these volatiles important for enhancing roast odour in chicken bouillon. The addition of YP known to contain these predictive volatiles can potentially be used to develop new products with the enhanced desired characteristics. However, it would first be advisable to use (food grade) spiking experiments with these individual or combined compounds to validate the sensory impact of these molecules.

The combined sensory-chemical approach and the proposed tools described here can be used to define sensory and chemical properties of similar processed food matrices with YP to monitor the flavour formation in food design using more efficient procedures. Careful attention is needed for the design of the sample products, so that both sensory and chemical space are properly covered. Prediction models might then be further applied on a larger set of food products to estimate the sensorial properties by using instrumental analysis only. This approach can be used to save some time in food design by removing the need for taste panels at the early stages of product design, however still sensory tests are always going to be needed to assess potentially interesting new products by testing overall flavour, texture and flavour-texture interactions.

Supplementary material

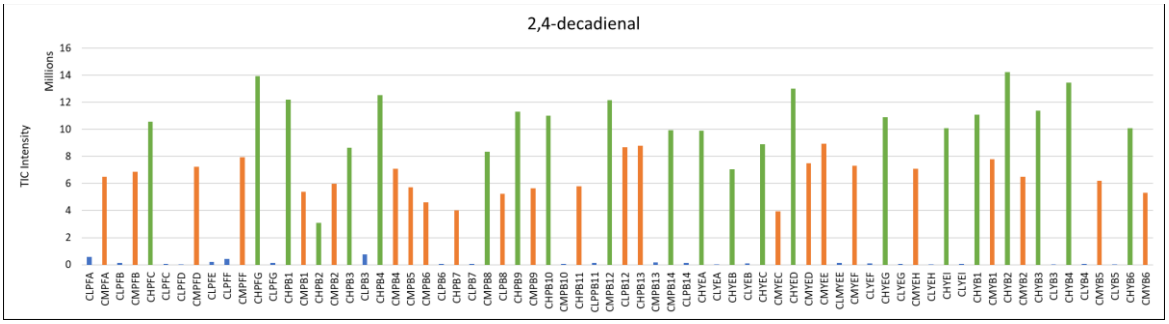


Figure S1. Bar graph showing the TIC intensity of the volatile 2,4-decadienal present in the chicken bouillon samples as ordered in **Table 1**. Samples with high chicken flavour level (green), medium chicken flavour level (orange) and no chicken flavour (blue).

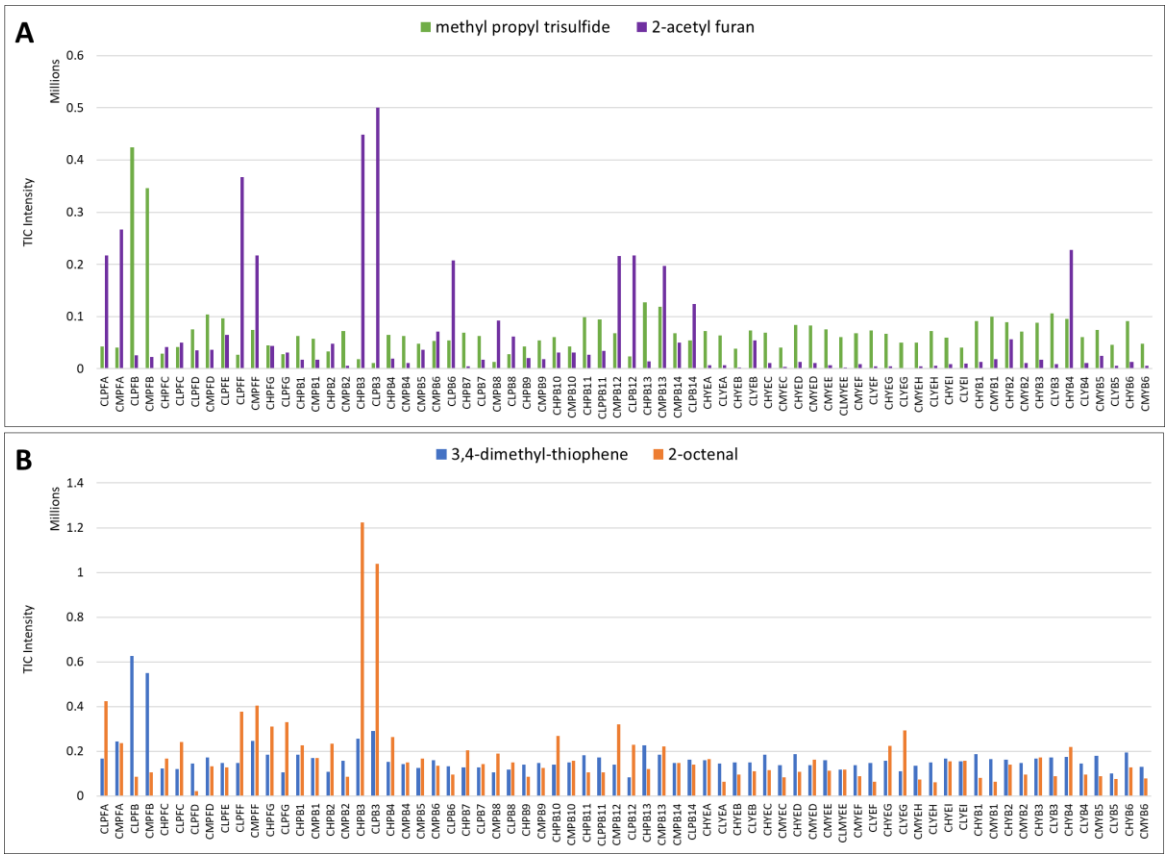
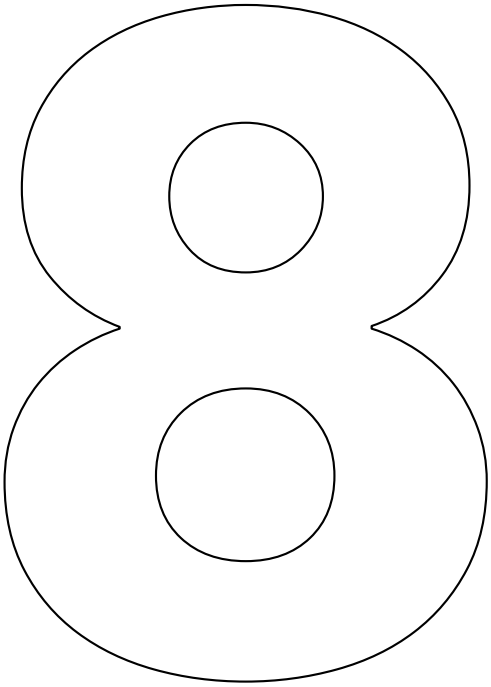


Figure S2. Bar graphs showing the TIC intensity of the selected volatiles for chicken odour present in the chicken bouillon samples as ordered in **Table 1**. **(A)** Methyl propyl trisulfide (green) and 2-acetyl furan (purple). **(B)** 3,4-Dimethyl thiophene (blue) and 2-octenal (orange).



General Discussion

Untargeted metabolomics approaches for volatile compounds: potential and limitations

The core of this thesis is centred around the development and comparison of advanced analytical methodologies to obtain a more comprehensive picture of the volatile compounds of a processed food sample. By improving metabolite coverage, more detailed conclusions can then be made on the changes and interactions occurring during the different steps involved in food processing. This section of the discussion focuses on examining and discussing different untargeted metabolomics methodologies that were designed to analyse a comprehensive number of volatile compounds present in a variety of complex savoury ingredients and place them in the context of current knowledge. These methodologies include particularly different approaches for extracting volatiles, namely SBSE, HSSE, SPME and DHS (all coupled with GC-MS) using complementary but contrasting sample types. These were two different yeast derived products (YP) (**Chapter 4**), a range of soy sauces (**Chapter 5**), tomato soups containing YP (**Chapter 6**) and chicken bouillons containing YP (**Chapter 7**). The choice to focus on these sample types was related to the technological challenges to be incurred as well as our limited knowledge of these materials despite their broad importance in the food industry.

For many years, metabolomics technologies have been progressing towards better accuracy, detectability, comprehensiveness and ease of use. Moreover, they have been applied to an ever-increasing range of samples, from contaminated air, to plant tissues, to human cells, to forensic samples, to processed foods (Beleggia et al. 2011; Hall 2006; Kim et al. 2016; Sussulini 2017; Szeremeta et al. 2021). Considering the diversity and complexity levels of the different sample types, it is evident that methodologies need to be adapted for each specific type of biological sample, as well as adapting them to the particular aim of the study. While the Holy Grail of a 'one size fits all' – type method would be ideal, in practice this will likely never be feasible. In the case of savoury food ingredients such as YP, the use of untargeted metabolomics methodologies has not yet been fully exploited. Only a few studies describe volatile analysis on YP, such as yeast extracts and yeast extract pastes which are made for different flavouring purposes (Lin et al. 2014; Mahadevan and Farmer 2006; Raza et al. 2019; Wang et al. 2020; Zhang et al. 2017). Finding out what are the typical volatiles that characterize YP is not simple, as each product has often been produced/prepared using different processing procedures (and with the addition of various base-ingredients) specifically to obtain a range of differentiating aromas. Therefore, the volatile composition will likely differ between the reported studies and the present study. Moreover, in each study, a distinct methodology to extract and then analyse the volatiles was used, which again makes findings difficult to compare. Nevertheless, the untargeted methods designed and applied here have clear added value as previously mainly only identified compounds are discussed which entails that yet unknown compounds of great potential relevance are ignored despite their equal importance from a compositional perspective. Therefore, for the studies described here, it was decided to give particular emphasis to the selection of the most appropriate technique(s) for our purpose in combination with an unbiased metabolomics approach taking both known and unknown compounds equally into account.

In **Chapter 4**, a comparison between four sorptive-based extraction techniques (SBSE, HSSE, SPME and DHS) was performed to select the one that detected the broadest number of volatiles (identified and non-identified) and in a reproducible manner using two contrasting PF samples that should contain a high diversity and range of volatiles. These PFs were chosen for their abundance in volatile chemical groups relevant to defining aroma and flavour profiles. The results were compared both in an unbiased way by looking at the whole spectrum of detected volatiles (identified and non-identified compounds) as well as by targeting specific compounds that were differentially trapped by the different techniques. The processed data of the different trapping techniques were analysed for their repeatability, sensitivity and selectivity. Results revealed that for both PFs, SBSE, where the volatiles are trapped directly from within the liquid, did extract the highest number of volatile compounds. With the static headspace SPME also a high number of compounds was detected, while the dynamic headspace trapping (DHS) and the headspace sorptive extraction (HSSE) appeared to detect the smallest number of volatiles for these PFs. As an example, in one of the PF (S99, Figure 2D, **Chapter 4**), 53 volatiles were commonly trapped by all four techniques however SBSE trapped an additional 43 unique volatiles, while only 1-12 volatiles were uniquely trapped by the other techniques. Those unique volatiles in SBSE were generally characterized by having higher molecular weights and semi-volatility. Many were also listed as 'unknowns'. SBSE has previously demonstrated its effectiveness in trapping both non-polar and semi-polar compounds from other liquid food samples (Ochiai et al. 2018). Moreover, SBSE has delivered more comprehensive profiles than SPME in liquid matrices, such as wine (Caven-Quantrill and Buglass 2011) and coffee (Bicchi et al. 2002). Interestingly, in this thesis, SBSE was also identified as being the technique having the highest volatile coverage, as compared to SPME, in soy sauce samples. SPME trapped a total of 246 volatiles, as compared to 542 for SBSE. Soy sauces were characterized by having a higher number of volatile acids, alcohols, esters, pyr(an)ones and pyrazines compared to the PFs, which contained more terpenoids, sulphur aliphatic and aromatics. SBSE trapped more, less-volatile compounds, including phenylethyl acetate and ethyl cinnamate. On the other hand, SPME exclusively trapped some highly volatile compounds such as 2-pentanone and 2-butenal. Both techniques also trapped a number of unknown compounds. Additionally, when applying these techniques to the analysis of more complex food products, such as tomato soups containing oils, sugars and YP, SBSE again appeared the more comprehensive technique, as compared to SPME. In this tomato soup study (**Chapter 6**), SPME trapped a total of 331 detected metabolites in the tomato soups, while 482 metabolites were detected by SBSE. So purely on the basis of numbers, SBSE out-performs SPME for both individual food ingredients (process flavours, PFs) as well as when these are also present in a more complex food matrix. This supports similar findings for other food products as reported by Bicchi et al. (2002), Lee et al. (2019) and Perestrelo et al. (2009).

It must, however, be borne in mind that comprehensiveness is not the only analytical trait of importance in the selection of a methodology. Aspects such as repeatability and sensitivity should also be considered, particularly in the context of the final objective of the analysis. In the analyses reported here, in terms of sensitivity, SBSE resulted in higher intensities for most compounds. This may be because SBSE methods, as compared to SPME, uses a larger volume of sorptive phase (24 µL in SBSE compared to 0.5 µL in SPME) and consequently, higher sensitivities can be achieved (Caven-Quantrill and Buglass 2011). The repeatability

was actually lower for the SBSE and SPME data compared to that of HSSE and DHS (Table 1, **Chapter 4**). Repeatability for the SPME and SBSE data in the soy sauce analyses was quite similar (RSD 11% and 13% respectively, **Chapter 5**). For the tomato soup data, while SPME resulted in better repeatability, both techniques had acceptable values for our purpose as most compounds had values below 20% (Figure 1, **Chapter 6**). Clearly, there is no perfect approach and each method has its own strengths and limitations in both qualitative and quantitative terms. Consequently, informed choices are needed - if the goal of the study is to obtain accurate and precise measurements of specific targeted compounds, SPME would be the choice, assuming they are trapped with this method. However, when the aim is to analyse the broadest possible metabolite profile of samples, avoiding any bias due to prior knowledge, SBSE would be most appropriate. However, practically it is not always easy to know beforehand which method should be selected when analysing a large-scale experiment of complex savoury products and without prior knowledge of the chemical composition. Therefore, scientists should also develop strategies that help in the decision making for selecting the most appropriate technique for a certain sample-metabolite study. In such situations, pilot analyses are to be highly recommended. **Chapter 6** focused on the comparison of two methodologies (SBSE and SPME) used in combination with an untargeted approach by creating a pipeline that inspects the performance of methodologies in an easy and straightforward manner - in this case for a small set tomato soup products containing YP in addition to a few differences in the raw ingredients added.

Looking in detail at some of the groups of identified compounds found using the different techniques revealed that some volatile classes are better represented by some extraction techniques than others. SPME is able to trap a considerable number of sesquiterpene hydrocarbons, and this may be related to the apolar nature of the compounds as well as the coating polymers of the SPME fibre used. Besides having less sorptive phase volume, SPME fibres are currently available in several combinations of different types of coating, such as the combined coating PDMS/DVB/CAR in which the properties of different polymers are blended/incorporated, while for SBSE and HSSE, the available coatings are limited to just two types (PDMS and EG-Silicone). Consequently, SPME as an approach is more versatile regarding the selectivity range of the technique. Another noticeable observation was the high contribution of SBSE to trapping pyrazines, as compared to the other techniques. Pyrazines are semi-polar volatiles which have medium-solubility in water, and thus they are trapped more effectively in the liquid phase. Additionally, SBSE also trapped more efficiently (or uniquely in some cases) sesquiterpene alcohols. The high molecular weight (hence less volatile) and higher polarity of these compounds could entail that they are poorly released into the headspace and hence are more surely available for trapping by that technique (Richter et al. 2017). Sesquiterpenoids are of potential particular relevance to the samples used here as they are characterized by conferring several green/herbal/fragrant/flower attributes and pyrazines are known to give roasted flavour aromas to food (www.foodb.ca).

It can be concluded that the volatile extraction method with GC-MS analyses has significant impact on the volatile profiles of savoury ingredients, such as YP and soy sauce as well as on foods themselves such as tomato soup. All these products are derived through complex fermentation and/or processing procedures which enrich their flavour composition. There is effectively no ideal approach as each has its limitations -

but being aware and taking account of these limitations is of key importance when data is being interpreted and conclusions drawn. The extraction techniques used in this thesis are sorptive-based (solventless) methods that have already been widely used in many food aroma analyses (Perestrelo et al. 2009). Here, we expanded the applicability of these methods to processed savoury food ingredients. The contrasting properties of each technique is important when looking at a high diversity of volatile aroma compounds, having differences and/or similarities in the volatility, solubility and polarity of the compounds. In fact, this is a crucial step when the focus of the analysis is the contribution of these volatiles to the aroma and taste of processed food, as in processed food, a broader range of compound classes are formed (Diez-Simon et al. 2019). Flavour is a result of a complex interaction between hundreds of molecules co-interacting together in the food matrix (Guichard 2002, 2015). Carefully weighed choices must be made regarding the best selection or combination of analytical methods to employ in a specific metabolite-sample study. For broadest comprehensiveness more than one protocol might be needed depending on the chemical complexity and diversity of the specific samples to be characterised. Combining methods however, does have the disadvantage of incurring extra labour and input costs. On the other hand, if comprehensiveness is not the goal, but rather more precise measurements are needed for specific (known) compounds, single tailor-made approaches might be preferable. Similarly, when quantification of the metabolites is the goal, different approaches would be taken. In the research described in this thesis, the main objectives were driven by the prediction power of the metabolite data to forecast flavour characteristics. Therefore, combining straightforward extraction techniques with untargeted methodologies and data-driven approaches that expand the number of molecules (volatiles) detected in a large set of samples is one way to improve our ability for flavour prediction in food (ingredient) analyses. Furthermore, by always including the 'unknown' molecules we have then an effective way of finding those new molecules of potential specific relevance which can still be used (as markers) 'as is', i.e. without chemical annotation, or can then become the target of dedicated metabolite identification approaches.

Untargeted volatile metabolomics to relate production procedures and ingredients

Fermentation is a natural process in which metabolic changes occur, such as the conversion of sugars under anaerobic conditions to form organic acids and alcohols. Microorganisms like yeast and bacteria play an important role, since they are usually providing the enzymes needed to breakdown the raw ingredients. Not only is fermentation used to preserve food, but also importantly, it can give food an extraordinary flavour and texture which cannot be found in fresh materials (Steinkraus 2002). Moreover, fermentation processes are considered to enrich our diet in terms of health-promoting components (e.g. prebiotics) essential for a healthy gut microbiome, and can make food taste more appetizing (Yang et al. 2020). Several production procedures are used in industry to create fermented products with superior sensory impact. Fermentation, heat treatment, and physical treatment (e.g. shredding) are often also combined to create a broader range of desirable aroma compounds. Although some of the reactions occurring have not yet been fully described, researchers are trying to monitor the processes involved according to the chemical composition obtained. Here is where multivariate analysis and nontargeted methodologies can effectively be used to analyse volatile compounds to relate production procedures. The previous section focused on nontargeted

methodologies, while this section will discuss the outcome of applying these methods for evaluating different ways of production processes influencing volatile formation.

In this thesis, we focused on analysing contrasting savoury food products that were characterized by the processes mentioned above (fermentations, heat treatments, etc.). One of these was soy sauce where a range of commercially-available products using distinct combinations of raw ingredients and fermentation conditions was studied for their aroma. More than 300 volatiles have been described in different soy sauces in earlier studies (Gao et al. 2010; Lee et al. 2006; Song et al. 2015a; Sun et al. 2010) with alcohols, acids, esters, and aldehydes being the most abundant volatile classes. However, some soy sauces on the market are non-fermented, meaning that no microorganisms have been involved in the production process. Nevertheless, interestingly, the final product still tastes like soy sauce, while some typical fermented notes may not be present. Consequently, in these soy sauce types, the flavour will have been formed by chemically hydrolysing the proteins and sugars in the raw ingredients (soybeans and wheat), in the presence of high concentrations of acid (HCl) under high temperature conditions (Diez-Simon et al. 2020b).

All soy sauces appear to have differences in their sensory characteristics, related to whether they have been fermented or non-fermented, are Chinese-style, Japanese-style, tamari, low-salt, etc. (Feng, Cai, et al. 2014; Kaneko et al. 2013). This suggests that the composition of aroma compounds may vary as a result of the production history of the soy sauces. **Chapter 5** reports a comprehensive metabolomics analysis of volatiles from twenty commercial soy sauces of different origins, production methods, or different raw ingredient composition. Multivariate analyses revealed that the largest difference in the soy sauce volatile profiles was related to the production procedure applied: i.e. fermented versus non-fermented (acid hydrolysed). Thus, although both types are being recognised as soy sauces their aroma profiles can be clearly distinguished analytically. Volatiles typical to the fermentation process were found highly or uniquely present in the fermented products (Group III, Figure 3, **Chapter 5**). Heat treatments during non-fermented production also appear to trigger other types of chemical reactions more typical to food processing, including Maillard reactions, lipid oxidation, and interactions between Maillard and breakdown processes (from sugars, lipids, amino acids, vitamins, etc.; **Chapter 2**). One example of typical Maillard reaction products are the heterocycles, like pyrazines and pyr(an)ones which have been found in many of the savoury products throughout this thesis. For instance, 2-ethyl-6-methylpyrazine is a well-known product of the reaction between Strecker aldehydes and alpha-aminoketones (Müller and Rappert 2010). Pyrazines are described as having roasted, nutty, cocoa and sweet notes. Due to their low odour threshold values, they are also often considered powerful aromatic compounds particularly in heat-treated foods (Yanfang and Wenyi 2009). When looking at the GC-O-MS data obtained in **Chapter 5**, sweet notes were associated with pyrazines, and were only detected in the non-fermented soy sauce sample (DAP, Table 3, **Chapter 5**). However, a lower abundance of 2-ethyl-6-methylpyrazine was also found in some of the fermented soy sauces (Group II, Figure 3, **Chapter 5**) suggesting that manufacturers may have employed heat treatments after fermentation has been completed (Gao et al. 2010). Pasteurization generally uses temperatures around 80°C and this may trigger some of the Maillard reaction products, at low level. Some authors have also linked the formation of pyrazines with fermentation processes occurring with specific microorganisms (Song et al. 2015b),

although their route of formation is still unclear. Other relevant compounds present in non-fermented soy sauces are acids, such as benzoic and sorbic acid. Although benzoic acid can be a bioconversion product of phenylalanine in yeast, its presence in soy sauce is mostly associated with the addition of sodium benzoate (E211) which is used as a preservative (Hazelwood et al. 2008b; Sun et al. 2010). Sorbic acid (E201) is another chemical preservative commonly used in soy sauce for its antifungal properties (Montaño et al. 1995). Notably, in our analyses, only soy sauces that were suspected (due to incomplete labelling) to be non-fermented, or were blended, had one of these acids added to the ingredient list. This is potentially important for food authentication as these acids might also serve as indicators of non-fermented/mixed soy sauces. While many of the annotated metabolites help validate previous findings, in addition, a large number of yet unknown volatiles also displayed highly contrasting patterns in their abundance between the different soy sauces. Examples are shown in Table 2 (**Chapter 5**). These unknowns may not only become useful chemical markers for acid hydrolysed (non-fermented) soy sauces, as well as other specific fermented variants, such as tamari and low-salt soy sauces but also, will help us to better understand the contrasting chemical reactions taking place during the production process. Thus, such unknowns, especially when they appear to be associated with production histories, warrant deeper analyses and annotation.

Most studies on volatile compounds in soy sauces have focused on the changes related to distinct fermentations/treatments, for example, during long-fermentation processes, using different yeast strains and other fungal species, etc. (Lee et al. 2019; Song et al. 2015a). However, we are aware of only a single study that examined the differences between fermented and non-fermented soy sauces (Lee et al. 2006). Through the research described in this thesis, our knowledge of volatile information of non-fermented soy sauces has thus been expanded, and new potential indicators of non-fermented processes have been annotated. This has given deeper insights into the chemical reactions occurring, from raw ingredients to the final product. Moreover, odour qualities could also be linked to non-fermented soy sauces, which were found to have certain distinct odours as compared to fermented sauces. What was also innovative in this study, was the description of the SBSE extraction method coupled to GC-MS as being comprehensive and robust. SBSE has only recently been used once for a small subset of soy sauces which had undergone long-fermentation treatments (Lee et al. 2019).

Manufacturers of acid hydrolysed (non-fermented) soy sauce often use limited labelling where the product label does not indicate the use of acids and high temperatures. The use of the words “fermented” and “naturally brewed” on the label are of course also avoided so the production process is generally unclear. Volatiles have proven in this study to be able to differentiate between these two processes, and thus, could prove an appropriate approach for use in food fraud and food authentication studies. Currently, the production of non-fermented soy sauces is starting to be banned in certain countries such as China, where also mixtures of fermented + hydrolysed sauces are also no longer marketable. In this thesis, we have also shown that natural ‘spontaneous’ complex fermentations, typical of the higher quality soy sauces, are difficult to mimic/simulate by artificial (non-fermented) processes (Carmen Diez-Simon et al. 2021).

The use of untargeted metabolomics analyses also proved effective when using processed (instant) soups where YP, such as the process flavours analysed in **Chapter 4,** or similar YP are added. Volatile differences

were also discussed in this thesis for processed tomato soups and chicken bouillon materials, which were prepared using different ingredients such as various YP, oil types or processing conditions, thus making the food matrix more complex. In the case of tomato soups with YP added, volatiles belonging to the group of S-containing compounds proved to correlate with the presence of one YP, named O31 (Figure 3, **Chapter 6**). This suggests that the processing used to prepare O31 triggered the formation of volatile compounds that were not present in the other tomato soups without this supplement. Or, on the contrary, the interaction of chemicals from the relatively complex soup matrix may disrupt the detection of these specific compounds in the other soup samples. Aroma retention by the presence of lipids in the matrix is often significant for sulphur compounds such as the disulfides (**Chapter 6**) and has been proven for dynamic headspace experiments (Gijs et al. 2000). In any manner, the addition of the yeast derived product O31 to instant tomato soups appears to result in higher amounts of S-containing compounds which could change the sensory perception of the product. This will be further discussed in the next section. Overall, after the analyses of all these products, we can conclude that untargeted methodologies developed in this thesis, combined with multivariate approaches are better able to relate changes in the volatile composition with changes in the preparation of processed food ingredients, as well as in the final (more complex) products.

Volatile content to predict sensory evaluations

For many years, scientists and food designers have tried to understand the relationship between chemistry and flavour in food ingredients, in order to evaluate and improve food processing strategies. The main challenge is related to the complex multiple interactions that occur in the food matrix and that, today, many are still unknown. For industry, formulating new products with desired sensory qualities is of great relevance, especially for YP often used as natural substitutes of artificial flavours to enhance the aroma and taste of instant soups and meals, bouillons and snacks (DSM). The purpose of this section is to discuss the fundamentals for obtaining good prediction models that can also be used for future formulation studies, and to propose predictive volatile compounds related to desired (or off-flavour) attributes.

Chapter 6 and 7 of this thesis have described a continuing experimental study to correlate volatiles with sensory characteristics using data-driven modelling approaches. In **Chapter 6** ('the pilot study'), the goal was to lay down a strategy to help in the selection of one of the two metabolomics techniques (SPME- and SBSE-GC-MS, developed in **Chapter 4 and 5** of this thesis) that had a better performance for the predictive models. For that, tomato soups with added YP were first tested and elastic net- and PLS-based strategies for predicting sensory (intensity odour and umami flavour) and ingredients (O31 and oil type) were evaluated (Fig 3 and 5, **Chapter 6**). By proposing three strategic levels to inspect the performance of the models, the study concluded that the SBSE method, although less repeatable than SPME, was able to predict better the intensity odour attribute. To continue the study of the sensory impact induced by added YP, new samples were used comprising a broader range of YP and leading to more diverse sensory perceptions and chemical composition. In **Chapter 7** ('the larger study'), the aim was to correlate volatile groups with sensory attributes by applying predictive models in a data-driven efficient pipeline, re-adjusted regarding the outcome of the pilot study. To do that, chicken bouillon samples with YP added were used and Random Forest-based strategies for predicting roast odour and chicken odour were evaluated (Fig 5, **Chapter 7**).

Two main limitations appeared on modelling YP additions to processed food. One was related to the sensory space and the other one was related to the chemical space. In the tomato pilot study, predictive models were generated and limitations regarding the variability in the individual sensory attributes were found, while in the bouillon samples, more variability was incorporated. This was partly due to the strong tomato base flavour which overlapped the differences between the flavour of the YP added. On the other hand, the volatile composition of the tomato soups was able to clearly distinguish between the YP added. In the bouillon samples, many volatile compounds appeared to be specific (or even unique) for many samples. These specific volatiles are very interesting since they can easily distinguish between the different YP. However, these unique volatiles were not able to explain the whole variance found in the data and as a result, the models became less strong. Random Forest was therefore used instead of PLS as it does not require transformation of data and also avoids problems with sparse data (Gromski et al. 2015). In conclusion, to obtain a strong prediction model, both chemical and sensory spaces need to be well covered by the samples. Furthermore, to be able to find molecular interactions relevant for sensory properties, it is important that these molecules are present in multiple samples. In other words, while data need to be variable, they still also need to share sufficient common variables. Finding a good balance of sample products which cover the sensory and chemical space is therefore key for generating good prediction models.

Along with learning the fundamentals for evaluating the performance of the models applied in the two studies, it is also important to account and evaluate the predicted outcome. For the tomato soup study, the most discriminative sensory attributes were intensity-od (odour) and garlic-fl (flavour), and the volatiles that were proposed in this thesis to be correlated with intensity-od were 2,6-diethyl pyrazine, methylbutanal and eugenol, highly trapped by the SBSE method (Fig 5, **Chapter 6**). On the other hand, polysulfides, such as allyl methyl disulfide and diallyl trisulfide, were proposed to be associated with the garlic flavour attribute, and have also been described in garlic samples before (Lee et al. 2003).

For the bouillon study, the most discriminative sensory attributes were roast-od, chicken-od, sulphury-od off-odour, salt-fl, umami-fl, roast-fl, chicken-fl, herbs-fl, sulphury-fl and balance-fl (Fig 5A, **Chapter 7**). As an example, for roast odour, the selected variables were 2-ethyl-5-methyl pyrazine, 2,3,5-trimethyl-6-isopentyl pyrazine and six 'unknown' compounds (Table 4A, **Chapter 7**). The addition of YP from the category process flavours (specifically in PFB, PFD, PB6, PB12 and PB14) may therefore enhance the overall roasted odour of the chicken bouillon product, since they contain higher amounts of these pyrazines, as well as the non-identified compounds, which interestingly had the same distribution as the pyrazines (Figure 6, **Chapter 7**) and the mass spectra profile showed to be chemically related, although further identification strategies need to follow to accurately identified the compounds. The chicken bouillons containing the process flavour PFB were particularly different from the other samples in terms of the volatile content, and were also scored high in off-flavour and off-odour, besides being high in herb-fl, intensity-fl and pungent-mf. These may be related to the unbalanced (or too high) amounts of S-compounds as well as sesquiterpenes which bring off-notes to the bouillons. The same PF when analysed just in water in **Chapter 4**, interestingly showed the same volatile content, although no sensory characteristics were tested for the PF itself.

In conclusion, this thesis has expanded the knowledge of modelling of sensory evaluations for processed food products (tomato soup and chicken bouillon) to study the sensory impact induced by YP. The overall sensory effect of adding YP proved difficult to establish by modelling, as in both pilot and large study, limitations were identified. Careful attention is needed for the design of the sample products, so that sensory and chemical space is properly covered. In the future, including the non-volatile composition (LC-MS) into the models may possibly improve the performance of the prediction models. Nevertheless, the present work suggests that the contribution of certain YP can influence the sensory perceptions of chicken bouillon, such as for roast odour and chicken odour. Therefore, the addition of YP known to contain a high abundance of the predictive volatiles can potentially be used to develop new products with enhancement of the desired characteristics in a more targeted fashion. It would however, first be advisable to validate the sensory impact of these molecules and this is discussed in the following section.

Future complementary research

The work described in this thesis has demonstrated the application of comprehensive analyses of volatile compounds in various processed food ingredients and products. This PhD project was part of a consortium between three Universities (WUR, UvA, LU) and two food industry companies (Unilever and DSM). In the food processing industry, there is a strong interest in using natural food ingredients, such as YP in ready-to-eat soups, to create savoury food products with superior sensory properties. This thesis specifically focused on the analysis of the volatile composition and developed methods that can detect a comprehensive number of volatiles (identified and non-identified). However, the work can also be seen in the context of a broader approach to combine diverse methodologies to broaden our knowledge of food ingredients and their potential sensory impact. For this broader context, several steps were taken. Two studies were designed, and two different set of samples were formulated and produced for characterisation. Sensory tests were performed by trained panellists who ranked the samples according to a large number of sensory perceptions. Subsequently, volatile and non-volatile characterization of the products was performed by using diverse approaches based on GC- and LC-MS metabolomics technologies, that were optimized and developed by the academic partners. Eventually, all these data, when combined, shall allow us to correlate sensory characteristics with chemical compounds, which are also linked to the chemical processes during the preparation of the samples. Nevertheless, the fusion between GC-MS and LC-MS data (volatiles and non-volatiles) has not been considered as part of the present thesis, and therefore, it is recommended to develop further the prediction models, using all the chemical data available. By doing so, we will also gain a deeper understanding of the chemical reactions that occur from the precursors (non-volatiles) to the chemical (breakdown) pathways that trigger the formation of volatile (aroma) compounds.

In this thesis, many detected metabolites could be putatively annotated, however, a large number of yet 'unknown' volatiles have also been reported throughout this thesis which currently remain unidentified. These compounds were ear-marked as being relevant for differentiating between trapping techniques (e.g. unique volatiles trapped by SBSE), sample treatments (e.g. soy sauce fermentation processes) and sensory characteristics (e.g. roasted odour of bouillon), and are of great potential relevance for the creation of new tasty savoury ingredients. While as unknowns they can still effectively be used as 'markers', ideally we

would have great value in knowing their chemical identity. The identification of unknowns is not a straight-forward approach and requires advanced technologies and labour-intensive strategies in order to confirm and validate the chemical structures of the compounds which are not yet found in databases. Therefore, for future studies in relation to savoury processed products, more attention and time needs to be given to the annotation of the sub-set of non-identified compounds that have been recognised to be of strongest specific relevance to flavour development. This will help the food industry to continue designing and developing natural food ingredients in a more directed manner that have the desired flavour, while avoiding off-flavour formation. In the same way, industry can achieve this by being able to replace artificial flavours with natural source flavours.

The validation of the prediction models will also need to involve the evaluation of synthetic-food grade mixtures by a sensorial panel. However, this importantly depends on the availability of food grade versions of the selected molecules to create fully food-grade synthetic compositions needed to cross-validate the predictive models. Subsequently, the analytical methods developed in this thesis could then be applied to study the specific dynamic changes in the metabolome of these artificial mixtures and the data could then be added into the models. Relevant precursor molecules might also be synthesized with C^{13} labels and added to a food ingredient or food product before undergoing the processing step in order to follow the chemical reactions taking place in more specific detail. Furthermore, molecules that were selected in this study could as well be investigated for their potential direct interaction with human receptors by performing receptomics analyses. All these validation methods are included in the schematic follow-up overview in **Figure 1**. Last but not least, the untargeted metabolomics technologies developed in this thesis are now being applied by the industry partners. Application within the industry is ongoing and collaboration between academic and industry must continue for a better development and outcome of flavour research in the context of designing new and improved food products meeting the desires of the modern global consumer.

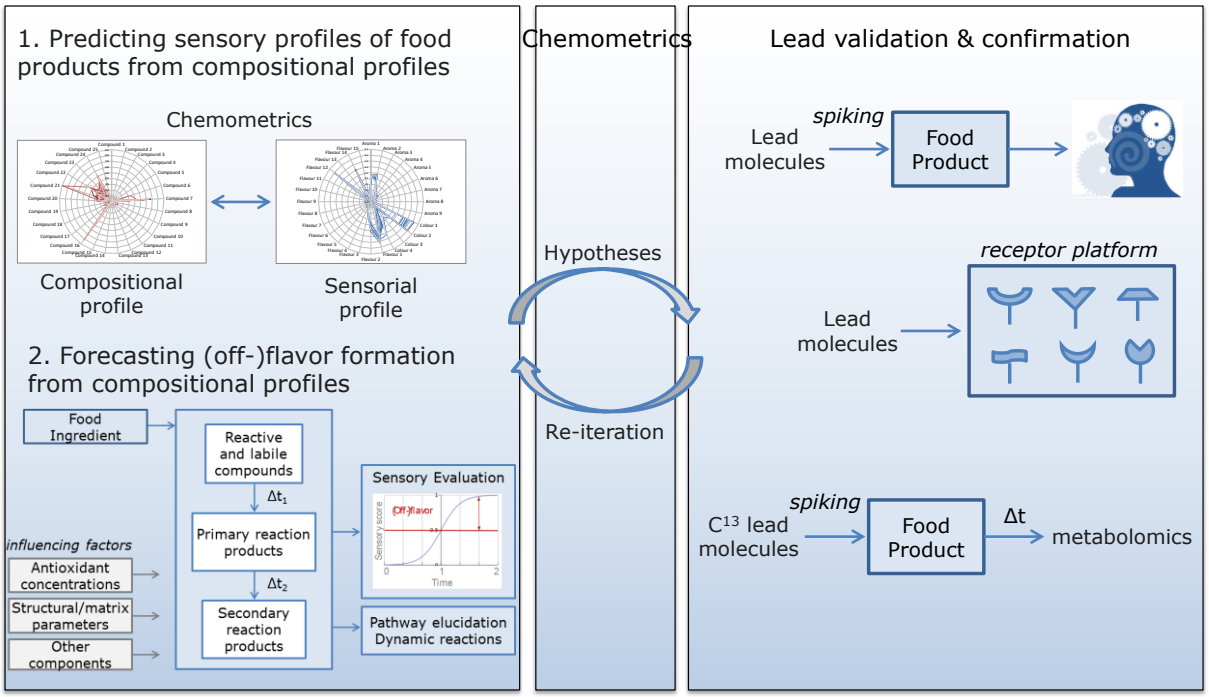


Figure 1. Schematic representation of the holistic data-driven approach required to understand the underlying chemistry and formulate hypotheses to be validated, through various methods proposed, in order to design new superior flavours.

Main conclusions

Nowadays, living in a society where we care more than ever about what we put in our mouths, and where eating food for many is more than ever becoming a great pleasure rather than being purely a necessity, consumers have a growing desire to better understand food ingredients, food composition and aspects of quality, such as aroma. Research in the past decade has rapidly led to an increased understanding of the role of volatile compounds in many types of food. However, our knowledge of the chemical reactions leading to the formation of these volatiles during food processing and which are present and impact the quality of the final products still needs to be expanded. This thesis has revealed improved methodologies for the analysis of volatiles and has broadened insights into the volatile compounds formed by the processing of specific food ingredients such as YP, which are increasingly being employed to improve natural flavour and reduce or avoid chemical supplements. During food processing, chemical complexity can increase considerably (e.g. as seen here with soy sauce) and involves a wide range of volatile groups having contrasting aroma and physicochemical properties. Although no single technique is ideal to determine the entire volatile composition of a sample, this thesis has developed more efficient approaches to expand the number of volatile compounds which can be detected. The combination of a simplistic sample preparation, hardly invasive extraction procedures, unbiased data processing, and data-driven statistical analyses offers an excellent strategy to perform predictive modelling of food aroma at large-scale application and in a time efficient manner (**Figure 2**). To place this work more in a global context, where food systems should transition towards healthier food alternatives, with minor (limited) environmental impact, technological development and food design innovations can greatly help move towards approaches that efficiently

monitor food processing properties, such as food quality flavour, all along the food processing chain. Here is where volatile compounds can play a role since these are some of the most important molecules influencing the flavour of our food and thus, they become a valuable way to monitor and develop more efficient food production systems.

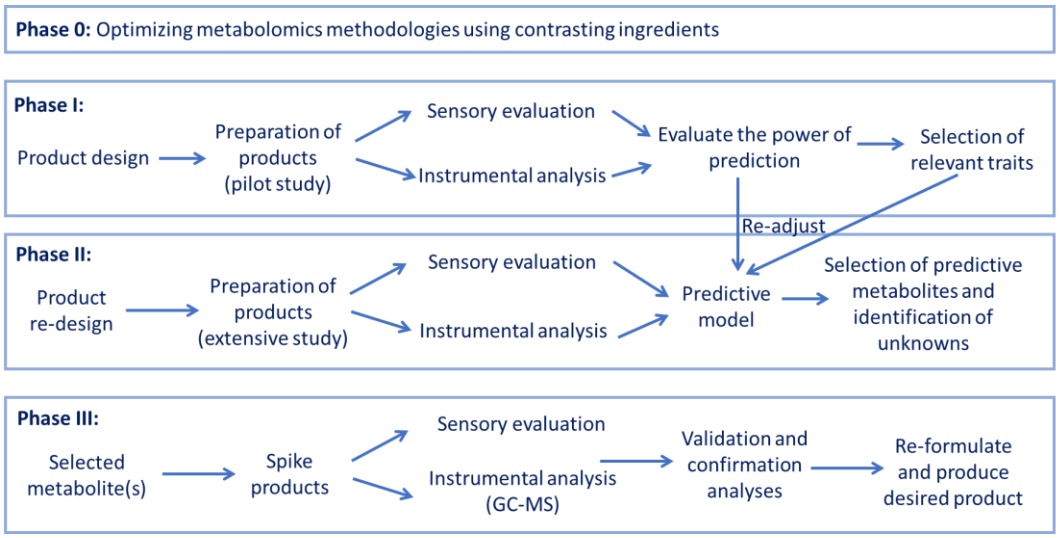


Figure 2. The schematic data-driven pipeline approach proposed in this thesis, for the formulation of superior flavour products using natural YP.

References

Adams, A. N., Polizzi, V., Van Boekel, M., & De Kimpe, N. (2008). Formation of pyrazines and a novel pyrrole in Maillard model systems of 1,3-dihydroxyacetone and 2-oxopropanal. *Journal of Agricultural and Food Chemistry*, *56*(6), 2147–2153.

Adams, R. L., Mottram, D. S., Parker, J. K., & Brown, H. M. (2001). Flavor - protein binding: Disulfide interchange reactions between ovalbumin- and volatile disulfides. *Journal of Agricultural and Food Chemistry*, *49*(9), 4333–4336.

Aernouts, B., Polshin, E., Saeys, W., & Lammertyn, J. (2011). Mid-infrared spectrometry of milk for dairy metabolomics: A comparison of two sampling techniques and effect of homogenization. *Analytica Chimica Acta*, *705*(1–2), 88–97.

Aliani, M., & Farmer, L. J. (2005). Precursors of chicken flavor. II. Identification of key flavor precursors using sensory methods. *Journal of Agricultural and Food Chemistry*, *53*(16), 6455–6462.

Allwood, J. W., Cheung, W., Xu, Y., Mumm, R., De Vos, R. C. H., Deborde, C., et al. (2014). Metabolomics in melon: A new opportunity for aroma analysis. *Phytochemistry*, *99*, 61–72.

Alonso, A., Marsal, S., & Julia, A. (2015). Analytical methods in untargeted metabolomics: state of the art in 2015. *Frontiers in Bioengineering and Biotechnology*, *3*, 1–20.

Alstrup, J., Petersen, M. A., Larsen, F. H., & Münchow, M. (2020). The effect of roast development time modulations on the sensory profile and chemical composition of the coffee brew as measured by nmr and dhs-gc–ms. *Beverages*, *6*(4), 1–14.

Ames, J. M. (1994). Volatile sulfur compounds in yeast extracts. In *Sulfur Compounds in Foods* (Vol. 564, pp. 147–159).

Ames, J. M., & Elmore, J. S. (1992). Aroma components of yeast extracts. *Flavour and Fragrance Journal*.

Antonio, C. (2018). Plant metabolomics : methods and protocols. New York: Humana Press.

Aprea, E., Corollaro, M. L., Betta, E., Endrizzi, I., Demattè, M. L., Biasioli, F., & Gasperi, F. (2012). Sensory and instrumental profiling of 18 apple cultivars to investigate the relation between perceived quality and odour and flavour. *Food Research International*, *49*(2), 677–686.

Attokaran, M. (2017). *Natural Food Flavors and Colorants*. (John Wiley & Sons Ltd., Ed.). Wiley-Blackwell.

Auvray, M., & Spence, C. (2008). The multisensory perception of flavor. *Consciousness and Cognition*, *17*(3), 1016–1031.

Baker, J. M., Ward, J. L., & Beale, M. H. (2012). Combined NMR and Flow Injection ESI-MS for Brassicaceae Metabolomics. In N. W. Hardy & R. D. Hall (Eds.), *Plant Metabolomics: Methods and Protocols* (pp. 177–191). Totowa, NJ: Humana Press.

Batenburg, A. M., De Joode, T., & Gouka, R. J. (2016). Characterization and modulation of the bitterness of polymethoxyflavones using sensory and receptor-based methods. *Journal of Agricultural and Food Chemistry*, *64*(12), 2619–2626.

Batenburg, M., & van der Velden, R. (2011). Saltiness enhancement by savory aroma compounds. *Journal of Food Science*, *76*(5), 280–288.

Běhalová, B., & Beran, K. (1979). Activation of proteolytic enzymes during autolysis of disintegrated baker's Yeast. *Folia Microbiologica*, *24*(6), 455–461.

Behrens, M., Meyerhof, W., Hellfritsch, C., & Hofmann, T. (2011). Sweet and umami taste: Natural products, their chemosensory targets, and beyond. *Angewandte Chemie International Edition*, *50*(10), 2220–2242.

Beleggia, R., Platani, C., Papa, R., Di Chio, A., Barros, E., Mashaba, C., et al. (2011). Metabolomics and food processing: From semolina to pasta. *Journal of Agricultural and Food Chemistry*, *59*(17), 9366–9377.

Belitz, H.-D., Grosch, W., & Schieberle, P. (2004). Aroma Compounds. In *Food Chemistry* (3rd ed., pp. 342–408). Berlin: Springer-Verlag.

Belitz, H. D., Grosch, W., & Schieberle, P. (2008). *Lehrbuch der Lebensmittelchemie*. Berlin: Springer.

Berdyshev, E. V. (2011). Mass spectormetry of fatty aldehydes. *Biochimica et biophysica acta*, *1811*(11), 680–693.

Biasioli, F., Gasperi, F., Yeretzian, C., & Märk, T. D. (2011). PTR-MS monitoring of VOCs and BVOCs in food science and technology. *TrAC - Trends in Analytical Chemistry*, *30*(7), 968–977.

Bicchi, C., Iori, C., Rubiolo, P., & Sandra, P. (2002). Headspace sorptive extraction (HSSE), stir bar sorptive extraction (SBSE), and solid phase microextraction (SPME) applied to the analysis of roasted Arabica coffee and coffee brew. *Journal of Agricultural and Food Chemistry*, *50*(3), 449–459. https://doi.org/10.1021/jf010877x

Blank, I., & Fay, L. B. (1996). Formation of 4-Hydroxy-2,5-dimethyl-3(2H)-furanone and 4-Hydroxy-2(or 5)-ethyl-5(or 2)-methyl-3(2H)-furanone through Maillard Reaction Based on Pentose Sugars. *Journal of Agricultural and Food Chemistry*, *44*, 531–536.

Blasco, C., Fernández, M., Picó, Y., & Font, G. (2004). Comparison of solid-phase microextraction and stir bar sorptive extraction for determining six organophosphorus insecticides in honey by liquid chromatography-mass spectrometry. *Journal of Chromatography A*, *1030*(1–2), 77–85.

Bohlmann, J. S. J. (2015). *Biotechnology of Isoprenoids*. *Advances in Biochemical Engineering/Biotechnology* (Vol. 148).

Burdock, G. A. (2009). *Fenaroli's Handbook of Flavor Ingredients*. CRC Press Boca Raton.

Calingacion, M., Fang, L., Quiatchon-Baeza, L., Mumm, R., Riedel, A., Hall, R. D., & Fitzgerald, M. (2015). Delving deeper into technological innovations to understand differences in rice quality. *Rice*, *8*(1).

Calingacion, M., Mumm, R., Tan, K., Quiatchon-Baeza, L., Concepcion, J. C. T., Hageman, J. A., et al. (2017). A multidisciplinary phenotyping and genotyping analysis of a mapping population enables quality to be combined with yield in rice. *Frontiers in Molecular Biosciences*, *4*(MAY).

Camino-Sánchez, F. J., Rodríguez-Gómez, R., Zafra-Gómez, A., Santos-Fandila, A., & Vílchez, J. L. (2014). Stir bar sorptive

extraction: Recent applications, limitations and future trends. *Talanta*, *130*, 388–399.
Cao, C., Xie, J., Hou, L., Zhao, J., Chen, F., Xiao, Q., et al. (2017). Effect of glycine on reaction of cysteine-xylose: Insights on initial Maillard stage intermediates to develop meat flavor. *Food Research International*, *99*, 444–453.
Carneiro, S., Pereira, R., & Rocha, I. (2014). Yeast Metabolomics: Sample Preparation for a GC/MS-Based Analysis. In V. Mapelli (Ed.), *Yeast Metabolic Engineering: Methods and Protocols* (pp. 197–207). New York, NY: Springer New York.
Caven-Quantrill, D. J., & Buglass, A. J. (2011). Comparison of volatile constituents extracted from model grape juice and model wine by stir bar sorptive extraction-gas chromatography-mass spectrometry. *Journal of Chromatography A*, *1218*(7), 875–881.
Cevallos-Cevallos, Juan M., Reyes-De-Corcuera, J. I., Etxeberria, E., Danyluk, M. D., & Rodrick, G. E. (2009). Metabolomic analysis in food science: a review. *Trends in Food Science and Technology*, *20*(11–12), 557–566.
Cevallos-Cevallos, Juan Manuel, & Reyes-De-Corcuera, J. I. (2012). Metabolomics in Food Science. *Advances in Food and Nutrition Research*, *67*, 1–24.
Chambers IV, E., & Koppel, K. (2013). Associations of volatile compounds with sensory aroma and flavor: The complex nature of flavor. *Molecules*, *18*(5), 4887–4905.
Charve, J., Manganiello, S., & Glabasnia, A. (2018). Analysis of umami taste compounds in a fermented corn sauce by means of sensory-guided fractionation. *Journal of Agricultural and Food Chemistry*, *66*(8), 1863–1871.
Chen, Z. Y., Feng, Y. Z., Cui, C., Zhao, H. F., & Zhao, M. M. (2015). Effects of koji-making with mixed strains on physicochemical and sensory properties of Chinese-type soy sauce. *Journal of the Science of Food and Agriculture*, *95*(10), 2145–2154.
Chin, E., & Slupsky, C. M. (2013). *Applications of metabolomics in food science: Food composition and quality, sensory and nutritional attributes*. *Metabolomics in Food and Nutrition*. Elsevier Masson SAS.
Choi, S. M., Lee, D. J., Kim, J. Y., & Lim, S. T. (2017). Volatile composition and sensory characteristics of onion powders prepared by convective drying. *Food Chemistry*, *231*, 386–392.
Chris Liang. (2016). The Complete Guide to Soy Sauces.
Colina-Coca, C., González-Peña, D., Vega, E., de Ancos, B., & Sánchez-Moreno, C. (2013). Novel approach for the determination of volatile compounds in processed onion by headspace gas chromatography–mass spectrometry (HS GC–MS). *Talanta*, *103*, 137–144.
Cordero, C., Kiefl, J., Schieberle, P., Reichenbach, S. E., & Bicchi, C. (2015). Comprehensive two-dimensional gas chromatography and food sensory properties: Potential and challenges. *Analytical and Bioanalytical Chemistry*, *407*(1), 169–191.
Corollaro, M. L., Aprea, E., Endrizzi, I., Betta, E., Demattè, M. L., Charles, M., et al. (2014). A combined sensory-instrumental tool for apple quality evaluation. *Postharvest Biology and Technology*, *96*, 135–144.
Cozzolino, D. (2016). Metabolomics in grape and wine: Definition, current status and future prospects. *Food Analytical Methods*, *9*(11), 2986–2997.
Croissant, A. E., Watson, D. M., & Drake, M. A. (2011). Application of sensory and instrumental volatile analyses to dairy products. *Annu. Rev. Food Sci. Technol*, *2*, 395–421.
Cubero-Leon, E., Peñalver, R., & Maquet, A. (2014). Review on metabolomics for food authentication. *Food Research International*, *60*, 95–107.
Cuicui, L., & Lixia, H. (2018). Review on volatile flavor components of roasted oilseeds and their products. *Grain & Oil Science and Technology*, *1*(4), 151–156.
Curioni, P. M. G., & Bosset, J. O. (2002). Key odorants in various cheese types as determined by gas chromatography-olfactometry. *International Dairy Journal*, *12*(12), 959–984.
Czerny, M., Christlbauer, M., Christlbauer, M., Fischer, A., Granvogl, M., Hammer, M., et al. (2008). Re-investigation on odour thresholds of key food aroma compounds and development of an aroma language based on odour qualities of defined aqueous odorant solutions. *European Food Research and Technology*, *228*(2), 265–273.
Dagan, L., Schneider, R., Lepoutre, J. P., & Baumes, R. (2006). Stability of sotolon in acidic and basic aqueous solutions: Application to the synthesis of a deuterated analogue for its quantitative determination in wine. *Analytica Chimica Acta*, *563*, 365–374.
David, F., Ochiai, N., & Sandra, P. (2019). Two decades of stir bar sorptive extraction: A retrospective and future outlook. *TrAC - Trends in Analytical Chemistry*, *112*, 102–111. https://doi.org/10.1016/j.trac.2018.12.006
David, F., & Sandra, P. (2007). Stir bar sorptive extraction for trace analysis. *Journal of Chromatography A*, *1152*(1–2), 54–69.
De Toffoli, A., Monteleone, E., Bucalossi, G., Veneziani, G., Fia, G., Servili, M., et al. (2019). Sensory and chemical profile of a phenolic extract from olive mill waste waters in plant-base food with varied macro-composition. *Food Research International*, *119*, 236–243.
Devanthi, P. V. P., & Gkatzionis, K. (2019). Soy sauce fermentation: Microorganisms, aroma formation, and process modification. *Food Research International*, *120*, 364–374.
Diez-Simon, C., Eichelsheim, C., Mumm, R., & Hall, R. D. (2020). Chemical and sensory characteristics of soy sauce: A review. *Journal of agricultural and food chemistry*, *68*(42).
Diez-Simon, C., Mumm, R., & Hall, R. D. (2019). Mass spectrometry-based metabolomics of volatiles as a new tool for understanding aroma and flavour chemistry in processed food products. *Metabolomics*, *15*(3).
Diez-Simon, Carmen, Ammerlaan, B., van den Berg, M., van Duynhoven, J., Jacobs, D., Mumm, R., & Hall, R. D. (2020). Comparison of volatile trapping techniques for the comprehensive analysis of food flavourings by Gas Chromatography-Mass Spectrometry. *Journal of Chromatography A*, *1624*, 461191.
Diez-Simon, Carmen, Eichelsheim, C., Jacobs, D. M., Mumm, R., & Hall, R. D. (2021). Stir bar sorptive extraction of

aroma compounds in soy sauce : Revealing the chemical diversity. *Food Research International*, *144*, 110348.
Dongen, W. van, & Wiggers, B. (n.d.). VCF Online. Volatile Compounds in Food.
Dresow, J. F., & Böhm, H. (2009). The influence of volatile compounds of the flavour of raw, boiled and baked potatoes: impact of agricultural measures on the volatile components. *Landbauforschung - vTI Agriculture and Forestry Research*, *59*(4), 309–337.
DSM. (n.d.). Savory taste solutions by DSM.
Dunkel, A., Steinhaus, M., Kotthoff, M., Nowak, B., Krautwurst, D., Schieberle, P., & Hofmann, T. (2014). Nature's chemical signatures in human olfaction: A foodborne perspective for future biotechnology. *Angewandte Chemie International Edition*, *53*(28), 7124–7143.
Effenberger, I., Hoffmann, T., Jonczyk, R., & Schwab, W. (2019). Novel biotechnological glucosylation of high-impact aroma chemicals, 3(2H)- and 2(5H)-furanones. *Scientific Reports*, *9*(1), 1–9.
Elmore, J. S., Campo, M. M., Enser, M., & Mottram, D. S. (2002). Effect of lipid composition on meat-like model systems containing cysteine, ribose, and polyunsaturated fatty acids. *Journal of Agricultural and Food Chemistry*, *50*(5), 1126–1132.
Elmore, J. S., Mottram, D. S., Enser, M., & Wood, J. D. (1997). Novel thiazoles and 3-thiazolines in cooked beef aroma. *Journal of Agricultural and Food Chemistry*, *45*(9), 3603–3607.
Esslinger, S., Riedl, J., & Fauhl-Hassek, C. (2014). Potential and limitations of non-targeted fingerprinting for authentication of food in official control. *Food Research International*, *60*, 189–204.
Fan, W., & Qian, M. C. (2006). Characterization of aroma compounds of Chinese "Wuliangye" and "Jiannanchun" liquors by aroma extract dilution analysis. *Journal of Agricultural and Food Chemistry*, *54*(7), 2695–2704.
Fan, W., Shen, H., & Xu, Y. (2011). Quantification of volatile compounds in Chinese soy sauce aroma type liquor by stir bar sorptive extraction and gas chromatography-mass spectrometry. *Journal of the Science of Food and Agriculture*, *91*(7), 1187–1198.
FAOSTAT. (2017).
Feng, Y., Cai, Y., Fu, X., Zheng, L., Xiao, Z., & Zhao, M. (2018). Comparison of aroma-active compounds in broiler broth and native chicken broth by aroma extract dilution analysis (AEDA), odor activity value (OAV) and omission experiment. *Food Chemistry*, *265*, 274–280.
Feng, Y., Cai, Y., Su, G., Zhao, H., Wang, C., & Zhao, M. (2014). Evaluation of aroma differences between high-salt liquid-state fermentation and low-salt solid-state fermentation soy sauces from China. *Food Chemistry*, *145*, 126–134.
Feng, Y., Cai, Y., Sun-Waterhouse, D., Cui, C., Su, G., Lin, L., & Zhao, M. (2015). Approaches of aroma extraction dilution analysis (AEDA) for headspace solid phase microextraction and gas chromatography-olfactometry (HS-SPME-GC-O): Altering sample amount, diluting the sample or adjusting split ratio? *Food Chemistry*, *187*, 44–52.
Feng, Y., Chen, Z., Liu, N., Zhao, H., Cui, C., & Zhao, M. (2014). Changes in fatty acid composition and lipid profile during koji fermentation and their relationships with soy sauce flavour. *Food Chemistry*, *158*, 438–444.
Feng, Y., Cui, C., Zhao, H., Gao, X., Zhao, M., & Sun, W. (2013). Effect of koji fermentation on generation of volatile compounds in soy sauce production. *International Journal of Food Science and Technology*, *48*(3), 609–619.
Feng, Y., Su, G., Sun-Waterhouse, D., Cai, Y., Zhao, H., Cui, C., & Zhao, M. (2017). Optimization of headspace solid-phase micro-extraction (HS-SPME) for analyzing soy sauce aroma compounds via coupling with direct GC-Olfactometry (D-GC-O) and Gas Chromatography-Mass Spectrometry (GC-MS). *Food Analytical Methods*, *10*(3), 713–726.
Feng, Y., Su, G., Zhao, H., Cai, Y., Cui, C., Sun-Waterhouse, D., & Zhao, M. (2015). Characterisation of aroma profiles of commercial soy sauce by odour activity value and omission test. *Food Chemistry*, *167*, 220–228.
Fiddler, W., Parker, W. E., Wasserman, A. E., & Doerr, R. C. (1967). Thermal decomposition of ferulic acid. *Journal of Agricultural and Food Chemistry*, *15*(5), 757–761.
Filzmoser, P., Liebmann, B., & Varmuza, K. (2009). Repeated double cross validation. *Journal of Chemometrics*, *23*(4), 160–171.
Forss, D. A. (1973). Odor and flavor compounds from lipids. *Progress in the Chemistry of Fats and Other Lipids*, *13*(C), 177–258.
Frank, D. C., Owen, C. M., & Patterson, J. (2004). Solid phase microextraction (SPME) combined with gas-chromatography and olfactometry-mass spectrometry for characterization of cheese aroma compounds. *LWT - Food Science and Technology*, *37*(2), 139–154.
Frankel, E. N. (1980). Lipid oxidation. *Progress in Lipid Research*, *19*(1–2), 1–22.
Frerot, E., & Chen, T. (2013). Identification and quantitation of new glutamic acid derivatives in soy sauce by UPLC/MS/MS. *Chemistry and Biodiversity*, *10*(10), 1842–1850.
Gambacorta, G., Sinigaglia, M., Schena, A., Baiano, A., Lamacchia, C., Pati, S., & La Notte, E. (2009). Changes in free fatty acid and diacylglycerol compounds in short-ripening dry-cured sausage. *Journal of Food Lipids*, *16*(1), 1–18.
Gao, X. L., Cui, C., Zhao, H. F., Zhao, M. M., Yang, L., & Ren, J. Y. (2010). Changes in volatile aroma compounds of traditional chinese-type soy sauce during moromi fermentation and heat treatment. *Food Science and Biotechnology*, *19*(4), 889–898.
Gao, X., Liu, E., Yin, Y., Yang, L., Huang, Q., Chen, S., & Ho, C. T. (2020). Enhancing activities of salt-tolerant proteases secreted by Aspergillus oryzae using atmospheric and room-temperature plasma mutagenesis. *Journal of Agricultural and Food Chemistry*, *68*(9), 2757–2764.
Gao, X., Liu, E., Zhang, J., Yang, L., Huang, Q., Chen, S., et al. (2020). Accelerating aroma formation of raw soy sauce using low intensity sonication. *Food Chemistry*, *329*, 127118.
Gao, X., Liu, E., Zhang, J., Yang, M., Chen, S., Liu, Z., et al. (2019). Effects of sonication during moromi fermentation on antioxidant activities of compounds in raw soy sauce. *Lwt*, *116*, 108605.

Gao, X., Zhang, J., Liu, E., Yang, M., Chen, S., Hu, F., et al. (2019). Enhancing the taste of raw soy sauce using low intensity ultrasound treatment during moromi fermentation. *Food Chemistry*, *298*, 124928.

García-Lomillo, J., González-Sanjosé, M. L., Del Pino-García, R., Ortega-Heras, M., & Muñiz-Rodríguez, P. (2016). Effect of a new natural seasoning on the formation of pyrazines in barbecued beef patties. *Journal of Chemistry*, *2016*.

Gilart, N., Marcé, R. M., Borrull, F., & Fontanals, N. (2014). New coatings for stir-bar sorptive extraction of polar emerging organic contaminants. *TrAC Trends in Analytical Chemistry*, *54*, 11–23.

Giri, A., Osako, K., Okamoto, A., & Ohshima, T. (2010). Olfactometric characterization of aroma active compounds in fermented fish paste in comparison with fish sauce, fermented soy paste and sauce products. *Food Research International*, *43*(4), 1027–1040.

Goodacre, R., Graham, D., & Faulds, K. (2018). Recent developments in quantitative SERS: Moving towards absolute quantification. *TrAC - Trends in Analytical Chemistry*, *102*, 359–368.

Grabež, V., Bjelanović, M., Rohloff, J., Martinović, A., Berg, P., Tomović, V., et al. (2019). The relationship between volatile compounds, metabolites and sensory attributes: A case study using lamb and sheep meat. *Small Ruminant Research*, *181*, 12–20.

Greenberg, P. (2000). Soy Sauce, China's Liquid Spice. *Sauces, Seasonings, and Spices*, (7(2)), 21–22.

Gromski, P. S., Muhamadali, H., Ellis, D. I., Xu, Y., Correa, E., Turner, M. L., & Goodacre, R. (2015). A tutorial review: Metabolomics and partial least squares-discriminant analysis - a marriage of convenience or a shotgun wedding. *Analytica Chimica Acta*, *879*, 10–23.

Guichard, E. (2002). Interactions between flavor compounds and food ingredients and their influence on flavor perception. *Food Reviews International*, *18*(1), 49–70.

Guichard, E. (2015). *Interaction of aroma compounds with food matrices*. *Flavour Development, Analysis and Perception in Food and Beverages*. Elsevier Ltd.

Guichard, Elisabeth, & Salles, C. (2016). Retention and release of taste and aroma compounds from the food matrix during mastication and ingestion. In *Flavor* (Vol. 299, pp. 3–22). Elsevier.

Guo, S., Na Jom, K., & Ge, Y. (2019). Influence of roasting condition on flavor profile of sunflower seeds: A flavoromics approach. *Scientific Reports*, *9*(1), 1–10.

Halasz, A., & Lasztity, R. (1991). *Use of Yeast Biomass in Food Production*. CRC Press Boca Raton.

Hall, R. D. (2006). Plant metabolomics: From holistic hope, to hype, to hot topic. *New Phytologist*, *169*(3), 453–468.

Hall, R. D., Brouwer, I. D., & Fitzgerald, M. A. (2008). Plant metabolomics and its potential application for human nutrition. *Physiologia Plantarum*, *132*(2), 162–175.

Hamada, T., Ishiyama, T., & Motai, H. (1989). Applied microbiology biotechnology continuous fermentation of soy sauce by immobilized cells of Zygosaccharomyces rouxii in an airlift reactor. *Applied Microbiology and Biotechnology*, *31*, 346–350.

Harada, R., Yuzuki, M., Ito, K., Shiga, K., Bamba, T., & Fukusaki, E. (2018). Microbe participation in aroma production during soy sauce fermentation. *Journal of Bioscience and Bioengineering*, *125*(6), 688–694.

Hasnip, S., Crews, C., Potter, N., Brereton, P., Diserens, H., Oberson, J.-M., et al. (2005). Determination of 1,3-dichloropropanol in soy sauce and related products by headspace gas chromatography with mass spectrometric detection: Interlaboratory study. *Journal of AOAC INTERNATIONAL*, *88*(5), 1404–1412.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*. New York: Springer.

Hazelwood, L. A., Daran, J. M., Van Maris, A. J. A., Pronk, J. T., & Dickinson, J. R. (2008). The Ehrlich pathway for fusel alcohol production: A century of research on Saccharomyces cerevisiae metabolism. *Applied and Environmental Microbiology*, *74*(12), 3920.

Ho, C.-T., & Chen, Q. (1994). Lipids in food flavors. *ACS Symposium Series; American Chemical Society: Washington, DC, 1994.*, *558*, 2–14.

Hodge, J. E. (1953). Dehydrated foods, chemistry of browning reactions in model systems. *Journal of Agricultural and Food Chemistry*, *1*(15), 928–943.

Hofmann, T., Schieberle, P., & Grosch, W. . (1996). Model studies on the oxidative stability of odor-active thiols occurring in food flavors. *Journal of Agricultural and Food Chemistry*, *44*(1), 251–255.

Hosking, R. (1996). Soy Sauce. In N. Ishige (Ed.), *A dictionary of Japanese food: ingredients & culture* (1st ed., pp. 204–206). Tokyo: Tuttle Publishing.

In, M.-J., Kim, D. C., & Chae, H. J. (2005). Downstream process for the production of yeast extract using brewer's yeast cells. *Biotechnology and Bioprocess Engineering*, *10*(1), 85–90.

Inoue, Y., Kato, S., Saikusa, M., Suzuki, C., Otsubo, Y., Tanaka, Y., et al. (2016). Analysis of the cooked aroma and odorants that contribute to umami aftertaste of soy miso (Japanese soybean paste). *Food Chemistry*, *213*, 521–528.

Izzo, H. V, & Ho, C.-T. (1991). Isolation and identification of the volatile components of an extruded autolyzed yeast extract. *J. Agric. Food Chem*, *39*, 2245–2248.

Jaeger, H., Janositz, A., & Knorr, D. (2010). The Maillard reaction and its control during food processing. The potential of emerging technologies. *Pathologie Biologie*, *58*(3), 207–213.

Jayasena, D. D., Ahn, D. U., Nam, K. C., & Jo, C. (2013). Flavour chemistry of chicken meat: A review. *Asian-Australasian Journal of Animal Sciences*, *26*(5), 732–742.

Jeleń, H. H., Majcher, M., & Dziadas, M. (2012). Microextraction techniques in the analysis of food flavor compounds: A review. *Analytica Chimica Acta*, *738*, 13–26.

Jerković, I., Mastelić, J., & Tartaglia, S. (2007). A study of volatile flavour substances in Dalmatian traditional smoked ham: Impact of dry-curing and frying. *Food Chemistry*, *104*(3), 1030–1039.

Jo, Y. J., Cho, I. H., Song, C. K., Shin, H. W., & Kim, Y. S. (2011). Comparison of fermented soybean paste (Doenjang) prepared by different methods based on profiling of volatile compounds. *Journal of Food Science*, *76*(3), 368–379.

Johanningsmeier, S. D., Harris, G. K., & Klevorn, C. M. (2016). Metabolomic technologies for improving the quality of food: Practice and promise. *Annual Review of Food Science and Technology*, *7*(1), 413–438.

Jones, M. G., Hughes, J., Tregova, A., Milne, J., Tomsett, A. B., & Collin, H. A. (2004). Biosynthesis of the flavour precursors of onion and garlic. *Journal of Experimental Botany*, *55*(404), 1903–1918.

Kadam, S. U., Misra, N. N., & Zaima, N. (2016). Mass spectrometry based chemical imaging of foods. *RSC Advances*, *6*(40), 33537–33546.

Kamani, M. H., Meera, M. S., Bhaskar, N., & Modi, V. K. (2019). Partial and total replacement of meat by plant-based proteins in chicken sausage: evaluation of mechanical, physico-chemical and sensory characteristics. *Journal of Food Science and Technology*, *56*(5), 2660–2669.

Kaneko, S., Kumazawa, K., & Nishimura, O. (2012). Comparison of key aroma compounds in five different types of Japanese soy sauces by aroma extract dilution analysis (AEDA). *Journal of Agricultural and Food Chemistry*, *60*(15), 3831–3836.

Kaneko, S., Kumazawa, K., & Nishimura, O. (2013). Studies on the key aroma compounds in raw (unheated) and heated Japanese soy sauce. *Journal of Agricultural and Food Chemistry*, *61*(14), 3396–3402.

Kataoka, H., Lord, H. L., & Pawliszyn, J. (2000). Applications of solid-phase microextraction in food analysis. *Journal of Chromatography A*, *880*(1), 35–62.

Kerth, C. R., & Miller, R. K. (2015). Beef flavor: A review from chemistry to consumer. *Journal of the Science of Food and Agriculture*, *95*(14), 2783–2798.

Khan, M. I., Jo, C., & Tariq, M. R. (2015). Meat flavor precursors and factors influencing flavor precursors-A systematic review. *Meat Science*, *110*, 278–284.

Khan, M. K., Paniwnyk, L., & Hassan, S. (2019). Polyphenols as Natural Antioxidants: Sources, Extraction and Applications in Food, Cosmetics and Drugs. In Y. Li & F. Chemat (Eds.), *Plant Based "Green Chemistry 2.0": Moving from Evolutionary to Revolutionary* (pp. 197–235). Singapore: Springer.

Kim, S., Kim, J., Yun, E. J., & Kim, K. H. (2016). Food metabolomics: from farm to human. *Current Opinion in Biotechnology*, *37*, 16–23.

Klevorn, C. M., & Dean, L. L. (2018). A metabolomics-based approach identifies changes in the small molecular weight compound composition of the peanut as a result of dry-roasting. *Food Chemistry*, *240*, 1193–1200.

Koehler, P. E., Mason, M. E., & Odell, G. V. (1971). Odor threshold levels of pyrazine compounds and assessment of their role in the flavor of roasted foods. *Journal of Food Science*, *36*(5), 816–818.

Kong, Y., Yang, X., Ding, Q., Zhang, Y. Y., Sun, B. G., Chen, H. T., & Sun, Y. (2017). Comparison of non-volatile umami components in chicken soup and chicken enzymatic hydrolysate. *Food Research International*, *102*, 559–566.

Kong, Y., Zhang, L. L., Zhang, Y. Y., Sun, B. G., Sun, Y., Zhao, J., & Chen, H. T. (2018). Evaluation of non-volatile taste components in commercial soy sauces. *International Journal of Food Properties*, *21*(1), 1854–1866. https://doi.org/10.1080/10942912.2018.1497061

Kuroda, M., Kato, Y., Yamazaki, J., Kai, Y., Mizukoshi, T., Miyano, H., & Eto, Y. (2013). Determination and quantification of the kokumi peptide , γ-glutamyl-valyl-glycine , in commercial soy sauces. *Food Chemistry*, *141*(2), 823–828.

Kuroda, M., & Miyamura, N. (2015). Mechanism of the perception of "kokumi" substances and the sensory characteristics of the "kokumi" peptide, γ-Glu-Val-Gly. *Flavour*, *4*(1), 1–3.

Landaud, S., Helinck, S., & Bonnarme, P. (2008). Formation of volatile sulfur compounds and metabolism of methionine and other sulfur compounds in fermented food. *Applied Microbiology and Biotechnology*, *77*(6), 1191–1205.

Lawless, H., & Heymann, H. (2010). *Sensory Evaluation of Food Principles and Practices*. *Elementary Food Science*. https://doi.org/10.1007/978-1-4419-6488-5

Lawless, H. T., & Heymann, H. (2010). *Sensory evaluation of food: principles of good practice*. *Sensory Evaluation of Food*. Springer Science & Business Media.

Ledl, F., & Schleicher, E. (1990). New aspects of the Maillard reaction in foods and in the human body. *Angewandte Chemie International Edition in English*, *29*(6), 565–594.

Lee, B. Q., & Khor, S. M. (2015). 3-Chloropropane-1,2-diol (3-MCPD) in soy sauce: A review on the formation, reduction, and detection of this potential carcinogen. *Comprehensive Reviews in Food Science and Food Safety*, *14*(1), 48–66.

Lee, Jeehyun, Vázquez-Araújo, L., Adhikari, K., Warmund, M., & Elmore, J. (2011). Volatile compounds in light, medium, and dark black walnut and their influence on the sensory aromatic profile. *Journal of Food Science*, *76*(2), C199–C204.

Lee, Jun-young, Seok, W., Yun, K., Lee, Y., Choi, Y., Choi, H., & Jang, H. W. (2019). Solid-phase microextraction Arrow for the volatile organic compounds in soy sauce Running title : SPME Arrow GC-MS analysis of VOCs in soy sauce, 1–24.

Lee, K. E., Lee, S. M., Choi, Y. H., Hurh, B. S., & Kim, Y.-S. (2013). Comparative volatile profiles in soy sauce according to inoculated microorganisms. *Bioscience, Biotechnology, and Biochemistry*, *77*(11), 2192–2200.

Lee, L. C., Liong, C.-Y., & Jemain, A. A. (2018). Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *The Analyst*, *143*(15), 3526–3539.

Lee, S.-N., Kim, N.-S., & Lee, D.-S. (2003). Comparative study of extraction techniques for determination of garlic flavor components by gas chromatography-mass spectrometry. *Analytical and Bioanalytical Chemistry*, *377*(4), 749–

756.
Lee, S. M., Seo, B. C., & Kim, Y. S. (2006). Volatile compounds in fermented and acid-hydrolyzed soy sauces. *Journal of Food Science*, *71*(3), C146–C156.
Lee, Sang Mi, Kim, S. Bin, & Kim, Y. S. (2019). Determination of key volatile compounds related to long-term fermentation of soy sauce. *Journal of Food Science*, *84*(10), 2758–2776.
Lesschaeve, I., & Noble, A. C. (2005). Polyphenols: factors influencing their sensory properties and their effects on food and beverage preferences. *The American journal of clinical nutrition*, *81*, 330–335.
Lestremau, F., Andersson, F. a. T., & Desauziers, V. (2004). Investigation of artefact formation during analysis of volatile sulphur compounds using solid phase microextraction (SPME). *Chromatographia*, *59*(9), 607–613.
Li, X, Hiramoto, K., Yoshida, M., Kato, T., & Kikugawa, K. (1998). Identification of 2,5-dimethyl-4-hydroxy-3(2H)-furanone (DMHF) and 4-hydroxy-2(or 5)-ethyl-5(or 2)-methyl-3(2H)-furanone(HEMF) with DNA breaking activity in soy sauce. *Food and Chemical Toxicology*, *36*, 305–314.
Li, Xiaodong, Staszewski, L., Xu, H., Durick, K., Zoller, M., & Adler, E. (2002). Human receptors for sweet and umami taste. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(7), 4692–4696.
Lin, M., Liu, X., Xu, Q., Song, H., Li, P., & Yao, J. (2014). Aroma-active components of yeast extract pastes with a basic and characteristic meaty flavour. *Journal of the Science of Food and Agriculture*, *94*(5), 882–889.
Lin, S. Y., Lo, Y. C., Chen, Y. K., Yang, S. C., & Mau, J. L. (2015a). Nonvolatile Taste Components and Functional Compounds of Commercial Soy Sauce Products. *Journal of Food Processing and Preservation*, *39*(6), 2680–2686.
Lin, S. Y., Lo, Y. C., Chen, Y. K., Yang, S. C., & Mau, J. L. (2015b). Nonvolatile taste components and functional compounds of commercial soy sauce products. *Journal of Food Processing and Preservation*, *39*(6), 2680–2686.
Lindinger, C., Pollien, P., De Vos, R. C. H., Tikunov, Y., Hageman, J. A., Lambot, C., et al. (2009). Identification of ethyl formate as a quality marker of the fermented off-note in coffee by a nontargeted chemometric approach. *Journal of Agricultural and Food Chemistry*, *57*(21), 9972–9978.
Lioe, H. N. (2007). Chemical and sensory characteristics of low molecular weight fractions obtained from three types of Japanese soy sauce (shoyu) – Koikuchi , tamari and shiro shoyu. *Food Chemistry*, *100*, 1669–1677.
Lioe, H. N., Selamat, J., & Yasuda, M. (2010). Soy sauce and its umami taste: A link from the past to current situation. *Journal of Food Science*, *75*(3).
Liu, Mengru, Li, H., & Zhan, H. (2014). A novel method for the determination of the ethanol content in soy sauce by full evaporation headspace Gas Chromatography. *Food Analytical Methods*, *7*(5), 1043–1046.
Liu, Miao, Wang, M., Wang, J., & Li, D. (2013). Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar. *Sensors and Actuators, B: Chemical*, *177*, 970–980.
Lommen, A. (2009). MetAlign: Interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. *Analytical Chemistry*, *81*(8), 3079–3086.
Lopez-Sanchez, P., de Vos, R. C. H., Jonker, H. H., Mumm, R., Hall, R. D., Bialek, L., et al. (2015). Comprehensive metabolomics to evaluate the impact of industrial processing on the phytochemical composition of vegetable purees. *Food Chemistry*, *168*, 348–355.
Lubes, G., & Goodarzi, M. (2017). Analysis of volatile compounds by advanced analytical techniques and multivariate chemometrics. *Chemical Reviews*, *117*(9), 6399–6422.
Luh, B. S. (1995). Industrial production of soy sauce. *Journal of industrial Microbiology*, *14*, 467–471.
Maarse, H., & Visscher, C. A. (1989). *Volatile compounds in food: qualitative and quantitative data* (6th ed.). Zeist: TNO-CIVO Food Analysis Institute.
Macnamara, K., Ochiai, N., Sasamoto, K., & Hoffmann, A. (2018). A combination of standard (SBSE) and solvent-assisted (SA-SBSE) stir bar sorptive extraction for comprehensive analysis of flavor compounds in beverages. *Gerstel*, (198).
Mahadevan, K., & Farmer, L. (2006). Key odor impact compounds in three yeast extract pastes. *Journal of Agricultural and Food Chemistry*, *54*(19), 7242–7250.
Maillard, L. C. (1912). Action des acides amines sur les sucres: Formation des melanoidines par voie methodique. *Comptes Rendus de l'Académie des Sciences*, *154*, 66–68.
Malherbe, S., Menichelli, E., du Toit, M., Tredoux, A., Muller, N., Næs, T., & Nieuwoudt, H. (2013). The relationships between consumer liking, sensory and chemical attributes of Vitis vinifera L. cv. Pinotage wines elaborated with different Oenococcus oeni starter cultures. *Journal of the Science of Food and Agriculture*, *93*(11), 2829–2840.
Mamedov, V. A. (2016). Quinoxaline–As a Parent Heterocycle. In *Quinoxalines: Synthesis, Reactions, Mechanisms and Structure* (pp. 1–437). Springer.
Manzocco, L., Mastrocola, D., Nicoli, M. C., & Marangoni, V. (2001). Review of non- enzymatic browning and antioxidant capacity in processed foods. *Trends in Food Science Technology*, *11*(9–10), 340–346.
Marshall, D. D., & Powers, R. (2017). Beyond the paradigm: Combining mass spectrometry and nuclear magnetic resonance for metabolomics. *Progress in Nuclear Magnetic Resonance Spectroscopy*, *100*, 1–16.
Maruyama, Y., Yasuda, R., Kuroda, M., & Eto, Y. (2012). Kokumi substances, enhancers of basic tastes, induce responses in calcium-sensing receptor expressing taste cells. *PLoS ONE*, *7*(4), 1–8.
Matros, A., & Mock, H.-P. (2013). Mass spectrometry based imaging techniques for spatially resolved analysis of molecules. *Frontiers in Plant Science*, *4*(89), 1–7.
McGorrin, R. J. (2011). The Significance of Volatile Sulfur Compounds in Food Flavors. In *Volatile Sulfur Compounds in Food* (Vol. 1068, pp. 1–3). American Chemical Society.
McKay, M., Bauer, F. F., Panzeri, V., Mokwena, L., & Buica, A. (2019). Profiling potentially smoke tainted red wines: Volatile phenols and aroma attributes. *South African Journal of Enology and Viticulture*, *40*(2), 1–16.

Mebazaa, R., Mahmoudi, A., Fouchet, M., Santos, M. Dos, Kamissoko, F., Nafti, A., et al. (2009). Characterisation of volatile compounds in Tunisian fenugreek seeds. *Food Chemistry*, *115*(4), 1326–1336.

Miyamura, N., & Kuroda, M. (2015). Distribution of a kokumi peptide, γ-Glu-Val-Gly, in various fermented foods and the possibility of its contribution to the sensory quality of fermented foods. *Fermentation Technology*, *04*(02), 3–5.

Montaño, A., Sánchez, A. H., & Rejano, L. (1995). Determination of benzoic and sorbic acids in packaged vegetable products. Comparative evaluation of methods. *The Analyst*, *120*(10), 2483–2487.

Morton, I. D., Akroyd, P., & May, C. G. (1960). Flavoring substances. U.S. Patent 2934437. *Chemical Abstracts*, *54*, 17746a.

Mottram, D. S., Wedzicha, B. L., & Dodson, A. T. (2002). Acrylamide is formed in the Maillard reaction. *Nature*, *419*, 448.

Müller, R., & Rappert, S. (2010). Pyrazines: Occurrence, formation and biodegradation. *Applied Microbiology and Biotechnology*, *85*(5), 1315–1320.

Mumm, R., Hageman, J. A., Calingacion, M. N., de Vos, R. C. H., Jonker, H. H., Erban, A., et al. (2016). Multi-platform metabolomics analyses of a broad collection of fragrant and non-fragrant rice varieties reveals the high complexity of grain quality characteristics. *Metabolomics*, *12*(2), 38.

Munafo, J. P., Didzbalis, J., Schnell, R. J., Schieberle, P., & Steinhaus, M. (2014). Characterization of the major aroma-active compounds in mango (Mangifera indica L.) cultivars haden, white alfonso, praya sowoy, royal special, and malindi by application of a comparative aroma extract dilution analysis. *Journal of Agricultural and Food Chemistry*, *62*(20), 4544–4551.

Münch, P., Hofmann, T., & Schieberle, P. (1997). Comparison of key odorants generated by thermal treatment of commercial and self-prepared yeast extracts: Influence of the amino acid composition on odorant formation. *Journal of Agricultural and Food Chemistry*, *45*(4), 1338–1344.

Murray, R. A. (2001). Limitation to the use of solid-phase microextraction for quantitation of mixtures of volatile organic sulfur compounds. *Analytical Chemistry*, *73*(7), 1646–1649.

Naknean, P., & Meenune, M. (2010). Factors affecting retention and release of flavour compounds in food carbohydrates. *International Food Research Journal*, *17*(1), 23–34.

Newton, A. E., Fairbanks, A. J., Golding, M., Andrewes, P., & Gerrard, J. A. (2012). The role of the Maillard reaction in the formation of flavour compounds in dairy products – not only a deleterious reaction but also a rich source of flavour compounds. *Food & Function*, *3*(12), 1231.

Nieva-Echevarría, B., Goicoechea, E., & Guillén, M. D. (2017). Behaviour of non-oxidized and oxidized flaxseed oils, as models of omega-3 rich lipids, during in vitro digestion. Occurrence of epoxidation reactions. *Food Research International*, *97*, 104–115.

Nishimura, T., Goto, S., Miura, K., Takakura, Y., Egusa, A. S., & Wakabayashi, H. (2016). Umami compounds enhance the intensity of retronasal sensation of aromas from model chicken soups. *Food Chemistry*, *196*, 577–583.

Nishimura, T., & Kuroda, M. (2019). *Koku in Food Science and Physiology*. *Koku in Food Science and Physiology*. Springer.

Noda, F., Hayashi, K., & Mizunuma, T. (1980). Antagonism between osmophilic lactic acid bacteria and yeasts in brine fermentation of soy sauce. *Applied and Environmental Microbiology*, *40*(3), 452–457.

Nogueira, J. M. F. (2015). Stir-bar sorptive extraction: 15 years making sample preparation more environment-friendly. *TiAC Trends in Analytical Chemistry*, *71*, 214–223.

Nunomura, N., Sasaki, M., Asao, Y., & Yokotsuka, T. (1976). Identification of volatile components in Shoyu (soy sauce) by Gas Chromatography-Mass Spectrometry. *Agricultural and Biological Chemistry*, *40*(3), 485–490.

Nursten, H. (2005). *The Maillard Reaction: Chemistry, Biochemistry and Implications*. (Royal Society of Chemistry, Ed.). Cambridge, UK.

Ochiai, N., Sasamoto, K., David, F., & Sandra, P. (2018). Recent developments of stir bar sorptive extraction for food applications: Extension to polar solutes. *Journal of Agricultural and Food Chemistry*, *66*(28), 7249–7255. review-article.

Ochiai, N., Sasamoto, K., Hoffmann, A., & Okanoya, K. (2012). Full evaporation dynamic headspace and gas chromatography-mass spectrometry for uniform enrichment of odor compounds in aqueous samples. *Journal of Chromatography A*, *1240*, 59–68.

Ochiai, N., Tsunokawa, J., Sasamoto, K., & Hoffmann, A. (2014). Multi-volatile method for aroma analysis using sequential dynamic headspace sampling with an application to brewed coffee. *Journal of Chromatography A*, *1371*, 65–73.

Ohyama, T. (2013). *Traditional and modern Japanese soy foods : manufacturing, nutrition and cuisine of a variety of soy foods for health and joy of taste*. *Food Science and Technology TA - TT* -. Hauppauge, New York SE: Nova Science Publishers, Inc.

Omar, S. H., & Al-Wabel, N. A. (2010). Organosulfur compounds and possible mechanism of garlic in cancer. *Saudi Pharmaceutical Journal*, *18*(1), 51–58.

Onuma, T., Maruyama, H., & Sakai, N. (2018). Enhancement of saltiness perception by monosodium glutamate taste and soy sauce odor: A near-infrared spectroscopy study. *Chemical Senses*, *43*(3), 151–167.

Ottinger, H., & Hofmann, T. (2003). Identification of the taste enhancer alapyridaine in beef broth and evaluation of its sensory impact by taste reconstitution experiments. *Journal of Agricultural and Food Chemistry*, *51*(23), 6791–6796.

Parker, J. K. (2015a). *Thermal generation or aroma*. *Flavour Development, Analysis and Perception in Food and Beverages*. Elsevier Ltd.

Parker, J. K. (2015b). *Introduction to aroma compounds in foods*. *Flavour Development, Analysis and Perception in Food*

and Beverages. Elsevier Ltd.
Parliament, T. H. (1997). Solvent Extraction and Distillation Techniques. In R. Marsili (Ed.), *Flavor, fragrance, and odor analysis* (pp. 1–26). New York: Taylor & Francis.
Patana-Anake, P. (2015). Effect of temperature, pH, and food additives on tomato product volatile levels. *International Food Research Journal*, *22*(2), 561–571.
Patti, G. J., Yanes, O., & Siuzdak, G. (2012). Innovation: Metabolomics: the apogee of the omics trilogy. *Nat Rev Mol Cell Biol*, *13*(4), 263–269.
Perestrelo, R., Nogueira, J. M. F., & Câmara, J. S. (2009). Potentialities of two solventless extraction approaches-Stir bar sorptive extraction and headspace solid-phase microextraction for determination of higher alcohol acetates, isoamyl esters and ethyl esters in wines. *Talanta*, *80*(2), 622–630.
Petretto, G. L., Urgeghe, P. P., Mascia, I., Fadda, C., Rourke, J. P., & Pintore, G. (2017). Stir bar sorptive extraction coupled with GC/MS applied to honey: optimization of method and comparative study with headspace extraction techniques. *European Food Research and Technology*, *243*(5), 735–741.
Pham, T. T., Guichard, E., Schlich, P., & Charpentiei4, C. (1995). Optimal conditions for the formation of sotolon from a-ketobutyric acid in the french "Vin Jaune." *Journal of Agricultural and Food Chemistry*, *43*, 2616–2619.
Phillips, M., & Goss, M. J. (1932). Chemistry of lignin VII. Distillation of alkali lignin in reduced atmosphere of carbon dioxide. *Industrial and Engineering Chemistry*, *24*(12), 1436–1441.
Picard, M., Franc, C., de Revel, G., & Marchand, S. (2018). 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. *Analytica Chimica Acta*, *1001*, 168–178.
Pico, J., Gómez, M., Bernal, J., & Bernal, J. L. (2016). Analytical methods for volatile compounds in wheat bread. *Journal of Chromatography A*, *1428*, 55–71.
Piraprez, G., Hérent, M. F., & Collin, S. (1998). Flavour retention by lipids measured in a fresh cheese matrix. *Food Chemistry*, *61*(1–2), 119–125.
Poehlmann, S., & Schieberle, P. (2013). Characterization of the aroma signature of Styrian pumpkin seed oil (Cucurbita pepo subsp. pepo var. Styriaca) by molecular sensory science. *Journal of Agricultural and Food Chemistry*, *61*(12), 2933–2942.
Poinot, P., Arvisenet, G., Ledauphin, J., Gaillard, J. L., & Prost, C. (2013). How can aroma-related cross-modal interactions be analysed? A review of current methodologies. *Food Quality and Preference*, *28*(1), 304–316.
Pripis-Nicolau, L., De Revel, G., Bertrand, A., & Maujean, A. (2000). Formation of flavor components by the reaction of amino acid and carbonyl compounds in mild conditions. *Journal of Agricultural and Food Chemistry*, *48*(9), 3761–3766.
Procida, G., Cichelli, A., Lagazio, C., & Conte, L. S. (2016). Relationships between volatile compounds and sensory characteristics in virgin olive oil by analytical and chemometric approaches. *Journal of the Science of Food and Agriculture*, *96*(1), 311–318.
Qin, L., & Ding, X. (2007). Formation of taste and odor compounds during preparation of Douchiba, a Chinese traditional soy-fermented appetizer. *Journal of Food Biochemistry*, *31*(2), 230–251.
Rajalahti, T., Arneberg, R., Berven, F. S., Myhr, K. M., Ulvik, R. J., & Kvalheim, O. M. (2009). Biomarker discovery in mass spectral profiles by means of selectivity ratio plot. *Chemometrics and Intelligent Laboratory Systems*, *95*(1), 35–48.
Raynie, D. E. (2006). Modern extraction techniques, *78*, 3997–4004.
Raza, A., Begum, N., Song, H., Li, K., & Li, P. (2019). Optimization of headspace solid-phase microextraction (HS-SPME) parameters for the analysis of pyrazines in yeast extract via Gas Chromatography Mass Spectrometry (GC-MS). *Journal of Food Science*, *84*(8), 2031–2041.
Richter, T. M., Eyres, G. T., Silcock, P., & Bremer, P. J. (2017). Comparison of four extraction methods for analysis of volatile hop-derived aroma compounds in beer. *Journal of Separation Science*, *40*(22), 4366–4376.
Ridgway, K. (2015). *Analysis of taints and off-flavours*. *Flavour Development, Analysis and Perception in Food and Beverages*. Elsevier Ltd.
Riedel, K., Sombroek, D., Fiedler, B., Siems, K., & Krohn, M. (2017). Human cell-based taste perception-a bittersweet job for industry. *Natural Product Reports*, *34*(5), 484–495.
Rizzi, G. P. (2008). The Strecker degradation of amino acids: Newer avenues for flavor formation. *Food Reviews International*, *24*(4), 416–435.
Rohart, F., Eslami, A., Matigian, N., Bougeard, S., & Lê Cao, K. A. (2017). MINT: A multivariate integrative method to identify reproducible molecular signatures across independent experiments and platforms. *BMC Bioinformatics*, *18*(1), 1–13.
Romano, R., Brockhoff, P. B., Hersleth, M., Tomic, O., & Næs, T. (2008). Correcting for different use of the scale and the need for further analysis of individual differences in sensory analysis. *Food Quality and Preference*, *19*(2), 197–209.
Ross, C. F. (2009). Sensory science at the human-machine interface. *Trends in Food Science and Technology*, *20*(2), 63–72.
Rubert, J., Zachariasova, M., & Hajslova, J. (2015). Advances in high-resolution mass spectrometry based on metabolomics studies for food – a review. *Food Additives and Contaminants - Part A*, *32*(10), 1685–1708.
Sano, A., Satoh, T., Oguma, T., Nakatoh, A., Satoh, J. ichi, & Ohgawara, T. (2007). Determination of levulinic acid in soy sauce by liquid chromatography with mass spectrometric detection. *Food Chemistry*, *105*(3), 1242–1247.
Scalbert, A., Brennan, L., Fiehn, O., Hankemeier, T., Kristal, B. S., van Ommen, B., et al. (2009). Mass-spectrometry-based metabolomics: Limitations and recommendations for future progress with particular focus on nutrition

research. *Metabolomics*, *5*(4), 435–458.
Schaich, K. M. (2012). Thinking outside the classical chain reaction box of lipid oxidation. *Lipid Technology*, *24*(3), 55–58.
Sebald, K., Dunkel, A., Schäfer, J., Hinrichs, J., & Hofmann, T. (2018). Sensoproteomics: A new approach for the identification of taste-active peptides in fermented foods. *Journal of Agricultural and Food Chemistry*, *66*(42), 11092–11104.
Seisonen, S., Vene, K., & Koppel, K. (2016). The current practice in the application of chemometrics for correlation of sensory and gas chromatographic data. *Food Chemistry*, *210*, 530–540.
Seok, Y. J., Her, J. Y., Kim, Y. G., Kim, M. Y., Jeong, S. Y., Kim, M. K., et al. (2015). Furan in thermally processed foods - A review. *Toxicological Research*, *31*(3), 241–253.
Shahidi, F., Samaranayaka, A. G. P., & Pegg, R. B. (2014). *Maillard Reaction and Browning*. *Encyclopedia of Meat Sciences* (Vol. 1). Elsevier Ltd.
Shahidi, Fereidoon, Rubin, L. J., & D'Souza, L. A. (1986). Meat flavor volatiles: A review of the composition, techniques of analysis, and sensory evaluation. *C R C Critical Reviews in Food Science and Nutrition*, *24*(2), 141–243.
Shepherd, G. M. (2005). Outline of a theory of olfactory processing and its relevance to humans. *Chemical Senses*, *30*, 3–5.
Sheskin, D. J. (2003). Parametric and non parametric statistical procedures. *Handbook of Parametric and Nonparametric Statistical Procedures:*, 1–1193.
Shi, L., Westerhuis, J. A., Rosén, J., Landberg, R., & Brunius, C. (2019). Variable selection and validation in multivariate modelling. *Bioinformatics*, *35*(6), 972–980.
Shiga, K., Yamamoto, S., Nakajima, A., Kodama, Y., Imamura, M., Sato, T., et al. (2014). Metabolic profiling approach to explore compounds related to the umami intensity of soy sauce. *Journal of Agricultural and Food Chemistry*, *62*(29), 7317–7322.
Shim, J., Son, H. J., Kim, Y., Kim, K. H., Kim, J. T., Moon, H., et al. (2015). Modulation of sweet taste by umami compounds via sweet taste receptor. *PLoS ONE*, *10*(4), 1–9.
Shurtleff, W., & Aoyagi, A. (2012). *History of soy sauce (160 CE to 2012)* (1st ed.). Lafayette: Soyinfo Center.
Siskos, A. P., Jain, P., Römisch-Margl, W., Bennett, M., Achaintre, D., Asad, Y., et al. (2017). Interlaboratory reproducibility of a targeted metabolomics platform for analysis of human serum and plasma. *Analytical Chemistry*, *89*(1), 656–665.
Smit, B. A., Engels, W. J. M., & Smit, G. (2009). Branched chain aldehydes: Production and breakdown pathways and relevance for flavour in foods. *Applied Microbiology and Biotechnology*, *81*(6), 987–999.
Soares, S., Silva, M. S., García-Estevez, I., Groβmann, P., Brás, N., Brandão, E., et al. (2018). Human bitter taste receptors are activated by different classes of polyphenols. *Journal of Agricultural and Food Chemistry*, *66*(33), 8814–8823.
Song, Y.-R., Jeong, D.-Y., & Baik, S.-H. (2015a). Effects of indigenous yeasts on physicochemical and microbial properties of Korean soy sauce prepared by low-salt fermentation. *Food Microbiology*, *51*, 171–178.
Song, Y.-R., Jeong, D.-Y., & Baik, S.-H. (2015b). Monitoring of yeast communities and volatile flavor changes during traditional Korean soy sauce fermentation. *Journal of Food Science*, *80*(9), M2005–M2014.
Spence, C. (2015). Just how much of what we taste derives from the sense of smell? *Flavour*, *4*(1), 1–10.
Steinhaus, P., & Schieberle, P. (2007). Characterization of the key aroma compounds in soy sauce using approaches of molecular sensory science. *Journal of Agricultural and Food Chemistry*, *55*(15), 6262–6269.
Steinkraus, H. K. (2002). Fermentations in world food processing. *Comprehensive reviews in food science and food safety*, *1*(1), 23–32.
Stolle, T., Grondinger, F., Dunkel, A., & Hofmann, T. (2018). Quantitative proteomics and SWATH-MS to elucidate peri-receptor mechanisms in human salt taste sensitivity. *Food Chemistry*, *254*, 95–102.
Sucan, M. K., & Weerasinghe, D. K. (2005). Process and reaction flavors: an overview. *Process and Reaction Flavors*, *905*, 1–23.
Sugimoto, M., Koseki, T., Hirayama, A., Abe, S., Sano, T., Tomita, M., & Soga, T. (2010). Correlation between sensory evaluation scores of Japanese sake and metabolome profiles. *Journal of Agricultural and Food Chemistry*, *58*(1), 374–383.
Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., et al. (2007). Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics*, *3*(3), 211–221.
Sun, H. M., Wang, J. Z., Zhang, C. H., Li, X., Xu, X., Dong, X. B., et al. (2014). Changes of flavor compounds of hydrolyzed chicken bone extracts during maillard reaction. *Journal of Food Science*, *79*(12), C2415–C2426.
Sun, S. Y., Jiang, W. G., & Zhao, Y. P. (2010). Profile of volatile compounds in 12 Chinese soy sauces produced by a high-salt-diluted state fermentation. *Journal of the Institute of Brewing*, *116*(3), 316–328.
Sussulini, A. (2017). Metabolomics : from fundamentals to clinical applications. Cham, Switzerland: Springer.
Szeremeta, M., Pietrowska, K., Niemcunowicz-Janica, A., Kretowski, A., & Ciborowski, M. (2021). Applications of metabolomics in forensic toxicology and forensic medicine. *International Journal of Molecular Sciences*, *22*(6), 1–16.
Tamanna, N., & Mahmood, N. (2015). Food processing and Maillard reaction products: Effect on human health and nutrition. *International Journal of Food Science & Technology*, *2015*, 1–6.
Tareke, E., Rydberg, P., Karlsson, P., Eriksson, S., & Törnqvist, M. (2000). Acrylamide: A cooking carcinogen? *Chemical Research in Toxicology*, *13*(6), 517–522.
Thissen, U., Coulier, L., Overkamp, K. M., Jetten, J., van der Werff, B. J. C., van de Ven, T., & van der Werf, M. J. (2011). A proper metabolomics strategy supports efficient food quality improvement: A case study on tomato sensory

properties. *Food Quality and Preference*, *22*(6), 499–506.
Thomsen, B. R., Yesiltas, B., Sørensen, A.-D. M., Hermund, D. B., Glastrup, J., & Jacobsen, C. (2016). Comparison of three methods for extraction of volatile lipid oxidation products from food matrices for GC–MS analysis. *Journal of the American Oil Chemists' Society*, *93*(7), 929–942.
Tikunov, Y. M., Laptenok, S., Hall, R. D., Bovy, A., & de Vos, R. C. H. (2012). MSClust: a tool for unsupervised mass spectra extraction of chromatography-mass spectrometry ion-wise aligned data. *Metabolomics*, *8*(4), 714–718.
Tikunov, Y. M., Molthoff, J., de Vos, R. C. H., Beekwilder, J., van Houwelingen, A., van der Hooft, J. J. J., et al. (2013). NON-SMOKY GLYCOSYLTRANSFERASE1 prevents the release of smoky aroma from tomato fruit. *The Plant Cell*, *25*(8), 3067–3078.
Tikunov, Yury M., Roohanitaziani, R., Meijer-Dekens, F., Molthoff, J., Paulo, J., Finkers, R., et al. (2020). The genetic and functional analysis of flavor in commercial tomato: the FLORAL4 gene underlies a QTL for floral aroma volatiles in tomato fruit. *Plant Journal*, *103*(3), 1189–1204.
Tocmo, R., Wu, Y., Liang, D., Fogliano, V., & Huang, D. (2017). Boiling enriches the linear polysulfides and the hydrogen sulfide-releasing activity of garlic. *Food Chemistry*, *221*, 1867–1873.
Toelstede, S., Dunkel, A., & Hofmann, T. (2009). A series of kokumi peptides impart the long-lasting mouthfulness of matured gouda cheese. *Journal of Agricultural and Food Chemistry*, *57*(4), 1440–1448.
U.S. Environmental Protection Agency. CompTox Chemicals Dashboard. (n.d.).
Ueda, Y., Sakaguchi, M., Hirayama, K., Miyajima, R., & Kimizuka, A. (1990). Characteristic flavor constituents in water extract of garlic. *Agricultural and Biological Chemistry*, *54*(1), 163–169.
Van Boekel, M. A. J. S. (2006). Formation of flavour compounds in the Maillard reaction. *Biotechnology Advances*, *24*(2), 230–233.
van Boekel, M., Fogliano, V., Pellegrini, N., Stanton, C., Scholz, G., Lalljie, S., et al. (2010). A review on the beneficial aspects of food processing. *Molecular Nutrition and Food Research*, *54*(9), 1215–1247.
Van Duynhoven, J. P. M., & Jacobs, D. M. (2016). Assessment of dietary exposure and effect in humans: The role of NMR. *Progress in Nuclear Magnetic Resonance Spectroscopy*, *96*, 58–72.
Vidal, V. A. S., Santana, J. B., Paglarini, C. S., da Silva, M. A. A. P., Freitas, M. Q., Esmerino, E. A., et al. (2020). Adding lysine and yeast extract improves sensory properties of low sodium salted meat. *Meat Science*, *159*, 107911.
Vilgis, T., & Vierich, T. (2020). *Aroma essenziell: Die praktische Kunst des Würzen*. Stiftung Warentest.
Vinci, R. M., Mestdagh, F., Van Poucke, C., Van Peteghem, C., & De Meulenaer, B. (2012). A two-year investigation towards an effective quality control of incoming potatoes as an acrylamide mitigation strategy in french fries. *Food Additives and Contaminants*, *29*(3), 362–370.
Visessanguan, W., Benjakul, S., Riebroy, S., Yarchai, M., & Tapingkae, W. (2006). Changes in lipid composition and fatty acid profile of Nham, a Thai fermented pork sausage, during fermentation. *Food Chemistry*, *94*(4), 580–588.
Wanakhachornkrai, P., & Lertsiri, S. (2003). Comparison of determination method for volatile compounds in Thai soy sauce. *Food Chemistry*, *83*, 619–629.
Wang, X., Guo, M., Song, H., & Meng, Q. (2020). Characterization of key aroma compounds in traditional Chinese soy sauce through the molecular sensory science technique. *Lwt*, *128*, 109413.
Wang, Z., Xiao, Q., Zhuang, J., Feng, T., Ho, C. T., & Song, S. (2020). Characterization of aroma-active compounds in four yeast extracts using instrumental and sensory techniques. *Journal of Agricultural and Food Chemistry*, *68*(1), 267–278.
Wei, Q., Wang, H., Chen, Z., Lv, Z., Xie, Y., & Lu, F. (2013). Profiling of dynamic changes in the microbial community during the soy sauce fermentation process. *Applied Microbiology and Biotechnology*, *97*(20), 9111–9119.
Wei, Q., Wang, H., Lv, Z., Hu, G., Li, Y., Liu, Y., et al. (2013). Search for potential molecular indices for the fermentation progress of soy sauce through dynamic changes of volatile compounds. *Food Research International*, *53*(1), 189–194.
Welzenbach, J., Neuhoff, C., Looft, C., Schellander, K., Tholen, E., & Große-Brinkhaus, C. (2016). Different statistical approaches to investigate porcine muscle metabolome profiles to highlight new biomarkers for pork quality assessment. *PLoS ONE*, *11*(2), 1–24.
Werkhoff, P., Brennecke, S., Bretschneider, W., & Bertram, H.-J. (2001). Modern methods for isolating and quantifying volatile flavor and fragrance compounds. *Flavor, Fragrance and Odor Analysis*, 139–204.
Whitfield, F. B. (1992). Volatiles from interactions of Maillard reactions and lipids. *Critical Reviews in food Science and Nutrition*, *31*, 1–58.
Wishart, D. S. (2008). Metabolomics: applications to food science and nutrition research. *Trends in Food Science and Technology*, *19*(9), 482–493.
Xia, W., & Budge, S. M. (2017). Techniques for the analysis of minor lipid oxidation products derived from triacylglycerols: Epoxides, alcohols, and ketones. *Comprehensive Reviews in Food Science and Food Safety*, *16*(4), 735–758.
Xu, Y., Fan, W., & Qian, M. C. (2007). Characterization of aroma compounds in apple cider using solvent-assisted flavor evaporation and headspace solid-phase microextraction. *Journal of Agricultural and Food Chemistry*, *55*(8), 3051–3057.
Yanfang, Z., & Wenyi, T. (2009). Flavor and taste compounds analysis in Chinese solid fermented soy sauce. *African Journal of Biotechnology*, *8*(4), 673–681.
Yang, Juan, Bai, W., Zeng, X., & Cui, C. (2019). Gamma glutamyl peptides: The food source, enzymatic synthesis, kokumi-active and the potential functional properties – A review. *Trends in Food Science & Technology*, *91*, 339–346.
Yang, Jun, Zhao, X., Lu, X., Lin, X., & Xu, G. (2015). A data preprocessing strategy for metabolomics to reduce the mask

effect in data analysis. *Frontiers in molecular biosciences*, *2*(4), 1–9.
Yang, X., Zhou, X., Zhang, M., Zhang, Z., Song, L., Wang, G., & Zhang, J. (2020). Metabolism analysis for enhanced nutritional profile of chestnuts subjected to anerobic solid-state fermentation by probiotic lactic acid bacteria. *Journal of Food Processing and Preservation*, *44*(3), 1–11.
Yang, Y., Deng, Y., Jin, Y., Liu, Y., Xia, B., & Sun, Q. (2016). Dynamics of microbial community during the extremely long-term fermentation process of a traditional soy sauce: Microbial community in traditional soy sauce fermentation. *Journal of the Science of Food and Agriculture*, *97*.
Yaylayan, V. A., & Mandeville, S. (1994). Stereochemical control of maltol formation in Maillard reaction. *Journal of Agricultural and Food Chemistry*, *42*, 771–775.
Yokotsuka, T., & Sasaki, M. (1998). Fermented protein foods in the Orient: shoyu and miso in Japan. In B. J. B. Wood (Ed.), *Microbiology of Fermented Foods* (2nd ed., pp. 351–415). Boston: Springer.
Yu, A.-N., Tan, Z.-W., & Shi, B.-A. (2011). Influence of the pH on the formation of pyrazine compounds by the Maillard reaction of L-ascorbic acid with acidic, basic and neutral amino acids. *Asia-Pac. J. Chem. Eng.*, (7), 455–462.
Yu, A. N., & Zhang, A. D. (2010). Aroma compounds generated from thermal reaction of l-ascorbic acid with l-cysteine. *Food Chemistry*, *121*(4), 1060–1065.
Zhang, A., Sun, H., Wang, P., Han, Y., & Wang, X. (2012). Modern analytical techniques in metabolomics analysis. *The Analyst*, *137*(2), 293–300.
Zhang, Y., Song, H., Li, P., Yao, J., & Xiong, J. (2017). Determination of potential off-flavour in yeast extract. *LWT - Food Science and Technology*, *82*, 184–191. https://doi.org/10.1016/j.lwt.2017.04.030
Zhao, C. J., Schieber, A., & Gänzle, M. G. (2016). Formation of taste-active amino acids, amino acid derivatives and peptides in food fermentations – A review. *Food Research International*, *89*, 39–47.
Zhao, G. Q., Zhang, Y., Hoon, M. A., Chandrashekar, J., Erlenbach, I., Ryba, N. J. P., & Zuker, C. S. (2003). The receptors for mammalian sweet and umami taste. *Cell*, *115*(3), 255–266.
Zheng, J., Wu, C. De, Huang, J., Zhou, R. Q., & Liao, X. P. (2013). Analysis of volatile compounds in Chinese soy sauces moromi cultured by different fermentation processes. *Food Science and Biotechnology*, *22*(3), 605–612.
Zhu, W., & Cadwallader, K. R. (2019). Streamlined approach for careful and exhaustive aroma characterization of aged distilled liquors. *Food Chemistry: X*, *3*, 100038.
Zhuang, M., Lin, L., Zhao, M., Dong, Y., & Sun-waterhouse, D. (2016). Sequence, taste and umami-enhancing effect of the peptides separated from soy sauce. *Food Chemistry*, *206*, 174–181.
Zielinski, A. A. F., Haminiuk, C. W. I., Nunes, C. A., Schnitzler, E., van Ruth, S. M., & Granato, D. (2014). Chemical composition, sensory properties, provenance, and bioactivity of fruit juices as assessed by chemometrics: A critical review and guideline. *Comprehensive Reviews in Food Science and Food Safety*, *13*(3), 300–316.
Zou, G., Xiao, Y., Wang, M., & Zhang, H. (2018). Detection of bitterness and astringency of green tea with different taste by electronic nose and tongue. *PLoS ONE*, *13*(12), 1–11.

Summary

Aroma is considered as the superior component of flavour in food. The molecular compounds responsible for aroma are the so-called volatiles. Volatiles are small molecules that exhibit high vapour pressure under ambient conditions and have low boiling points, thus, they are present in the air phase as well as in the liquid and solid phases. Many volatiles are naturally present in plants, animals, microorganisms and the environment, however they can also be formed by the action of (enzymatic or non-enzymatic) processing reactions. When food is processed or cooked, different volatiles can be formed and these may influence the final flavour of food. Various analytical techniques (including metabolomics approaches) are being used to analyse the composition of food samples in order to generate new knowledge of the volatile composition of food samples, which is of great relevance when designing and formulating new superior flavourful products for the market.

In this PhD thesis I explored several methodologies to extract volatiles from complex food matrices, such as in processed food products (e.g. yeast derivative ingredients and instant soups). The major aim of this thesis was to design novel approaches to be able to zoom in on the volatile content of savoury food ingredients to better understand which ingredients and processing strategies can lead to superior and desired flavour qualities.

Therefore, first I focused on performing two literature studies, which covered the background knowledge of the savoury ingredients used in this thesis: yeast derivative ingredients in Chapter 2 and soy sauces in Chapter 3. **Chapter 2** provides a comprehensive and up-to-date review of the major flavour compound classes described in processed savoury ingredients, including the formation of these compounds from their precursors. Special attention was given to the interconnections between Maillard reactions and the different amino acid, lipid, and carbohydrate degradation pathways. Furthermore, the chapter provides insights into advancing metabolomics applications that have not yet being exploited in depth for processed food ingredients.

As a result of the learning outcome from the first review, I proceeded with the reported literature on soy sauce composition. **Chapter 3** summarises our knowledge on the chemical compounds known to be present in the most common soy sauces found on the market, and their potential sensory relevance. This review also presents a sensory wheel of taste and aroma attributes that characterize soy sauce flavour. Soy sauce is a condiment used worldwide and is perhaps one of the most complex fermented condiments containing many small molecules that boost the flavour of many dishes. Here also little has yet been done in implementing comprehensive metabolomics analyses to study the metabolite composition of soy sauces.

Intrigued by the knowledge of the volatiles reported in the literature and the lack of untargeted metabolomics techniques used for these specific food ingredients, I focused **Chapter 4** on developing and comparing new methodologies that broaden the volatile spectrum as well as taking into consideration the 'unknown' compounds, which are also of great relevance in flavour studies. Four volatile trapping

techniques were optimized and compared (SBSE, SPME, HSSE and DHS), and the results showed that SBSE performed better in terms of coverage of volatiles, while SPME was the most repeatable technique of all.

The lack of implementation of new methodologies, such as SBSE techniques, in soy sauce analysis led to **Chapter 5**, which focuses on analysing the volatile content of different soy sauce products made using contrasting production procedures and flavour characteristics. Fermented and non-fermented soy sauces showed the greatest differences in terms of the volatile content, and specific classes of volatiles (pyrazines in non-fermented and esters in fermented soy sauces) could be linked to the different types of soy sauces analysed. These results lead to the proposition that volatile content, and thus aroma, are greatly influenced both by the process used to produce the final product as well as the ratio of raw ingredients added. This resulted in a list of potential biomarkers for specific soy sauce types found on the global market. The implementation of new methodologies sets the basis for future research in the application of metabolomics for soy sauce products.

As a result of observing contrasting volatile profiles when using different volatile trapping techniques, I focused **Chapter 6** on finding applicable tools/pipelines that help in deciding, in the early stages of large-scale experiments, which technique is best to use in a specific study. The workflow proposed showed that SPME is more repeatable than SBSE, however, SBSE was able to trap more volatiles than SPME and also that the additional (unique) volatiles that SBSE trapped increased the power of prediction for certain sensory attributes, such as for garlic flavour, as compared to SPME. Therefore, depending on what is the precise goal of a study, an informed decision can be taken on which technique is best to apply.

In order to confirm that volatiles measured by SBSE-GC-MS can be used to predict sensory characteristics in savoury products, results in **Chapter 7** show that sensory attributes typical of bouillon-type soups, such as roast odour or chicken flavour, could be associated with certain volatile classes (i.e. short-chain aldehydes with chicken flavour), and thus the ingredients and processing steps of the bouillon samples that contained those volatile compounds could be linked.

All these chapters combined have shown that this untargeted metabolomics, data-driven approach is valuable for 1) scientists to understand aroma formation and chemical changes which occur during processing and; 2) industry to formulate improved food products in a more targeted way.

Samenvatting

Aroma wordt beschouwd als de leidende component van smaak in voedsel. De moleculaire stoffen die verantwoordelijk zijn voor aroma zijn de zogeheten vluchtige stoffen. Onder vluchtige stoffen verstaat men kleine moleculen die onder normale temperatuur en druk een hoge dampspanning en een laag kookpunt hebben, waardoor ze zowel in de gasfase als in de vloeibare en vaste toestand voorkomen. Veel vluchtige stoffen zijn van nature aanwezig in planten, dieren, micro-organismen en het milieu, maar kunnen ook gevormd worden door (enzymatische of niet-enzymatische) omzettingen. Wanneer voedsel verwerkt of gekookt wordt, kunnen er verschillende vluchtige stoffen gevormd worden en deze kunnen de uiteindelijke smaak van voedsel beïnvloeden. Diverse analysetechnieken (waaronder die gebaseerd op metabolomics) worden gebruikt voor het analyseren van de samenstelling van voedselmonsters om zo nieuwe kennis te genereren over de vluchtige samenstelling ervan, hetgeen van groot belang is bij het vormgeven en formuleren van nieuwe superieure smaakvolle producten voor de markt.

In dit proefschrift heb ik verschillende methodieken onderzocht om vluchtige stoffen te extraheren uit complexe voedselmatrices, zoals bewerkte voedingsmiddelen (bijv. gistderivaten en instant soepen). Het hoofddoel van dit proefschrift was om nieuwe benaderingen te ontwikkelen om in te kunnen zoomen op het gehalte aan vluchtige bestanddelen in hartige voedingsingrediënten, teneinde beter te begrijpen welke ingrediënten en verwerkingsstrategieën kunnen bijdragen aan een betere en gewenste smaak.

Daarom heb ik me eerst geconcentreerd op het uitvoeren van twee literatuurstudies, gericht op de achtergrond van de hartige ingrediënten gebruikt in dit proefschrift: gistderivaten in Hoofdstuk 2 en sojasauzen in Hoofdstuk 3. **Hoofdstuk 2** biedt een uitgebreid en actueel overzicht van de belangrijkste groepen van smaakstoffen die vermeld worden in bewerkte hartige ingrediënten, evenals hoe deze verbindingen uit hun precursors gevormd worden. Er wordt bijzondere aandacht besteed aan het verband tussen Maillard reacties en de diverse reactiemechanismen voor de afbraak van aminozuren, vetten en koolhydraten. Verder biedt het hoofdstuk inzicht in geavanceerde toepassingen van metabolomics die nog niet optimaal worden benut voor bewerkte voedingsingrediënten.

Naar aanleiding van de uitkomst van het eerste review artikel, richtte ik mij tot de beschikbare literatuur over de samenstelling van sojasaus. **Hoofdstuk 3** bundelt onze kennis over chemische verbindingen waarvan bekend is dat ze aanwezig zijn in de meest voorkomende sojasauzen op de markt, en hun mogelijke sensorische relevantie. Deze review presenteert tevens een sensorisch wiel met smaak- en aromakwaliteiten die de smaak van sojasaus kenmerken. Sojasaus wordt wereldwijd gebruikt als smaakmaker en is misschien wel een van de meest complexe gefermenteerde sauzen, met veel kleine moleculen die smaak toevoegen aan tal van gerechten. Ook hier is er nog weinig gedaan om uitgebreide metabolomics analyses te implementeren voor het bestuderen van de metabolietsamenstelling van sojasauzen.

Geïntrigeerd door de kennis van vluchtige stoffen uit de literatuur en de beperkte inzet van untargeted metabolomics technieken voor deze specifieke voedingsingrediënten, heb ik mij in **Hoofdstuk 4** gericht op

het ontwikkelen en vergelijken van nieuwe methodieken die zowel het spectrum van vluchtige stoffen uitbreiden als rekening houden met "onbekende" verbindingen, die ook uiterst relevant zijn voor smaakonderzoeken. Er zijn vier technieken voor de extractie en concentratie van vluchtige stoffen geoptimaliseerd en vergeleken (SBSE, SPME, HSSE en DHS); uit de resultaten bleek dat SBSE beter presteerde qua dekking van vluchtige stoffen, terwijl SPME de meest reproduceerbare techniek was.

De gebrekkige toepassing van nieuwe methodieken, waaronder SBSE technieken, in de analyse van sojasaus vormde de aanleiding voor **Hoofdstuk 5**, waar de nadruk ligt op het analyseren van het gehalte aan vluchtige bestanddelen in verschillende sojasaus producten gemaakt met behulp van contrasterende productieprocedures en smaakkenmerken. Gefermenteerde en niet-gefermenteerde sojasauzen vertoonden de grootste verschillen op het gebied van hun gehalte aan vluchtige bestanddelen; specifieke groepen van vluchtige stoffen (pyrazines in niet-gefermenteerde en esters in gefermenteerde sojasauzen) konden gerelateerd worden aan de verschillende soorten sojasaus die waren geanalyseerd. Deze resultaten leidden tot het idee dat het gehalte aan vluchtige bestanddelen, en dus aroma, sterk wordt beïnvloed door zowel het proces dat gebruikt wordt om het eindproduct te produceren als de verhouding tussen de gebruikte grondstoffen. Dit resulteerde in een lijst met potentiële biomarkers voor specifieke soorten sojasaus die op de wereldmarkt beschikbaar zijn. De implementatie van nieuwe methodieken vormt de basis voor toekomstig onderzoek naar de toepassing van metabolomics voor sojasausproducten.

Naar aanleiding van het observeren van contrasterende vluchtige profielen bij het gebruik van verschillende technieken voor de extractie en concentratie van vluchtige stoffen, heb ik mij in **Hoofdstuk 6** gericht op het vinden van toepasbare handvatten/workflows die ondersteuning bieden bij het beslissen welke techniek het beste is om te gebruiken voor een specifiek onderzoek, in de beginfase van grootschalige experimenten. De voorgestelde workflow wees uit dat SPME reproduceerbaarder is dan SBSE. Echter was SBSE in staat om meer vluchtige componenten op te vangen, waarbij de additionele (unieke) componenten die SBSE opving de voorspellingskracht verhoogden voor bepaalde sensorische kenmerken, zoals voor knoflooksmaak, in vergelijking met SPME. Afhankelijk van wat het precieze doel van een onderzoek is, dient men dus een weloverwogen beslissing nemen over welke techniek het beste kan worden toegepast.

Om te bevestigen dat de vluchtige stoffen gemeten door SBSE-GC-MS gebruikt kunnen worden voor het voorspellen van sensorische eigenschappen van hartige producten, laten de resultaten in **Hoofdstuk 7** zien dat sensorische kenmerken van bouillon-achtige soepen, zoals braadgeur of kipsmaak, geassocieerd kunnen worden met bepaalde groepen van vluchtige stoffen (namelijk korte-keten aldehyden met de smaak van kip). Derhalve konden de ingrediënten en verwerkingsstappen van de bouillonmonsters die deze vluchtige verbindingen bevatten, gelinkt worden.

Alle hoofdstukken tezamen hebben aangetoond dat deze untargeted metabolomics, datagedreven benadering waardevol is voor 1) wetenschappers, om de vorming van aroma's en chemische veranderingen die optreden tijdens verwerking te begrijpen en; 2) industrie, om op een nog gerichtere manier betere voedingsmiddelen te ontwikkelen.

Resumen

El aroma de los alimentos es el componente superior del sabor, en perjuicio del otro componente, el gusto. Las moléculas responsables del aroma son los denominados volátiles. Los compuestos volátiles son moléculas pequeñas que exhiben alta presión de vapor en condiciones ambientales y tienen puntos de ebullición bajos, por lo que pueden presentarse en estado gaseoso, así como en estado líquido y sólido. Muchos de los volátiles se presentan naturalmente en plantas, animales, microorganismos y en el medio ambiente, sin embargo, también pueden formarse a partir de reacciones químicas (enzimáticas o no enzimáticas) durante el cocinado. Cuando los alimentos se procesan o cocinan, se forman un gran número de compuestos volátiles los cuales pueden llegar a influenciar el sabor final de los alimentos. Hoy en día, existen varias técnicas analíticas (incluyendo las técnicas metabolómicas) para analizar la composición química de los alimentos, con el fin de conocer más sobre la estructura y el tipo de volátiles (y no volátiles) presentes en las muestras. Estos compuestos son de gran importancia a la hora de diseñar y formular nuevos productos alimenticios con sabores y aromas superiores.

En esta tesis, he explorado diferentes técnicas para extraer los compuestos volátiles de alimentos con matriz compleja como, por ejemplo, ingredientes derivados de levadura, sopas instantáneas u otros alimentos procesados. El objetivo principal se ha basado en diseñar nuevas metodologías para ampliar el contenido de volátiles que se pueden analizar en una muestra, con la finalidad de entender cuáles son los ingredientes y procesos que llevan a obtener las mejores y más deseadas cualidades del sabor.

Para obtener resultados satisfactorios al objetivo, en primer lugar, me centré en el estudio de la literatura presente sobre los compuestos volátiles identificados en ingredientes salados: en derivados de levadura (Capítulo 2), y en la salsa de soja (Capítulo 3). En el **Capítulo 2**, he redactado una revisión literaria, exhaustiva y actualizada de las principales moléculas causantes del sabor característico de ingredientes salados, incluyendo las reacciones que llevan a la formación de estos compuestos sabrosos. Se describen con especial atención las interconexiones entre reacciones de Maillard y las reacciones de degradación de aminoácidos, lípidos y carbohidratos. Por otra parte, también se discute información sobre el avance de las aplicaciones en metabolómica, las cuales aún no se han explotado en profundidad para los ingredientes alimenticios procesados.

Como resultado del aprendizaje de redactar una revisión literaria, en segundo lugar procedí a estudiar la literatura científica publicada sobre la salsa de soja. El **Capítulo 3** de esta tesis resume la información recopilada sobre los compuestos químicos (aromáticos) conocidos hasta ahora de las salsas de soja más comunes que se encuentran en el mercado, así como su relevancia sensorial. Esta revisión literaria muestra una rueda sensorial que contiene atributos del aroma y el gusto característicos del sabor de la salsa de soja. Tal condimento, la salsa de soja, se usa globalmente y, siendo probablemente, uno de los condimentos fermentados más complejos que existen, conteniendo numerosos compuestos que dan aroma y elevan el sabor de muchos alimentos. A su vez he destacado el uso de técnicas metabolómicas exhaustivas para analizar la composición química de la salsa de soja.

Intrigada por la ciencia de los compuestos volátiles conocida hasta ahora, y la falta del uso de técnicas metabolómicas exhaustivas en alimentos procesados, el **Capítulo 4** está enfocado en el desarrollo y la comparación de nuevas metodologías que expanden el espectro de volátiles analizados, así como los compuestos no identificados, ya que son de igual importancia en estudios del sabor. Optimizamos y comparamos cuatro técnicas para extraer volátiles (SBSE, SPME, HSSE y DHS), y los resultados mostraron que la técnica SBSE funcionó mejor en términos de mayor cobertura de volátiles, mientras que SPME resultó ser la técnica más reproducible de todas.

La falta de implementación de nuevas metodologías, como la técnica SBSE en el análisis de salsas de soja, me llevó a redactar el **Capítulo 5**. Este Capítulo está enfocado en el análisis de los volátiles presentes en varios tipos de salsas de soja preparadas usando procesos de producción variados, y con sabores contrastados. Los resultados mostraron que la mayor diferencia en términos del contenido en volátiles se encontraba entre las salsas de soja fermentadas y no fermentadas. De tal modo que se pudieron vincular volátiles de clases específicas con las diferentes salsas de soja analizadas (pirazinas en salsas no fermentadas y esteres en salsas fermentadas). Estos resultados llevaron a proponer que el contenido en volátiles, así como el aroma final, están mayormente influenciados por los procesos de producción usados y por la proporción de los ingredientes crudos añadidos. Lo que me llevó a generar una lista de biomarcadores para salsas de soja típicas encontradas en el mercado global. La implementación de nuevas metodologías en esta tesis pone la base para futuras investigaciones en la aplicación de metabolómica para productos de salsa de soja.

Como resultado de haber observado perfiles de volátiles diferentes al hacer uso de diferentes técnicas para extraer volátiles, enfoqué el **Capítulo 6** en buscar herramientas automatizadas que ayudan a decidir, en las primeras etapas de experimentos a gran escala, qué técnica es la mejor para cada estudio específico. La logística propuesta mostró que SPME es la técnica más reproducible, sin embargo, SBSE pudo extraer mayor número de volátiles que SPME. Los volátiles adicionales que SBSE extrajo incrementaron la potencia de predicción de ciertos atributos sensoriales, como por ejemplo el sabor a ajo. En consecuencia, haciendo uso de estas herramientas podemos llegar a tomar una decisión automatizada e informativa para conocer que técnica es mejor aplicar, dependiendo de cuál es el objetivo principal del estudio.

Para poder confirmar qué volátiles se pueden usar para predecir características sensoriales en productos salados, usando SBSE-GC-MS, el **Capítulo 7** muestra que ciertos atributos sensoriales típicos de sopas instantáneas de caldo, como el aroma tostado y el sabor a pollo, se pudieron asociar con un tipo de clases de volátiles (por ejemplo, aldehídos de cadena corta con el sabor a pollo), y así se pudieron asociar ingredientes y etapas del proceso específicos que llevaron a la formación de ese tipo de volátiles.

Todos estos capítulos en conjunto han demostrado que las técnicas metabolómicas exhaustivas y las herramientas automatizadas desarrolladas son valiosos para: 1) Los científicos, ya que ayudan a entender las reacciones de formación del aroma y los cambios químicos que ocurren durante el procesamiento de alimentos; 2) la industria, permitiendo diseñar y formular productos procesados mejorados de una manera más directa y rápida.

Acknowledgements

When I look back at the last four and half years, it feels like a trip full of abundant knowledge, bizarre experiences, and (self-)awareness, yet that trip could not have been possible without ~~all~~ the wonderful people that I met all along.

First, I would like to thank Robert. Since the first time we met, about seven and half years ago, we have shared many experiences, both professional and personal, which I remember with joy, peacefulness, and wisdom. I am incredibly thankful for the numerous opportunities you have always given me to continue working in research, and for your implacable support throughout the past years. I am very grateful to have met you and learned from you. I hope that in the coming years we will still meet, here and there, and share our newest anecdotes.

Special thanks to Roland. I first met you as a colleague and a couple of years after you became my supervisor, which brought us to a closer friendship. You are my master, you teach me everything I asked for, you always had time for my silly questions, and you are the reason that I felt (and still feel) so fascinated about volatiles. I am so grateful for the time you invested with me in the labs, and the nice moments full of humour and interesting stories. I hope that, soon, you could come to my future farm and enjoy some wine together.

À mon amour, Antoine, whom with I have shared most of my routine in the past years, including the cheerful moments and the struggles. Thank you for always supporting me with the PhD, and to make me smile at the end of every day. Being with you made me stronger than ever. Now this is the end of a long chapter, but for us, it means the beginning of a new, and greatest adventure! I cannot wait to live it with you. Merci beaucoup.

To my dear friend and colleague Eirini. Since the first day you joined us in the group, as a new PhD member, I knew that you were going to be of good company, not only at work but also at home. You spark the day of everyone that crosses with you, and to me, you spark my whole existence! Thank you for the help you often provided me during the PhD. I wish to live more joyful moments with you, and of course together with Nefèli and Len. I am very grateful to have met the incredible family you are. Efcharistó ólous.

Very special thanks to my parents and siblings: Luis, Bibi, Sara, Elena, and Miguel. Sin vosotros no habría llegado nunca hasta aquí, me habéis hecho sentir tan orgullosa, por mi trabajo y por tener la familia que tengo, que no me la merezco! Muchas gracias por todo el apoyo que he recibido de vosotros. Incluso cuando aun no sabéis ni lo que he hecho durante estos cuatro años, hehe. De cualquier modo, el apoyo de la familia es lo que más cuenta. Gracias a todos mis tíos y tías, primos y primas que también se han preocupado tanto por mi trabajo. Y, por último pero no menos importante, mis adorables sobrina/os que, aunque no lo sepan, su nacimiento durante mi tesis ha sido un golpe de alegría en mi vida. Gracias por ser tan buenos padres, Elena y Javi.

I want to profoundly thank Charlotte, my one and only student. You are an exceptional human, not just because of the incredible work you do, but also as a very good friend. I will never be able to fully thank you for the work you performed during your master thesis, which became a very important part of my thesis as

well. You are such a very hard worker and kind person, I truly appreciate that we are keeping contact, and I hope we can continue seeing each other wherever we will be! I wish you a lot of success.

Very special thanks to my colleagues, the metabolomics group (Bert, Eirini, Henriëtte, Jeroen, Ric, Robert, and Roland) whom with I shared so many good Monday morning meetings, and entertaining times in our bbq/dinners/outings. Thank you for all the help you always offered to me during these past seven and half years. Working with you made everything much easier and fun!

I would also like to thank my colleagues at PPH and Bioscience. Even though I have not been very involved in the coffee breaks and lunches, I knew I could always count on you for anything I needed. Just seeing you every day in our meetings or outings was always transmitting me a lot of good energy. I am very proud of having such a big team of fighters surrounding me, what you all do is stunning. Thank you. I would also like to thank Justin and the Young NMC crew, who gave me the opportunity to be part of the organizing committee, and from whom I have learned so much along with.

I am thankful to the Chemosensers that supported and contributed so much to this thesis. Apart from the supervision I have received from Robert and Roland, having such a team of experts from different fields in the project helped tremendously. I would like to thank all of you for the useful training, the valuable inputs on the data generated and the necessary feedback (so many manuscripts to review…). I can say we came up with very interesting results! Thank you, Age, Amy, Anne-Charlotte, Brenda, Doris, Emilie, Johan, John, Justus, Marco, Margriet, Naser, Rudi, Simon, and Thomas.

To my friends in Spain, especially my village family "los montonawers" (Adrià, Angel, Aurora, Carles, Cea, David, Elena, Elsa, Ester, Javi, Jona, Marta, Nerea, Nuria, Raquel, Sara, and Silvia). Se que no hemos podido disfrutar mucho juntos estos últimos años, pero para mí sois mi familia, mis hermanos, siempre. Gracias por todo el apoyo que me habéis dado estos últimos años ¡Seguir luchando, y pronto (esperemos el próximo año) disfrutaremos de unas fiestas de Montón en condiciones! To my dear Paula and my other chemistry musketeers (Andrea, Carmen, Natalia, Nuria, Pilar, and Silvia), thank you for all those incredible years of Bachelor we spent together. Keep sparkling!

To my friends in The Netherlands. Since I arrived in Wageningen seven and half years ago, I have met so many beautiful people that shaped my life, I am very grateful to all of you:

First, I would like to thank my old-time friends from Wageningen. You were the reason why I extended my stay in Wageningen and ended up doing a PhD. Thank you for sharing those years together: Carlos, Carmen, Jaione, Jordi, Mariana, Marlene, Marta, Nicho, Oriana, Oscar, Ramón, Percy, Xabi, H14 housemates, and many more.

Special thanks to my dear "solitarias" Agata, María, and Serena. I think you are the ones who suffered more from my struggles, ups and downs of the PhD, thank you for supporting me always and for giving me the company and positive energy I needed every day after work. I am sad to see that we are becoming more

solitarias than ever, now that we are living far away, but I have no doubts we will still meet in the coming trips and have more beautiful adventures together!

I would like to thank all my housemates from House 31 (Alex, Clem, Georges, Hakuna, Hanno, Hestia, Jan, Luca, Milan, Reinout, Roberto, Serena, and Tobia,). Jan, you have been my housemate since my arrival at 31 until my departure. I still remember when we started hanging out more often. One day you asked me "please, tell me everything about your PhD" (lol). I really appreciated your interest in my research, and the nice conversations we had are unforgettable for me. Thank you 31ers for the support, the cheerful years we spent together. I am gratefully happy to have you in my life. This thesis is part of all of you as well!

Many thanks to my other friends and neighbours in Droevendaal, where I have lived three years of my PhD. I think that combining a work that I enjoyed, with the right surroundings, is the best combination you can ever get. The evenings I spent in good company or in solitude, were the best way to disconnect and re-fresh before going to sleep and start a new day. Thank you all for the nice vibes we shared. My beautiful girls, Alessia and Xristina ("malakita") who supported me so much in the last stretch, I love you! Adrà, Ángel, Charlie, Coco, Emilio, Gotham, Javier, Lison, Michelle, Miguel, Philip, Seba, Rowan, Vicky, Verónica, Romanos and the rest of the people of 39, B2 house, Guillitos, Sacha and the rest of 89ers, 51ers, arboretumlaaners, and so many more… Last but not least, my special thanks to my dear friend Alok ("Alokito") whom I dedicated part of this thesis. You were always so supportive when I needed the most, in my last year of the PhD, you always tried to make me laugh. For me, laughing or smiling is the truly reason to be at ease and joyful in the present moment, and that is how you made me feel. I will always miss you and I will always carry you with me everywhere.

My last special thanks to my grandparents who inspired me to live simpler and introduced me to the concept of living with less, in nature and under self-sustained principles. I miss you.


Foto

About the author

"This is not about being a superhuman,
this is about realizing being a human is super"
Jaggi Vasudev (Sadhguru)

Carmen Diez Simon was born in Calatayud (Spain) on 1st November 1990. During her childhood, she has lived in many different villages around Segovia and has always been passionate about exploring new places. In 2008, she moved to Zaragoza to study a double bachelor's degree in chemistry and biochemistry at the University of Zaragoza. During her Bachelor's, she specialized in analytical chemistry and mass spectrometry techniques. Her curiosity for the plant and food chemistry science brought her to Wageningen University and Research, in 2014, where she did her second BSc thesis with Prof. Dr Robert Hall on plant metabolomics. She learned greatly about high-throughput techniques to analyse volatile aroma compounds which allowed her to continue one more year working within the Plant Sciences Group, as a research assistant. Her specialization focused on mass spectrometry techniques, and the processing and interpretation of metabolomics data. As life can sometimes call for breaks, and since she always seeks for new adventures, in 2016 she started a year travelling and discovering other type of jobs (e.g. as a cook in a restaurant). By doing so, she gained confidence in herself as well as experience in other branches of knowledge.

One year later, in 2017, she returned to research and started her PhD in Wageningen University and Research within the same group, supervised by Prof Dr Robert Hall and Dr Roland Mumm. Her experience brought her new ideas and awe-inspiring values to follow. Due to her works on fermented soy sauces and the numerous wine tastings she attended, her interest in the chemistry behind natural fermented processes and wine aromas grew further. She is currently starting a new journey throughout Europe, by van, promoting natural wines and a sustainable lifestyle alternative for young professionals.

List of publications

This thesis:

C. Diez-Simon, R. Mumm, R.D. Hall. 2019. Mass spectrometry-based metabolomics of volatiles as a new tool for understanding aroma and flavour chemistry in processed food products. *Metabolomics*. 15.

C. Diez-Simon, B. Ammerlaan, M. van den Berg, J. van Duynhoven, D. Jacobs, R. Mumm, R.D. Hall. 2020. Comparison of volatile trapping techniques for the comprehensive analysis of food flavourings by Gas Chromatography-Mass Spectrometry. *J. Chromatogr. A.* 1624, 461191.

C. Diez-Simon*, C. Eichelsheim*, R. Mumm, R.D. Hall. 2020. Chemical and sensory characteristics of soy sauce: A review. *J. Agric. Food Chem*. 68.

C. Diez-Simon, C. Eichelsheim, D.M. Jacobs, R. Mumm, R.D. Hall. 2021. Stir bar sorptive extraction of aroma compounds in soy sauce: Revealing the chemical diversity. *Food Res. Int.* 144, 110348.

N. Davarzani*, **C. Diez-Simon***, J.L. Großmann, D.M. Jacobs, R. van Doorn, M.A. van den Berg, A.K. Smilde, R. Mumm, R.D. Hall, J.A. Westerhuis. 2021. Systematic selection of competing metabolomics methods in a metabolite-sensory relationship study. *Metabolomics*, .

Other publications in peer-reviewed journals and books:

N.S. Outchkourov, R. Karlova, M. Hölscher, X. Schrama, I. Blilou, E. Jongedijk, **C. Diez-Simon**, A-J. van Dijk, D. Bosch, R.D. Hall, J. Beekwilder. 2017. Transcription factor-mediated control of anthocyanin biosynthesis in vegetative tissues. *Plant Physiology*. 176 (2), 1862-1878.

Y. Zhang, X. Cheng, Y. Wang, K. Flokova, **C. Diez-Simon**, A. Bimbo, F. Péréfarres, K. Posthuma, H.J. Bouwmeester, C. Ruyter-Spira. 2018. The tomato MAX1 homolog, SlMAX1, is involved in the conversion of tomato strigolactones from carlactone. *New Phytologist*. 219, 297-309.

M. Tomas, J. Beekwilder, R.D. Hall, **C. Diez-Simon**, O. Sagdic, E. Capanoglu. 2018. Effect of dietary fiber (inulin) addition on phenolics and in vitro bioaccessibility of tomato sauce. *Food Res. Int*. 106, 129-135.

D.W. Etalo, **C. Diez-Simon**, R.C.H. de Vos, R.D. Hall. Part of the *Methods in Molecular Biology* book series, 2018. Laser ablation electrospray ionization-mass spectrometry imaging (LAESI-MS) for spatially resolved plant metabolomics. Plant Metabolomics, 253-267.

In preparation:

S. Leygeber, **C. Diez-Simon**, J.L. Großmann, …, T. Hankemeier. Flavour profiling by comprehensive MS analysis of metabolites in complex food matrices.

S. Leygeber*, **C. Diez-Simon***, J.L. Großmann*, …, D.M. Jacobs, R. Mumm. Multi-platform mass spectrometry analyses in combination with sensory profiling of a large yeast product cohort to explore chemical impact on taste and aroma.

** Equal first authors*

Overview of completed training activities

Discipline specific activities

Training course, Gerstel, Wageningen, 2018

Young AGErs Symposium, VLAG, Wageningen, 2018

National meeting Benelux, Netherlands Metabolomics Centre, Rotterdam, 2018

Chemometrics course, VLAG, Wageningen, 2019

International Metabolomics conference, Metabolomics Society, The Hague, 2019

Weurman Flavour Research Symposium, CSGA, Dijon, 2021

Training period at DSM laboratories, DSM Food Specialties, Delft, 2020-2021

General courses

VLAG PhD week, VLAG, Baarlo, 2017

Basic R course for researchers, Plant sciences, Wageningen, 2017

Introduction to R, VLAG, Wageningen, 2019

Scientific writing, VLAG, Wageningen, 2019

Presenting with impact, VLAG, Wageningen, 2019

Applied statistics, VLAG, Wageningen, 2019

Teaching obligations

Teaching Plants & Health course, Wageningen, 2019 and 2020

Teaching and supervision Food Chemistry, Wageningen, 2019

Supervision MSc student thesis, Wageningen, 2019

Optional activities

Preparation of research proposal, Wageningen, 2017

Weekly group meetings, Bioscience & PPH, 2017-2021

Literature discussion meetings, PPH, 2017-2020

Project meetings CHEMOSENSE, 2017-2021

The research described in this thesis was financially supported by DSM Food Specialties, Unilever Research, and the Netherlands Organization for Scientific Research (NWO) under grant agreement number 731.015.207. Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover design by Carmen Diez Simon

Printed by Proefschriftmaken.nl