Article

Artificial Intelligence Sensing: Effective Flavor Blueprinting of Tea Infusions for a Quality Control Perspective

Andrea Caratti , Angelica Fina , Fulvia Trapani , Carlo Bicchi , Erica Liberto , Chiara Cordero * , Federico Magagna

Posted Date: 9 January 2024

doi: 10.20944/preprints202401.0707.v1

Keywords: flavor blueprint; black tea infusions; sensomics-based-expert-system; industrial quality control; Artificial Intelligence sensing; accurate flavor screening

Article

Artificial Intelligence Sensing: Effective Flavor Blueprinting of Tea Infusions for a Quality Control Perspective

Andrea Caratti [$], Angelica Fina [$], Fulvia Trapani, Carlo Bicchi, Erica Liberto, Chiara Cordero * and Federico Magagna

Dipartimento di Scienza a Tecnologia del Farmaco, Università degli Studi di Torino, 10125 Turin, Italy; andrea.caratti@unito.it (A.C.); angelica.fina@unito.it (A.F.); fulvia.trapani@unito.it (F.T.); carlo.bicchi@unito.it (C.B.); Erica.liberto@unito.it (E.L.); chiara.cordero@unito.it (C.C.); federico.magagna@gmail.it (F.M.)

* Correspondence: chiara.cordero@unito.it (C.C); Tel.: +39 011 6702197

[$] Authors A.C. and A.F. listed in alphabetical order equally contributed to this work.

Abstract: Tea infusions are the most consumed beverages in the world after water, their pleasant yet peculiar flavor profile drives consumer choice and acceptance and becomes a fundamental benchmark for industry. Any qualification method capable of objectifying the product's sensory features effectively supports industrial quality control laboratories guaranteeing high sample throughputs even without the human panel intervention. The current study presents an integrated analytical strategy acting as an Artificial Intelligence decision tool for black tea infusion aroma and taste blueprinting. Key markers validated by sensomics are accurately quantified in a wide dynamic range of concentrations. Thirteen key aromas are quantitatively assessed by standard addition with in-solution solid-phase microextraction sampling followed by GC-MS. On the other hand, nineteen key taste and quality markers are quantified by external standard calibration and LV-UV/DAD. The large dynamic range of concentration for sensory markers is reflected in the selection of seven high-quality teas from different geographical areas (Ceylon, Darjeeling Testa Valley and Castleton, Assam, Yunnan, Azores, and Kenya). The strategy acts as an AI smelling and taste machine predicting teas sensory features without the human panel intervention.

Keywords: flavor blueprint; black tea infusions; sensomics-based-expert-system; industrial quality control; Artificial Intelligence sensing; accurate flavor screening

1. Introduction

Tea plant has a great economic importance and, after water, tea infusions are the most consumed beverages in the world; they are prepared by hot water extraction of dried leaves from *Camellia sinensis* L. [1–3]. According to their manufacturing, teas can commercially be classified into three main categories: non-fermented (*white* and *green* teas), partially fermented (*oolong* and *paochong* teas), and fully fermented (*black* tea) [4,5]. Green and black teas are widespread in the world, while the other kinds of tea (white, yellow, and oolong) are mainly consumed in producing countries [6].

Tea is characterized by important physiological and potential health benefits [7–10] due to the presence of many bioactive chemical constituents. The chemical composition of the dried tea plant (green or black) includes primary metabolites such as proteins, free amino acids, carbohydrates, vitamins, minerals, and specialized metabolites, *i.e.* polyphenols, purine alkaloids (methyl xanthines and, above all, caffeine), chlorophyll, and other volatile and non-volatile compounds [1].

The peculiar combination of non-volatile solids extractable from tea leaves and the complex volatile fraction evokes tea's characteristic flavor, which is fundamental in influencing its market value and consumer choices. It is well-known that the quality of food is not only guaranteed by its

safety and health properties but a significant role is played by its hedonic profile, especially flavor, and appearance [11–20].

Tea flavor has represented an intriguing research topic since the 1930s [5] and still today this area of investigation is active. The increase in the industrial production of *ready-to-drink* beverages requires an in-depth knowledge of the quality of raw materials (traceability, and flavor features) to obtain products with *excellent* yet benchmarked flavor.

The literature reports several studies focused on compounds responsible for the characteristic flavor of teas differing for geographical provenience/botanical variety. In 2006, Schuh and Schieberle [21] applied the *molecular sensory science* approach, nowadays referred to as *sensomics*, to identify and quantify key aroma compounds in the infusion prepared from Darjeeling (India) black teas. The approach includes solvent-assisted flavor evaporation (SAFE) followed by gas chromatography-olfactometry (GC-O) with aroma extract dilution analysis (AEDA) to identify potent odorants and their odor qualities. By stable isotope dilution assay (SIDA) accurate quantitative determination of odorants highlights those exceeding the odor threshold (OT) in the sample thereby indicating how to reconstruct an aroma recombinate that evokes the unique and distinctive aroma identity (*aroma blueprint*) of the product [22]. The authors identified a total of 42 impacting odorants both in the leaves and in the resulting infusions with some quantitative differences. Among 16 character impact odorants, terpenes alcohols (*geraniol* and *linalool*), Strecker aldehydes (*2-methyl propanal*, *2-* and *3-methyl butanal*), carotenoids derivates (*β-damascenone*), significantly increased their concentration after the infusion process.

A similar analytical procedure was adopted by Scharbert and co-workers [23] to identify the key molecules generating the taste perception after tea infusion consumption. Their studies on Darjeeling tea revealed as the main contributors to the astringent taste perceived upon black tea consumption a series of compounds belonging to the flavonoids class, *i.e. flavan-3-ol glycosides*. Borse *et al*. [24] investigated the characteristic distribution of volatiles and non-volatiles in black teas originating from different regions of India (Darjeeling, Assam, Nilgiris, etc.). In that study, tea chemical fingerprints were identified by associating physicochemical assays, spectrophotometric and liquid chromatography (LC) analyses for the determination of *flavonoids* and *caffeine*, while simultaneous distillation-extraction (SDE) followed by GC coupled to mass spectrometry (MS) were used to map the volatile fraction. In particular, 25 volatile flavor compounds were demonstrated to be diagnostic in discriminating Indian tea samples from the others.

In 2008, Wang and co-workers [25] profiled 56 teas with different degrees of fermentation (green, oolong, and black teas) using LC and head-space solid phase micro-extraction (HS-SPME) followed by gas chromatography-mass spectrometry (GC-MS). They found that neither the total nor individual *catechins* content was different in green and oolong teas among the investigated samples while the post-harvest process (fermentation) was responsible for major differences in volatiles distribution. Within the volatiles, five components [*(E)-2-hexenal*, *benzaldehyde*, *6-methyl-5-hepten-2-one*, *methyl salicylate*, and *indole*) were able to discriminate unfermented and fermented teas, while *(E)-2-hexenal* and *methyl salicylate* together supplied an index to differentiate semi- and fully-fermented teas. Special teas (*e.g.*, Chinese green tea *Jingshan cha* and *Longjing*) were recently investigated for aroma blueprinting by sensomic principles [26,27] posing the basis of a new concept/methodology termed sensomic-based-expert-system (SEBES) [28].

The methodological approach SEBES is capable of "characterize key food odorants with one single analytical platform and without using the human olfactory system, that is, by *artificial intelligence smelling*" [28]. The SEBES is capable of identifying the set of key food odorants (KFO) in a product and by combining OTs with quantitative data, the OAVs are automatically estimated and aroma features are predicted with great accuracy. This way the food odor code is "defined without using the human olfactory system" [28]. This application of Artificial Intelligence (AI) to food sensory properties is intriguing yet attractive for the industry quality control (QC) where primary materials and finished products have to be benchmarked and compared to references.

 doi:10.20944/preprints202401.0707.v1

In this study, the SEBES concept is developed, validated, and applied for aroma and taste (*i.e.*, flavor) blueprinting of a selection of black tea infusions. To support the high throughput required by quality controls (QC), flavor and quality markers are quantitatively profiled by fully automated GC-MS and LC with ultra-violet diode-array (UV/DAD) avoiding time-consuming sample preparation. The strategy, if connoted by a suitable dynamic range and accuracy could support product quality assessment with a focus on flavor features avoiding human panel.

In particular, for the aroma blueprinting in-solution (IS) SPME-GC-MS is implemented with Standard Addition (SA) quantification while for taste blueprinting LC-UV/DAD is with external standard (ES) calibration of key-aromas and marker compounds [29]. The suitability of the AI tea flavor blueprinting is tested on a selection of commercial teas from all over the world.

2. Results and Discussion

The number of chemicals effectively contributing to the flavor of food (key aroma, key taste, and trigeminally active compounds) is relatively small, and complex analytical procedures are required to detect, identify, and quantify flavor-active components often occurring at very low concentrations (µg/L or below) in the final product. Concerning aroma-active compounds, the adoption of a suitable sampling technique is fundamental to obtaining a meaningful picture of components evoking the sensory identity and quality profile of the product [30]. In this respect, the so-called high concentration capacity sampling techniques (HCC) [3,31–34] are the elective route to achieve suitable selectivity, sensitivity, and quantification accuracy to provide high-throughput informative analysis in full automation. In a previous study, several HCC approaches were tested for their capability to efficiently and accurately delineate the aroma blueprint of black teas from Ceylon [3]. Of the tested techniques, HS-SPME-GC-MS resulted as the most suitable for in-solution sampling of key-aroma compounds providing the possibility of accurate quantification of analytes by standard addition (SA) and calibration [35].

About taste active and trigeminally active compounds, teas are characterized by a complex pattern of specialized metabolites belonging to *flavonoids* and *methyl xanthines* classes which are responsible for taste and trigeminal sensations. In particular, *caffeine* is responsible for bitterness while glycosidic derivates of *flavonols* (*quercetin*, *myricetin*, and *kaempferol*) are the most taste-active analytes imparting a mouth-drying and velvety-like astringency. In addition, an important role in defining tea quality (dried plant and related infusion) is also played by *flavan-3-ols* which exert a minor but not negligible contribution on overall tea astringency perception. Depending on their concentration, *flavan-3-ols* (*catechins*) and dimeric derivates (*theaflavins*) evoke a puckering astringency and rough oral sensation [36].

The following sections present and critically discuss the experimental results on the quali-quantitative profiling of the most potent flavor components of selected teas after standardized infusion according to the EMA/HMPC/283630/2012 Committee on Herbal Medicinal Products (HMPC) protocol. Teas were all fermented (i.e., black teas) and were from Ceylon (*Flowery Orange Pekoe* – FOP), from India (Assam, Darjeeling Testa Valley, Darjeeling Castleton), Portugal (Azores), China (Yunnan), and Kenya.

2.1. Tea infusion volatiles profiling by IS-SPME-GC-MS

As a first step, GC-MS analyses acquired in full-scan mode enabled a comprehensive mapping of the volatile fraction of tea infusions. Volatiles were reliably identified by their EI-MS fragmentation patterns, compared to those present in commercial libraries (Wiley [37] and NIST 2014 [38]). Acceptability criteria for putative identification were direct match factor (DMF) > 900; linear I^T_S tolerance ±5 units. When available pure reference compound confirmation was done. **Table 1** reports the list of 44 volatile marker compounds together with their retention times, experimental I^T_S, odor descriptors, and presence in analyzed samples. Although the number of detected and reliably identified volatiles was lower if compared to the GC×GC methodology [2,3] adopted in previous investigations, the proposed strategy by IS-SPME-GC-MS allowed for efficiently mapping the most

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0707.v1

relevant markers carrying information about tea aroma profile, plant origin, and manufacturing practices.

Table 1. List of the targeted volatiles detected in selected tea infusions by IS-SPME-GC-MS together with retention times (t_R min), experimental I^T, and odor descriptors as reported in reference literature [12,26,27,39,40]. Detected (d) and non-detected (nd) analytes are also reported for each tea sample; moreover, Ceylon tea is rendered as an average profile obtained from different quality grades. Analytes with an asterisk (*) are the key odorants confirmed by sensomics and subjected to accurate quantification in this study.

#	Compound name	t_R min	I^T_S	Odor	Ceylon	Assam	Azores	Darjeeling Castelton	Darjeeling Testa Valley	Kenya	Yunnan
1	3-Methyl	2.61	541	Malty	d	d	d	d	d	d	d
2	2-Methyl butanal*	2.70	566	Malty	d	d	d	d	d	d	d
3	(E)-2-Pentenal	4.10	725	Green, apple	d	d	nd	nd	d	d	nd
4	(Z)-2-Penten-1-ol	4.85	788	-	d	d	nd	nd	d	nd	nd
5	Hexanal*	5.19	816	Grassy-	d	d	d	d	d	d	d
6	(E)-2-Hexenal	6.82	856	Bitter,	d	d	d	d	d	d	d
7	1-Hexanol	7.65	873	Fruity	nd	nd	nd	nd	d	d	d
8	2-Heptanone	8.40	890	Sweet,	d	d	d	nd	d	nd	nd
9	(Z)-4-Heptenal*	8.48	896	Fishy	d	d	nd	d	d	d	d
10	Heptanal	8.79	899	Oil, fatty	d	d	d	d	d	d	d
11	(E)-2-Heptenal	11.10	953	Fatty,	d	d	nd	nd	nd	d	nd
12	Benzaldehyde	11.10	954	Almond,	d	d	d	d	d	d	d
13	6-Methyl-5-	12.42	984	Pungent,	d	d	nd	nd	d	d	nd
14	(E,Z)-2,4-	12.79	993	Fatty,	d	d	nd	d	d	d	nd
15	(E,E)-2,4-	13.43	1008	Fatty,	d	d	nd	d	d	d	d
16	Limonene	14.17	1026	Citrus	d	nd	nd	d	d	d	d
17	2,2,6-Trimethyl	14.51	1030	-	d	nd	nd	nd	nd	nd	nd
18	Benzyl alcohol	14.73	1035	Sweet,	d	d	nd	nd	d	nd	nd
19	Phenyl	14.90	1039	Honey-like	d	d	d	d	d	d	d
20	(E)-2-Octenal	15.67	1055	Green, nut,	d	d	nd	nd	d	d	nd
21	Trans-linalool-	16.98	1071	Sweet,	d	d	d	d	d	d	d
22	Cis-linalool-3,6-	17.13	1087	floral,	d	d	d	d	d	d	d
23	Linalool*	17.67	1101	Citrus	d	d	d	d	d	d	d
24	Nonanal	18.03	1103	Fatty, waxy	d	nd	nd	nd	d	nd	d
25	2-Phenyl alcohol	18.30	1111	Honey-like	d	nd	nd	nd	d	d	nd
26	(E,Z)-2,6-	20.12	1151	Cucumber-	d	d	nd	d	d	d	nd
27	(E)-2-Nonenal*	20.43	1157	Fatty, green	d	d	nd	nd	d	d	d
28	Cis-linalool-3,7-	20.93	1168	Sweet,	d	d	nd	nd	d	d	nd
29	Trans-linalool-	21.18	1173	floral,	d	d	nd	nd	d	d	nd
30	Methyl salicylate	21.95	1192	-	d	d	d	d	d	d	d
31	Safranal	22.25	1196	Saffron	d	nd	nd	nd	d	d	nd
32	Decanal	22.45	1200	Penetrating,	d	nd	nd	nd	d	d	nd
33	(E,E)-2,4-	22.93	1215	Fatty, green	d	d	d	nd	d	d	d
34	Geraniol*	24.79	1254	Rose-like	d	d	d	d	d	d	d
35	Geranial	25.55	1267	Citrus	d	nd	nd	d	d	d	nd
36	Trans anethole	26.35	1280	Sweet	nd	d	nd	d	d	nd	nd

 doi:10.20944/preprints202401.0707.v1

37	(E,Z)-2,4-	26.54	1291	Deep-fried	d	d	nd	d	d	d	nd
38	(E,E)-2,4-	29.64	1319	Fatty, fried	d	d	nd	d	d	d	nd
39	β-Damascenone	30.48	1381	Fruity	d	d	nd	nd	d	d	nd
40	(Z)-Jasmone	31.01	1400	Floral,	d	nd	nd	d	d	nd	nd
41	α-Ionone	32.24	1424	Violet-like	d	d	nd	d	d	d	d
42	Geranyl acetone	33.31	1450	Magnolia,	d	d	nd	d	d	d	d
43	β-Ionone*	34.62	1483	Violet-like	d	d	nd	d	d	d	d
44	Caffeine	43.77	1841	-	d	d	d	d	d	d	d

Results on the volatiles profiling showed a different distribution of analytes within the selected samples, Ceylon and Darjeeling Testa Valley (India) were characterized by a more complex volatile fraction, showing a matching of 42 detected analytes over 44 targets. The same outcome cannot be observed with the other teas from Darjeeling (Castleton) where only 26 volatiles were detected, demonstrating how different tea gardens, located in the same country, lead to distinctive products. On the other hand, tea samples from Azores (Portugal) and Yunnan (China) showed a less complex fraction of volatiles, also in terms of key aroma compounds distribution; indeed, within all identified analytes, only 15 compounds were mapped in Azores tea and 23 in Yunnan tea.

Qualitative *profiling* results show that some volatiles are ubiquitous in all tea infusions, although quantitative differences deserve some comments. Of those detected in all samples, Strecker aldehydes (*2-methyl butanal* and *3-methyl butanal*) and volatile terpenes (*linalool* and its related *3,6-oxides*, and *geraniol*) are the most relevant in the aroma definition. Phenyl propanoid derivates, a group of characteristic components in tea, followed a slightly different behavior; *phenylacetaldehyde* and *benzaldehyde* were detected in all samples, while *benzyl alcohol* and *2-phenyl alcohol* were present only in some products (both for Ceylon and Darjeeling T.V, *benzyl alcohol* in Assam tea, the others in tea from Kenya). Within the entire set of volatiles, an important role is played by saturated and unsaturated aldehydes which originate from the oxidation of fatty acids and contribute to defining the aroma of tea infusions. In this case, it is interesting to point out that some short-chain linear aldehydes (*hexanal*, *heptanal*) are present in all infusions while unsaturated C7-C10 aldehydes have characteristic patterns in Ceylon, Assam, Darjeeling T.V. and Kenya samples.

The next step focused on key aroma quantitation by Standard Addition (SA), a well-established internal calibration approach suitable when the so-called *matrix effect* cannot be neglected and likely has an impact on method accuracy. Indeed, in the case of tea infusions, the release/behavior of volatiles is strictly influenced by their interaction with non-volatile components (e.g., *polyphenols*, *alkaloids*, organic acids, pigments). The next section introduces the SA procedure and quantitation results.

2.2. Key-aroma markers quantitation by Standard Addition (SA) and IS-SPME-GC-MS

The standard addition procedure, widely used as a quantitation approach, consists of a series of experiments in which the original sample and a suitable number (at least four concentration levels) of aliquots of the sample spiked with increasing and known amounts of reference compounds, are submitted to the analytical process.

When using the single addition method, the analyte concentration in the sample can be estimated from Equation (**1**):

$$A_{(0+a)}=\frac{A_0}{W_0}W_a+A_0 \quad (1)$$

where: W_0 is the amount of analyte in the matrix, W_a the amount of analyte added to the sample, A_0 the instrumental response obtained from analysis of the original sample, and $A_{(0+a)}$ the instrumental response of the analyte obtained from analysis of the spiked sample.

A preferable and more accurate procedure, which was applied in this study, includes multiple standard additions. With multiple SA a linear regression analysis evaluates the terms W_a

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0707.v1

and $A_{(0+a)}$ therefore the amount of analyte in the matrix (W_0) is given by the ratio between the intercept and the slope, Eq. (**2**):

$$\frac{b}{a} = \frac{A_0}{\frac{A_0}{W_0}} \quad (2)$$

Standard addition is a quantitation approach that can be carried out in different ways: (*a*) by spiking the target analyte(s), in a gaseous state, into the sample headspace (gas phase addition - GPA); (*b*) by spiking the analyte(s) dissolved in a suitable solvent, directly onto the sample (sample phase addition - SPA) or (*c*) by spiking the stable-isotope-labeled analyte(s) dissolved in a suitable solvent (stable isotope dilution analysis - SIDA) onto the sample. The present study adopted the SPA protocol, as being suitable for its ease of implementation and automation, and its cost-effectiveness compared to isotopically-labeled standards.

The analytical protocol, rationalized in the Experimental section, consisted of (*a*) three replicate infusions prepared from each tea sample; and (*b*) three standard addition levels for each infusion (plus the analysis of the original sample). Acetone was selected as a solvent for spiking solutions because it guaranteed full solubilization of all target analytes, being miscible with water. For each calibration step two analytical replicates were acquired.

Thirteen key odorants were accurately quantified: 3-methyl butanal, 2-methyl butanal, hexanal, (Z)-4-heptenal, phenyl acetaldehyde, linalool, (E,Z)-2,6-Nonadienal, (E)-2-nonenal, (E,E)-2,4-nonedienal, geraniol, (E,E)-2,4-decadienal, β-damascenone and β-ionone. They are listed in **Supplementary Table 1 (ST1)** together with chromatographic information on retention times (t_R min), Target/Qualifier Ions m/z, calibration functions and determination coefficients (R^2), precision and accuracy results. For the method's performance parameters evaluation see the experimental section, section 3.9.

Table 2 lists quantitation results, obtained by SA calibration technique and IS-SPME-GC-MS (SIM) for key aromas in tea infusions. The results referred to Ceylon teas are reported as the mean value obtained from four commercial batches from the same harvest year. Data are expressed as μg/L in the infusion.

Experimental results are consistent with those obtained by Schuh and Schieberle [21] which adopted a more complex procedure for isolation and accurate quantification of potent odorants, *i.e.*, SIDA and Solvent Assisted Flavour Evaporation (SAFE) followedby GC-ofactometry (GC-O) and aroma extract dilution assay AEDA).

Table 2. SA quantitative results on key aroma compounds of tea infusions, data are expressed in μg/L.

Compound name	Odor Threshold (μg/L)	Ceylon	Assam	Azores	Darjeeling Castelton	Darjeeling Testa Valley	Kenya	Yunnan
3-Methyl butanal	1.2	37.00	49.12	52.32	25.22	27.45	84.89	17.02
2-Methyl butanal	4.4	45.44	63.12	46.69	29.50	21.22	89.97	18.29
Hexanal	10	45.25	23.04	14.84	15.30	63.49	31.75	16.88
(Z)-4-Heptenal	0.06	0.98	0.68	0.00	0.20	0.82	0.41	0.47
Phenyl	6.3	32.73	62.73	29.76	10.77	77.52	38.76	16.11
Linalool	0.6	25.75	9.72	10.19	7.86	54.48	27.24	0.41
(E,Z)-2,6-	0.03	0.56	0.39	0.00	0.50	0.28	0.14	0.00
(E)-2-Nonenal	0.4	0.40	0.24	0.00	0.02	1.21	0.60	0.15
(E,E)-2,4-	0.2	0.29	0.39	0.14	0.00	1.46	0.73	0.29
Geraniol	3.2	13.20	1.07	16.38	11.31	24.83	12.41	2.78
(E,E)-2,4-	0.16	0.51	0.15	0.00	2.29	0.21	0.11	0.00
Beta	0.004	0.26	0.29	0.00	0.00	0.38	0.19	0.00
Beta ionone	0.2	2.16	0.78	0.00	0.18	3.84	1.92	0.29

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0707.v1

An unsupervised multivariate approach (*i.e.* Principal Component Analysis – PCA) provides prompt information on samples' natural clustering based on compositional similarities. **Figure 1a** shows the scores plot on the first and the second principal components (F1-F2 plane) obtained by analyzing the distribution of targeted odorants in all sample replicates. The first principal component (F1) explains 57.70% of the total variance (74.85%), contributing most to the discrimination of samples, while the second principal component (F2) has a minor informative influence (17.15%). Three main groupings (ellipses with dotted line) arbitrarily delineated by the authors can be observed in the scores plot of **Figure 1a**: from left to right with increasing F1 scores values the Darjeeling Castelton, Azores and Yunnan, followed by Assam, Ceylon and Kenya, and finally Darjeeling Testa Valley (TV) with higher scores values along F1. The distribution of variables as a function of the first two principal components is reported in the loadings plot of **Figure 2b**. Analytes providing the most information on the F1 axis, and likely more abundant in related infusions, directly correlate with Ceylon, Assam, Kenya, and Darjeeling TV teas.

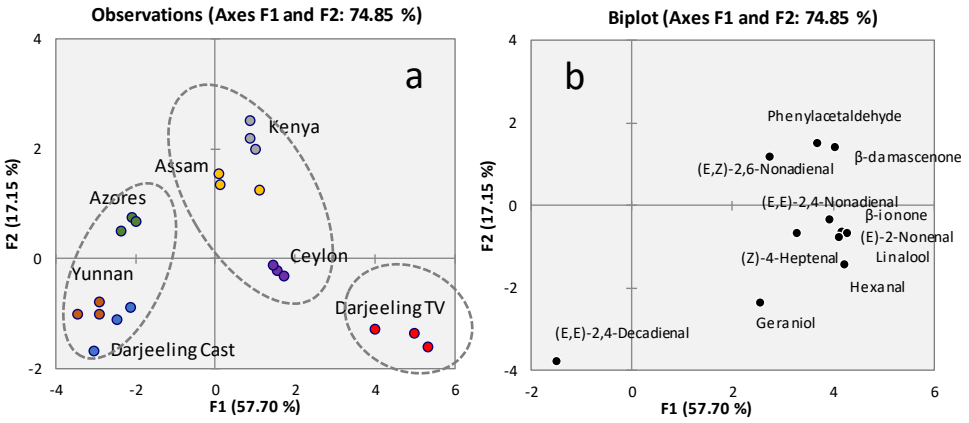

Figure 1. PCA results. (**1a**) scores plot on the first and the second principal components (F1-F2 plane) based on key- aroma compounds quantitative data determined in infusions prepared from tea leaves (Table 2). For Ceylon teas an average profile obtained from four lots was considered; (**2b**) loadings plot with the distribution of quantified analytes.

Interestingly Darjeeling teas, although produced in the same region of India, showed distinctive yet different aroma-active compound patterns, with Darjeeling TV characterized by higher amounts of these analytes. On the other hand, teas from Ceylon, Assam, and Kenya were clustered together suggesting similar aroma features.

Experimental results indicate that the most potent odorants characterizing these kinds of tea are present in a wide range of concentration, to be specific 0-4 μg/L for *(Z)-4-heptenal*, *(E,Z)-2,6-nonadienal*, *(E)-2-nonenal*, *(E,E)-2,4-nonadienal*, *(E,E)-2,4-decadienal*, *β-damascenone* and *β-ionone*, while in the range 0-100 μg/L for *hexanal*, *linalool*, *geraniol*, *phenyl acetaldehyde*, 2 and *3-methyl butanal*.

Darjeeling TV has a peculiar profile described by the flowery terpenes *linalool* (54.48 μg/L) and *geraniol* (24.83 μg/L), the green-grassy note from *hexanal* (63.49 μg/L) and fatty nuances likely modulated by unsaturated aldehydes *(E)-2-nonenal* and *(E,E)-2,4-nonadienal*. The commercial selection of Ceylon infusions are described by higher amounts of *(Z)-4-heptenal* (0.98 μg/L) and *(E,Z)-2,6-nonadienal* (0.56 μg/L). Other key volatiles such as *phenylacetaldehyde*, *β-damascenone*, and *β-ionone* play a role in defining the peculiar profile of Ceylon, Assam, Kenya, and Darjeeling TV.

Teas from Kenya and Assam are connoted by higher amounts of Strecker aldehydes *3-methyl butanal* (49.12 μg/L for Assam and 84.89 μg/L for Kenya) and *2-methyl butanal* (63.12 μg/L for Assam and 89.97 μg/L). Moreover, within the cluster of Ceylon, Assam and Kenya other major differences occur for *β-ionone*, more abundant in Ceylon (2.16 μg/L) and Kenya (1.92 μg/L).

In general, Azores, Yunnan, and Darjeeling Castleton showed a weaker profile of aroma-active analytes; in particular, many *key odorants* (unsaturated aldehydes and the nor-isoprenoids *β-damascenone* and *β-ionone*) were not detected and quantified in the related infusions. However,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0707.v1

tea from the Azores showed a fairly high amount of Strecker aldehydes (*3-methyl butanal*, *2-methyl butanal* 46.69 µg/L) which likely imparts malty notes.

Nevertheless, the great variability in the quantitative distribution of *key-odorants* does not necessarily lead to a meaningful characterisation of the selected samples for their sensory quality. Odorants, besides their intrinsic potency that is related to the binding with odor receptors (ORs), should be effectively released by the food matrix to reach the olfactory epithelium and trigger retronasal olfaction (i.e., aroma perception). The ratio between analyte's concentration in the sample (*i.e.*, tea infusion) and its odor threshold (OT, *i.e.*, the lowest concentration of a compound that is just enough for the recognition of its odour [36]) in water provides a more realistic perspective of the overall aroma quality and odorants balancing. For this reason, the contribution of the analytes in the prediction of samples sensory profile is evaluated by applying the SEBES concept.

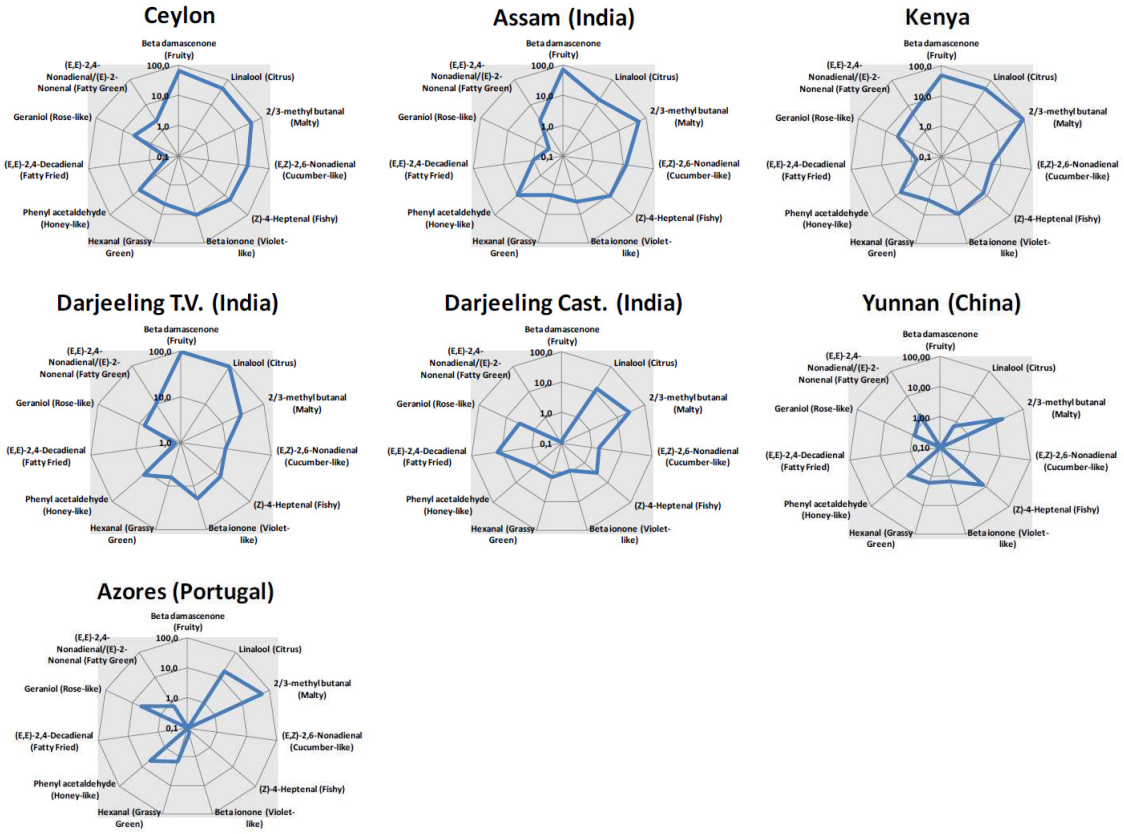

Figure 2. Spider diagrams with Odour Activity Values (OAVs) illustrating the contribution of each *key-aroma* compound to the aroma perception of selected teas. OAV data are reported in logarithmic scale. Ceylon tea aroma profile is reported as average value calculated on four different lots.

2.3. Aroma Blueprinting by AI smelling based on sensomics

As reported by Schieberle and co-workers in their studies on Darjeeling black tea (infusions and dried leaves) [21,36], a group of 24 odorants with high flavour dilution (FD) factors was recognized to play a prominent role in defining the characteristic aroma of the final infusion. Within these 24 chemicals, 16 were identified as having a high odor activity value (OAV), *i.e.*, the ratio between the odorant concentration in the food *vs.* its odor threshold. It is commonly assumed that the higher the OAV value the higher its contribution to the overall sensory perception.

Among the most-odor-active compounds revealed by sensomics (16), *i.e.* those with OAV values ≥ 1 (value recognized to be significant in contributing to the flavor of food [41]), the current method reliably monitors 13 of them with a fully automatized procedure that avoids laborious sample-preparation steps. By IS-SPME-GC-MS it is realized an effective aroma blueprinting to

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0707.v1

support, or even replace, sensory panel evaluation in the perspective of quality benchmarking and quality controls.

About aroma features, the Strecker aldehydes formed during the fermentation process [1] (*2-methyl butanal* and *3-methyl butanal*) concur to an intense malty perception while *phenylacetaldehyde* evokes a pleasant honey-like note. Responsible for floral notes are *geraniol* (rose-like) and *β-ionone* (violet-like), for fruity *linalool* and *β-damascenone*. An important contribution to the black tea aroma identity is provided also green and grassy notes of *hexanal*, *(E)-2-nonenal,* and *(E,E)-2,4-nonadienal*. Fatty perception is modulated by C9 unsaturated aldehydes while fishy notes are from [*(Z)-4- heptenal*] and fatty/fried by [*(E,E)-2,4-decadienal*]. All these aldehydes are formed from the enzyme-catalized oxidation of fatty acids during plant growth and manufacturing processes [40].

The aroma blueprint of the infusions is visualized as spider diagrams based on OAV values in a logarithmic scale (**Figure 2**). By this visualization for each compound the concentration is associated with the relative OT in water and the odor descriptor. The spider diagrams show that most tea samples have a specific aroma profile characterized by a well-balance of markers; in general teas from Ceylon, Assam, and Kenya likely have a similar aroma profile, especially for some notes such as fruity (*β-damascenone*), malty (*Strecker aldehydes*), honey-like (*phenyl acetaldehyde*), cucumber-like (*(E,Z)-2,6-nonadienal*). On the other hand, Assam infusion is characterized by a lower content of terpene derivates (*geraniol* and *linalool*) than the other teas, likely resulting in weaker citrus and rose-like notes. The aroma profile confirms the great diversity of samples from Darjeeling.

A very important outcome deriving from the calculation of OAVs consists in the explanation of the role played by each specific analyte in the definition of the aroma profile. Although present at very low concentrations (range 0-4 μg/L), some analytes such as *β-damascenone*, *β-ionone*, *(Z)-4-heptenal,* and *(E,Z)-2,6-nonadienal* are extremely potent being characterized by very low OTs (*e.g.* 0.004 μg/L for *β-damascenone*); as an example, *β-damascenone* generates the maximum OAV value of 94.8 in Darjeeling Testa Valley tea, where OAV values were higher also for other key-volatiles (*i.e.*, *linalool*, *3-methyl butanal*, and *2-methyl butanal*). On the contrary, despite their high concentration in the beverage, *hexanal* and *phenylacetaldehyde* likely play a minor role in the definition of tea aroma as being less active as odorants (higher OTs, 10 and 6.3 μg/L respectively).

2.4. Taste active compounds and quality markers accurate quantitative profiling by LC-UV/DAD

Tea infusions were then profiled by LC-UV/DAD to quantitatively map non-volatiles responsible for taste and trigeminal perception and for quality features [8,29]. The analytical method was optimized and verified for the accurate quantification of 19 informative chemicals in a single analytical run. Analytes are listed in **Table 3** together with taste threshold (TT) and average amounts (mg/L) from three replicate infusions for each sample. Analytical method figures of merit, including precision and accuracy are discussed in the Experimental section and detailed in **Supplementary Table 1** (**ST1**).

Principal Component Analysis was applied on the data matrix (quantitative results) of non-volatiles to evaluate the presence of natural clusters within selected samples. **Figure 3a** shows the scores plot on the first and the second principal components (F1-F2 plane); to note the variance explained by the first two components is quite high and similar to that resulted from key-aroma patterns shown in **Figure 1a** (58.55% F1, 20.10% F2, for a total variance of 78.65%). Moreover, sample clusters based on taste and quality markers distribution are similar to those shown in **Figure 1a** generated by volatiles patterns. Darjeeling Testa Valley is characterized by a fingerprint that drives its independent clustering; it is described by a peculiar pattern of *flavan-3-ols* (**Figure 3b**) such as *epigallocatechingallate* (EGCG), *epigallocatechin* (EGC), *catechin* (C) and *epicatechingallate* (ECG). Experimental results on Darjeeling black tea are consistent with those obtained by Scharbert and co-workers in 2004 [23]. Ceylon, Assam, and Kenya teas sub-classification is mostly driven by *theaflavins* and *flavonol-3-o-glycosides*. On the other hand, tea samples from Azores, Yunnan, and Darjeeling Castleton inversely correlate with these variables (weaker profiles).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0707.v1

Table 3. quantitative results on key-taste compounds (T) and quality markers (Q) of selected infusions. Amounts are expressed in mg/L and obtained by external calibration tecnique with LC-UV/DAD analysis. Results are provided as mean of three infusions and two analytical replicates. Taste threshold (TT) is expressed in μg/L.

Compound name	TT (μg/L)	Ceylon	Assam	Azores	Darjeeling Castelton	Darjeeling Testa Valley	Kenya	Yunnan
Epigallocatechin^{Q} (EGC)	159	72.38	29.24	16.09	25.27	71.08	38.62	16.61
Catechin^{Q} (C)	119	10.87	8.05	4.42	7.22	13.03	11.54	9.76
Epicatechin^{Q} (EC)	270	41.97	19.07	12.99	14.16	25.80	24.90	19.84
Epigallocatechingallate^{T,Q} (EGCG)	87.0	85.35	24.09	11.88	63.52	177.23	27.18	11.23
Gallocatechingallate^{Q} (GCG)	179	8.64	5.49	2.84	3.16	4.73	3.54	5.44
Epicatechingallate^{Q} (ECG)	115	41.40	21.90	7.92	29.68	39.79	20.31	18.80
Catechingallate^{Q} (CG)	239	5.37	5.18	3.07	5.03	3.93	3.59	2.76
Theaflavin^{Q} (TF)	9.00	7.55	4.39	2.92	2.25	2.16	4.50	2.33
Theaflavin-3-gallate^{Q} (TF3)	10.7	8.81	7.77	3.70	3.17	2.89	5.90	3.65
Theaflavin-3'-gallate^{Q} (TF3')	10.7	5.28	6.64	3.76	3.47	>LOQ	4.92	3.67
Theaflavin-3,3'-gallate^{Q} (TF3-3')	11.3	6.38	7.96	>LOQ	3.40	>LOQ	4.77	3.59
Myricetin-3-o-galactoside^{T} (M-3-o-gal)	1.3	2.39	>LOQ	0.28	0.42	0.32	1.51	>LOQ
Myricetin-3-o-glucoside^{T} (M-3-o-gluc)	1.0	4.10	0.24	0.37	0.85	1.02	1.33	>LOQ
Quercetin-3-o-rutinoside^{T} (Rutin)	0.0009	19.58	6.68	6.39	5.16	6.34	9.55	7.63
Quercetin-3-o-galactoside^{T} (Q-3-o-gal)	0.20	4.31	2.62	1.60	1.90	2.39	4.01	1.82
Quercetin-3-o-glucoside^{T} (Q-3-o-gluc)	0.30	9.42	3.93	1.01	0.77	1.12	5.53	2.68
Kaempferol-3-o-rutinoside^{T} (K-3-o-rut)	0.15	8.31	3.11	4.56	1.98	3.83	6.89	3.91
Kaempferol-3-o-glucoside^{T} (K-3-o-gluc)	0.30	4.19	1.32	1.07	0.62	1.24	3.50	1.43
Caffeine^{T}	97.1	250.50	290.30	143.70	241.52	265.88	263.36	228.03

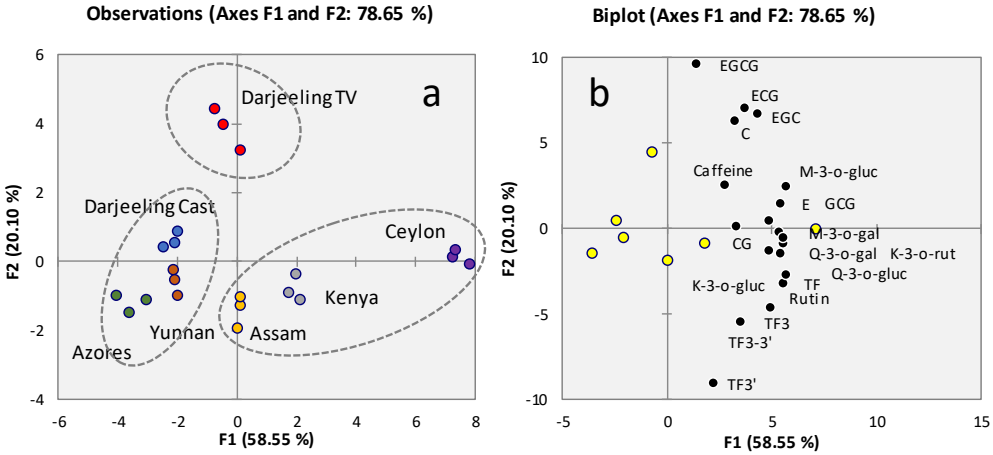

Figure 3. PCA results. (**3a**) scores plot on the first and the second principal components (F1-F2 plane) based on key-taste and quality compounds quantitative data determined in infusions prepared from tea leaves (Table 3). For Ceylon teas an average profile obtained from four lots was considered; (**3b**) loadings plot with the distribution of quantified analytes and centroids (yellow circles) for analyzed teas.

The quantitative distribution of monitored will have an impact on their taste and trigeminal perception. *Flavan-3-ols*, originally present in tea leaves, can undergo substantial changes during post-harvest treatments (oxidative phenomena during fermentation) leading to the formation of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0707.v1

high molecular weight dimeric (*theaflavins*) and oligomeric (*thearubigins*) derivates. For this reason, in fermented black tea, quantitative differences in *flavan-3-ols* and *theaflavins* can be ascribed to both the geographical origin and the technological processing. Conversely, the levels of *flavonols* and *caffeine* remain unchanged during tea manufacturing and thus, their variations are mostly influenced by the origin [3,40].

The high amount of *flavan-3-ols* but the low content of *theaflavins* in Darjeeling Testa Valley infusion suggests lower oxidation during post-harvest treatment which preserves the original amount of these analytes. On the other hand, teas from Ceylon, Assam, and Kenya due to the higher levels of dimeric analytes are likely more fermented. In addition, *theaflavins* are responsible for the characteristic dark orange-red color of black tea infusions, a piece of evidence that was confirmed by a visual inspection of the infusions. Indeed, teas from Darjeeling had a light yellow color due to the low concentration of these markers, in contrast with Ceylon tea. Tea from the Azores confirms to be characterized by a poor non-volatile fraction, for all compounds, suggesting a low astringency capacity.

Caffeine is present in similar concentrations in all samples (230-260 mg/L), with the only exception of Azores tea (143 mg/L); this outcome is probably related to the standardization of the infusion process (incomplete extraction), which does not allow its exhaustive extraction from the matrix.

2.5. Taste Blueprinting by AI tasting based on sensomics

In general, it can be assumed that the infusions with higher concentrations of *flavan-3-ols* and *flavonol-3-o-glycosides* (Ceylon and Darjeeling Testa Valley) are described by a more intense astringent sensation; however, to objectify the real contribution on the overall taste perception of the seven *flavan-3-ols* and the seven *flavonol-3-o-glycosides*, beyond the other markers, concentration data were associated with taste thresholds (**Table 3**).

Figure 4 visualizes as spider diagrams the dose over threshold (DoT, ratio of the concentration of each compound *vs.* its taste threshold) values, *i.e.*, the taste blueprint profile of tea infusions. Key tastants are associated with their taste descriptors, *e.g.* bitter, p. (puckering) astringency, m.d/v (mouth drying/velvety) astringency.

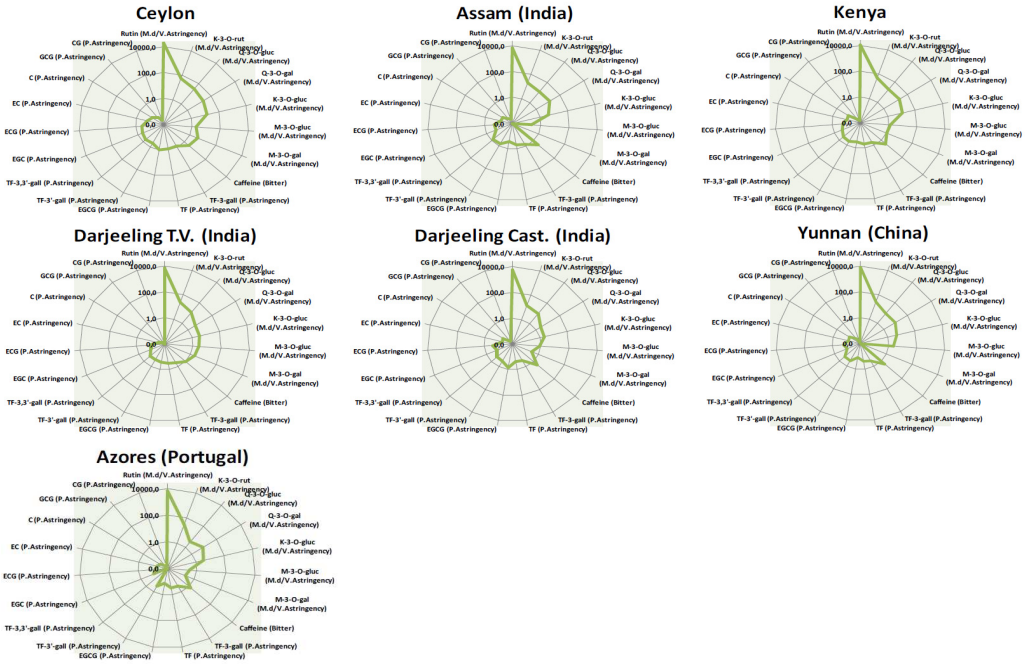

Figure 4. Spider diagrams with Dose over Threshold values (DoTs) illustrate the contribution of each key taste and quality marker to the sensory perception of selected teas. DoTs values are

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0707.v1

reported in logarithmic scale. Ceylon tea taste profile is reported as average values deriving from four different lots.

Within the entire class of monitored polyphenols, *flavonol-3-ol glycosides* are the most relevant for the definition of the taste blueprint, evoking a strong mouth-drying and velvety-like sensation [23,42]. In particular, *quercetin-3-o-rutinoside*, known as *rutin*, is the most representative marker, followed by *kaempferol-3-o-rutinoside* and the other derivates of *quercetin* (glucoside and galactoside). *Rutin* likely has a strong influence on the taste profile (upper part of the spider diagrams) since it is characterized by DoT values of three orders of magnitude higher than those of *flavan-3-ols* (very low taste thresholds, e.g. from 0.0009 mg/L for *rutin* to 0.30 mg/L for glucosidic compounds of *quercetin* and *kaempferol*). In addition, an important although not primary role for taste perception is also played by *flavan-3-ols*, especially those esterified with gallic acid; *epigallocatechingallate* (EGCG) is the most abundant *catechin* present in black teas (85.35 mg/L in Ceylon infusion, 177.23 mg/L in Darjeeling Testa Valley tea) [14,21]. The bitterness of the beverage is mainly related to the high content of caffeine which however shows a low DoT value, being characterized by high TT.

As a final consideration, the higher the concentration of flavonoidic markers is, the more intense the astringency sensation after tea consumption is; this outcome can be expected for tea samples from Ceylon, Darjeeling T.V., Assam and Kenya.

3. Materials and Methods

3.1. Reference compounds and solvents

Pure reference compounds for odorants and key aromas identity confirmation and quantitation [(E)-2-nonenal, (E,E)-2,4-nonadienal, (E,E)-2,4-decadienal, 3-methyl butanal, 2-methyl butanal, hexanal, phenyl acetaldehyde, (Z)-4-heptenal, linalool, (E,Z)-2,6-nonadienal, β-damascenone, β-ionone, geraniol, phenyl acetaldehyde) and n-alkanes (n-C7 to n-C25) for (I^{T}_{s}) determination were from Merck (Milan, Italy).

Pure reference compounds for quantitative determination of key tastants and quality markers [(-)-epigallocatechin, (+)-catechin, (-)-epigallocatechingallate, (-)-epicatechin, (-)-gallocatechingallate, (-)-epicatechingallate, (-)-catechingallate, theaflavin, theaflavin-3-gallate, theaflavin-3'-gallate, and theaflavin-3,3'-digallate] were supplied by Merck.

Tea extract from black tea (≥ 80% *theaflavins* and *theaflavin gallates*-basis) used as a reference for the theaflavins external standard calibration, was purchased from Fluka Analytical. Reference standard compounds of *myricetin-3-O-glucoside, myricetin-3-O-galactoside, quercetin-3-O-rutinoside* (rutin), *quercetin-3-O-galactoside, quercetin-3-O-glucoside, kaempferol-3-O-rutinoside, kaempferol-3-O-glucoside,* and *caffeine* were supplied by Extrasynthese (Lyon, France).

Internal standardization (ISTD) for volatile compounds was with *1,4-dibromobenzene* from Merck.

Solvents were all LC grade, from Merck: *acetone* (purity 99.5%), *acetonitrile* ACN (purity 99.9%), and *methanol* (purity 99.9%). *Water* used to prepare stock solutions, tea infusions, and LC mobile phases was obtained by a Milli-Q RG system (Millipore, Molsheim, France) in agreement with the ISO 9002 Quality Systems Standards.

3.2. Reference solutions and calibration mixtures

Standard stock solutions at 10 mg/mL, containing pure aroma reference compounds and ISTD, were prepared in acetone and stored in sealed vials at -18°C for a two weeks maximum. Standard spiking mixtures, adopted for quantification (Standard Addition SA procedure), were prepared by diluting standard stock solutions of each compound to different final concentrations in the range 5-500 ng/µL (5, 10, 25, 50, 100, 250, 500 ng/µL). The ISTF *1,4-dibromobenzene* was added to all calibration solutions at 120 ng/µL.

 doi:10.20944/preprints202401.0707.v1

For taste compounds, standard stock solutions were prepared in a water/ACN (9:1 v/v) mixture or methanol at the concentration of 1 mg/mL. Standard stock solutions were sealed and stored at -18°C for two weeks. Standard calibration solutions at 100, 50, 25, 10, 5, 2.5, and 1 ng/µL were prepared by diluting suitable amounts of standard stock solutions in the same solvent (water for *catechins* and *theaflavins*, methanol for *flavonols*) and stored at -18 °C until analyzed.

3.3. Tea infusions: samples and preparation

Four different lots of fermented dried tea leaves of homogeneous particle size from Ceylon (*Flowery Orange Pekoe* - FOP) were kindly supplied by Soremartec Italia Srl (Alba-Cuneo, Italy). Tea samples from India (Assam, Darjeeling Testa Valley, Darjeeling Castleton) were all graded as *Golden Flowery Orange Pekoe* – GFOP, from those from Portugal (Azores), China (Yunnan), and Kenya OP grading was not available. Excluding the Ceylon teas, all the others were bought in a specialized tea shop in Turin (Italy).

Infusions were prepared by following the EMA/HMPC/283630/2012 Committee on Herbal Medicinal Products (HMPC) protocol: 3.0 g of dried leaves were suspended in 300 mL of ultrapure boiling water for 90 seconds and then filtered. Infusions were left to reach ambient temperature in closed glass flasks and directly analyzed by IS-SPME followed by GC-MS for aroma blueprinting. For taste and quality markers quantitative profiling, infusions were filtered through 0.45 mm, 25 mm nylon membrane syringe filters (Agilent, Little Falls, DE, USA) and then analyzed by LC-UV/DAD.

3.4. Automated in-solution solid-phase microextraction: devices and sampling conditions

SPME sampling devices and fibers were from Merck Supelco (Bellefonte, PA, USA). A Divinylbenzene/Carboxen/Polydimethylsiloxane (DVB/CAR/PDMS) d_f 50/30 µm, 2 cm long fiber was chosen, and conditioned before use as recommended by the manufacturer.

Sampling conditions were set as follows: 20 mL of tea infusions were submitted to in-solution (IS) sampling for the quantitative profiling of key aroma compounds. In particular, for SA quantitation, 20 mL of tea infusion was sealed in a 20 mL headspace vial and spiked with suitable volumes of standard spiking solutions for each calibration level. Sampling was performed by exposing the SPME device to the tea infusion for 40 minutes at 50°C.

3.5. Automated IS-SPME-GC-MS instrumental set-up and analysis conditions

Automated SPME for IS sampling was performed using a MPS-2 multipurpose sampler (Gerstel, Mülheim a/d Ruhr, Germany) online integrated with an Agilent 7890 GC unit coupled to an Agilent 5977C MS spectrometer equipped with a High-Efficiency Source (HES) (Agilent, Little Falls, DE, USA) operating in EI mode at 70 eV. The transfer line was set to 280°C. A HES Tune was used and the scan range was set to m/z 35-350 (full scan acquisition) with a scanning rate of 1,000 amu/s, to obtain a suitable number of data points for each chromatographic peak to ensure reliable identification. Analyses were also acquired in Single Ion Monitoring (SIM) mode for the accurate quantification of selected (key-aroma) markers. A SE52 capillary column (95% polydimethylsiloxane, 5% phenyl – 30m × 0.25 mm d_c × 0.25 µm d_f) from Mega (Legnano, Milan, Italy) was used.

For Linear Retention Indexes (I^T_s) determination, the *n*-alkanes liquid sample solution (100 mg/L) was injected using the MPS-2 multipurpose sampler (Gerstel) under the following conditions: injection mode: split, split ratio 1:50, injector temperature 250°C, injection volume 1 µL.

The analytes sampled by IS-SPME were thermally desorbed from the fiber for 5 min, directly into the GC injector, under the following conditions: injection mode: split, split ratio 1:20, injector temperature 250°C. The carrier gas was helium, at a constant flow rate of 1.0 mL/min. The temperature program was 40°C (1 min) to 170°C at 3°C/min and to 260°C at 10°C/min (5 min).

3.6. LC-UV/DAD instrumental set-up and analysis conditions

LC-UV/DAD analyses were carried out on a Spectra System (SCM1000, P4000, AS3000) provided with a Spectra System UV6000LP Diode Array Detector (Thermo Fisher Scientific, Waltham, MA,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 January 2024 doi:10.20944/preprints202401.0707.v1

USA). LC column was a Supelco Ascentis® Express RP-C18 (150 mm × 4.6 mm; spherical particles with fused core® technology, 90 Å, 2.7µm) from Merck; a pre-column (5 mm × 4.6 mm; spherical particles with fused core® technology, 90 Å, 2.7µm) was installed to preserve the analytical column for pollution.

Operative conditions were as follows: injection volume 8 µL; detection wavelengths, 280 nm for flavan-3-ols and caffeine, 350 nm for flavonol-3-O-glycosides, 380 nm for theaflavins; mobile phases, (A) H_2O + 0.1% formic acid, (B) ACN + 0.1% formic acid; flow rate, 1.10 mL/min; mobile phase program, from 95% H_2O (4.7 min) to 85% H_2O (15.8 min), to 75% H_2O (9.2 min), hold for 12 min, then to 65% H_2O, and then to 100% ACN, hold for 3 min. Before re-injection, the LC system was stabilized for at least 5 min.

3.7. Data acquisition and data processing

GC-MS data were acquired by MassHunter WorkStation (Agilent Technologies) while LC analyses data acquisition and data handling were performed with ChromQuest 2.51 software (Thermo Fisher Scientific). Post-processing was by XLSTAT 2014 by Addinsoft (New York, USA).

3.8. Method validation parameters

Method validation was designed according to Eurachem Guidelines [43] and performance quality evaluation was based on reference parameters of Commission Implementing Regulation (EU) 2021/808 of 22 March 2021 for quantitative methods in food applications [44]. Infusions prepared from Ceylon tea (named Quality Controls QCs) were analyzed by IS-SPME-GC-MS and LC-UV/DAD and quantitative results collected over a period of six weeks were used for repeatability (intra-week) and intermediate precision (inter-week) assessment. Results on precision based on chromatographic responses (normalized over the internal standard for GC-MS and absolute areas for LC-UV/DAD) are reported in **Supplementary Table 1 (ST1)** expressed as percent coefficient of variation (CV%) together with chromatographic information on retention times (t_R min), Target/Qualifier Ions *m/z* (GC-MS), detection wavelengths nm (LC-UV/DAD), calibration functions and determination coefficients (R^2), and accuracy results from spiked QC infusions at two levels expressed as percent recovery % Rec. Accuracy was tested at +10 and +25 µg/L for aroma compounds and +10 and +25 mg/L for tastants and quality markers.

3.9. Method validation results: precision and accuracy

For aroma compounds results on quantified analytes referred to a fairly good precision with inter-week CV% never exceeding 25% [results reported in **Supplementary Table 1 (ST1)**]; this outcome was consistent with validation parameters included in the Commission Implementing Regulation (EU) 2021/808 in the case of analytes present at ppb levels (µg/L). Within aroma compounds, those with the best performances over the six weeks validation period were: *hexanal* (10.09%), *(E,Z)-2,6-nonadienal* (12.21%), *phenylacetaldehyde* (13.54%), and *geraniol* (14.25%). The most volatile odorants, *i.e*. Strecker aldehydes *3-methyl butanal* and *2-methyl butanal* reported CV% values close to the acceptability threshold with 24.56 and 22.13 % respectively.

In the case of taste active compounds results concerning precision were all within the acceptable range established by Commission Implementing Regulation (EU) 2021/808; CV% were always below 18%. The higher value was obtained for *flavan-3-ols* with an inter-week CV% of 14.55% while a very good intermediate precision was for *theaflavins* (5.16%), *caffeine* (6.0%), and *flavonols* derivates (10.60%).

Accuracy was determined by spiking suitable amounts of analytes within the working range. Percent Rec (% Rec) never exceeded the range ±20% with better performances, as expected, for non-volatile compounds (average Rec% 103).

4. Conclusions

The integrated analytical strategy proposed in this study enabled the successful definition of the aroma and taste blueprint of black tea infusions. Acting as a SEBES system, the approach supports the prediction of flavor features of teas without the need for human panel assessment. Moreover, thanks to the full automation and absence of laborious sample preparation, both methodologies can efficiently support a sustainable yet green analytical control with high throughput of samples as that requied in industrial laboratories.

The selected black teas showed a large dynamic range of variation of key-analytes concentrations, the integrated methodology is therefore suitable not only for standard infusion characterization but also to support industrial development of alternative extraction processes facilitating product benchmarking against a quality reference.

Experimental results demonstrate the tight relationship between the distribution of potent odorants (volatiles) and taste and quality markers (non-volatiles), a phenomenon taking place already during tea manufacturing; in fact, tea fermentation is characterized by the atmospheric oxidation of tea flavan-3-ols catalyzed by an endogenous catechol oxidase. In turn, these compounds are strongly oxidizing agents which can react with other chemicals (amino acids, carotenoids, unsaturated fatty acids, etc.) leading to the formation of the typical black tea aroma and its particular flavor profile [45,46].

Although limited in its representativeness, the sample set covered different geographical origins; results showed a quite good clustering of samples according to origin. This outcome suggests an additional feature related to origin qualification and discrimination, a characteristic that has relevance to tea price and value.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org, **Supplementary Table 1 ST1:** Analytes subjected to validation for Ai flavor blueprinting of tea. Analytes are listed together with chromatographic information on retention times (t_R min), Target/Qualifier Ions *m/z* (GC-MS), detection wavelengths nm (LC-UV/DAD), calibration functions and determination coefficients (R^2), precision (repeatability and intermediate precision) CV%, and accuracy results from spiked QC infusions at two levels expressed as percent recovery % Rec. Accuracy was tested at +10 and +25 µg/L for aroma compounds and +10 and +25 mg/L for tastants and quality markers.

Author Contributions: Conceptualization, C.B. and C.C.; methodology, C.C. and F.M.; validation, F.M., A.C and C.C.; formal analysis, F.M.; resources, C.B. and C.C.; data curation, F.M., A.C., A.F., and F.T.; writing—original draft preparation, F.M., C.C., and A.C.; writing—review and editing,C.C., A.C., A.F., and F.T.; visualization, F.M., A.C. and E.L.; supervision, C.C, E.L., and C.B.; funding acquisition, C.B. and C.C. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: This research was supported by Soremartec Italia srl (Alba, CN Italy) with the project entitled "Advanced analytical approaches for *sensomic* investigations of high quality food matrices of vegetable origin".

Data Availability Statement: Data will be made available on request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, L.; Cao, Q.Q.; Granato, D.; Xu, Y.Q.; Ho, C.T. Association between chemistry and taste of tea: A review. *Trends Food Sci. Technol.* **2020**, *101*, 139–149, doi:10.1016/j.tifs.2020.05.015.
2. Morimoto, J.; Rosso, M.C.; Kfoury, N.; Bicchi, C.; Cordero, C.; Robbat, A.; Cialiè Rosso, M.; Kfoury, N.; Bicchi, C.; Cordero, C.; et al. Untargeted/targeted 2D gas chromatography/mass spectrometry detection of the total volatile tea metabolome. *Molecules* **2019**, *24*, 1–14, doi:10.3390/molecules24203757.
3. Magagna, F.; Cordero, C.; Cagliero, C.; Liberto, E.; Rubiolo, P.; Sgorbini, B.; Bicchi, C. Black tea volatiles fingerprinting by comprehensive two-dimensional gas chromatography – Mass spectrometry combined with high concentration capacity sample preparation techniques: Toward a fully automated sensomic assessment. *Food Chem.* **2017**, *225*, 276–287, doi:10.1016/j.foodchem.2017.01.003.
4. Alcázar, A.; Ballesteros, O.; Jurado, J.M.; Pablos, F.; Martín, M.J.; Vilches, J.L.; Navalón, A. Differentiation of green, white, black, Oolong, and Pu-erh teas according to their free amino acids content. *J. Agric. Food Chem.* **2007**, *55*, 5960–5965, doi:10.1021/jf070601a.

5. Zhao, Y.; Chen, P.; Lin, L.; Harnly, J.M.; Yu, L. (Lucy); Li, Z. Tentative identification, quantitation, and principal component analysis of green pu-erh, green, and white teas using UPLC/DAD/MS. *Food Chem.* **2011**, *126*, 1269–1277, doi:10.1016/j.foodchem.2010.11.055.
6. Yang, Z.; Baldermann, S.; Watanabe, N. Recent studies of the volatile compounds in tea. *Food Res. Int.* 2013, *53*, 585–599.
7. Atoui, A.K.; Mansouri, A.; Boskou, G.; Kefalas, P. Tea and herbal infusions: Their antioxidant activity and phenolic profile. *Food Chem.* **2005**, doi:10.1016/j.foodchem.2004.01.075.
8. Del Rio, D.; Calani, L.; Scazzina, F.; Jechiu, L.; Cordero, C.; Brighenti, F. Bioavailability of catechins from ready-to-drink tea. *Nutrition* **2010**, *26*, 528–533, doi:10.1016/j.nut.2009.06.013.
9. Del Rio, D.; Calani, L.; Cordero, C.; Salvatore, S.; Pellegrini, N.; Brighenti, F. Bioavailability and catabolism of green tea flavan-3-ols in humans. *Nutrition* **2010**, *26*, 1110–1116, doi:10.1016/j.nut.2009.09.021.
10. Wheeler, D.S.; Wheeler, W.J. The Medicinal Chemistry of Tea. *Drug Dev. Res.* 2004.
11. Prescott, J. Multisensory processes in flavour perception and their influence on food choice. *Curr. Opin. Food Sci.* 2015, *3*, 47–52.
12. Bondarovich, H.A.; Giammarino, A.S.; Renner, J.A.; Shephard, F.W.; Shingler, A.J.; Gianturco, M.A. VOLATILES IN TEA Some Aspects of the Chemistry of Tea. A Contribution to the Knowledge of the Volatile Constituents. *J. Agric. Food Chem.* **1967**, *15*, 36–47, doi:10.1021/jf60149a011.
13. Kawakami, M.; Ganguly, S.N.; Baneijee, J.; Kobayashi, A. Aroma Composition of Oolong Tea and Black Tea by Brewed Extraction Method and Characterizing Compounds of Darjeeling Tea Aroma. *J. Agric. Food Chem.* **1995**, *43*, 200–207, doi:10.1021/jf00049a037.
14. Kumazawa, K.; Masuda, H. Change in the flavor of black tea drink during heat processing. *J. Agric. Food Chem.* **2001**, *49*, 3304–3309, doi:10.1021/jf001323h.
15. Lv, H.P.; Zhong, Q.S.; Lin, Z.; Wang, L.; Tan, J.F.; Guo, L. Aroma characterisation of Pu-erh tea using headspace-solid phase microextraction combined with GC/MS and GC-olfactometry. *Food Chem.* **2012**, doi:10.1016/j.foodchem.2011.07.135.
16. Hu, S.; He, C.; Li, Y.; Yu, Z.; Chen, Y.; Wang, Y.; Ni, D. The formation of aroma quality of dark tea during pile-fermentation based on multi-omics. *LWT* **2021**, doi:10.1016/j.lwt.2021.111491.
17. Togari, N.; Kobayashi, A.; Aishima, T. Pattern recognition applied to gas chromatographic profiles of volatile components in three tea categories. *Food Res. Int.* **1995**, doi:10.1016/0963-9969(95)00029-1.
18. Wang, C.; Lv, S.; Wu, Y.; Gao, X.; Li, J.; Zhang, W.; Meng, Q. Oolong tea made from tea plants from different locations in Yunnan and Fujian, China showed similar aroma but different taste characteristics. *Springerplus* **2016**, doi:10.1186/s40064-016-2229-y.
19. Zhang, L.; Zeng, Z.; Zhao, C.; Kong, H.; Lu, X.; Xu, G. A comparative study of volatile components in green, oolong and black teas by using comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry and multivariate data analysis. *J. Chromatogr. A* **2013**, *1313*, 245–252, doi:10.1016/j.chroma.2013.06.022.
20. Vitzthum, O.G.; Werkhoff, P.; Hubert, P. New Volatile Constituents of Black Tea Aroma. *J. Agric. Food Chem.* **1975**, doi:10.1021/jf60201a032.
21. Schuh, C.; Schieberle, P. Characterization of the key aroma compounds in the beverage prepared from Darjeeling black tea: Quantitative differences between tea leaves and infusion. *J. Agric. Food Chem.* **2006**, *54*, 916–924, doi:10.1021/jf052495n.
22. Cordero, C.; Kiefl, J.; Reichenbach, S.E.; Bicchi, C. Characterization of odorant patterns by comprehensive two-dimensional gas chromatography: A challenge in omic studies. *TrAC Trends Anal. Chem.* **2019**, *113*, 364–378, doi:10.1016/j.trac.2018.06.005.
23. Scharbert, S.; Hofmann, T. Molecular Definition of Black Tea Taste by Means of Quantitative Studies, Taste Reconstitution, and Omission Experiments. *J. Agric. Food Chem.* **2005**, *53*, 5377–5384, doi:10.1021/jf050294d.
24. Borse, B..; Jagan Mohan Rao, L.; Nagalakshmi, S.; Krishnamurthy, N. Fingerprint of black teas from India: identification of the regio-specific characteristics. *Food Chem.* **2002**, *79*, 419–424, doi:10.1016/S0308-8146(02)00191-7.
25. Wang, L.F.; Lee, J.Y.; Chung, J.O.; Baik, J.H.; So, S.; Park, S.K. Discrimination of teas with different degrees of fermentation by SPME-GC analysis of the characteristic volatile flavour compounds. *Food Chem.* **2008**, doi:10.1016/j.foodchem.2007.12.054.
26. Flaig, M.; Qi, S.C.; Wei, G.; Yang, X.; Schieberle, P. Characterisation of the key aroma compounds in a Longjing green tea infusion (Camellia sinensis) by the sensomics approach and their quantitative changes during processing of the tea leaves. *Eur. Food Res. Technol.* **2020**, *246*, 2411–2425, doi:10.1007/s00217-020-03584-y.
27. Flaig, M.; Qi, S.; Wei, G.; Yang, X.; Schieberle, P. Characterization of the Key Odorants in a High-Grade Chinese Green Tea Beverage (Camellia sinensis; Jingshan cha) by Means of the Sensomics Approach and Elucidation of Odorant Changes in Tea Leaves Caused by the Tea Manufacturing Process. *J. Agric. Food Chem.* **2020**, doi:10.1021/acs.jafc.0c01300.

28. Nicolotti, L.; Mall, V.; Schieberle, P. Characterization of Key Aroma Compounds in a Commercial Rum and an Australian Red Wine by Means of a New Sensomics-Based Expert System (SEBES) - An Approach to Use Artificial Intelligence in Determining Food Odor Codes. *J. Agric. Food Chem.* **2019**, *67*, 4011–4022, doi:10.1021/acs.jafc.9b00708.
29. Cordero, C.; Canale, F.; Rio, D. Del; Bicchi, C. Identification, quantitation, and method validation for flavan-3-ols in fermented ready-to-drink teas from the Italian market using HPLC-UV/DAD and LC-MS/MS. *J. Sep. Sci.* **2009**, *32*, 3643–3651, doi:10.1002/jssc.200900369.
30. Dunkel, A.; Steinhaus, M.; Kotthoff, M.; Nowak, B.; Krautwurst, D.; Schieberle, P.; Hofmann, T. Nature's chemical signatures in human olfaction: A foodborne perspective for future biotechnology. *Angew. Chemie - Int. Ed.* **2014**, *53*, 7124–7143, doi:10.1002/anie.201309508.
31. Ruosi, M.R.; Cordero, C.; Cagliero, C.; Rubiolo, P.; Bicchi, C.; Sgorbini, B.; Liberto, E. A Further Tool To Monitor the Coffee Roasting Process: Aroma Composition and Chemical Indices. *J. Agric. Food Chem.* **2012**, *60*, 11283–11291, doi:10.1021/jf3031716.
32. Nicolotti, L.; Cordero, C.; Cagliero, C.; Liberto, E.; Sgorbini, B.; Rubiolo, P.; Bicchi, C. Quantitative fingerprinting by headspace-Two-dimensional comprehensive gas chromatography-mass spectrometry of solid matrices: Some challenging aspects of the exhaustive assessment of food volatiles. *Anal. Chim. Acta* **2013**, *798*, 115–125, doi:10.1016/j.aca.2013.08.052.
33. Cordero, C.; Cagliero, C.; Liberto, E.; Nicolotti, L.; Rubiolo, P.; Sgorbini, B.; Bicchi, C. High concentration capacity sample preparation techniques to improve the informative potential of two-dimensional comprehensive gas chromatography–mass spectrometry: Application to sensomics. *J. Chromatogr. A* **2013**, *1318*, 1–11, doi:10.1016/j.chroma.2013.09.065.
34. Stilo, F.; Cordero, C.; Sgorbini, B.; Bicchi, C.; Liberto, E. Highly Informative Fingerprinting of Extra-Virgin Olive Oil Volatiles: The Role of High Concentration-Capacity Sampling in Combination with Comprehensive Two-Dimensional Gas Chromatography. *Separations* **2019**, *6*, 34, doi:10.3390/separations6030034.
35. Sgorbini, B.; Cagliero, C.; Liberto, E.; Rubiolo, P.; Bicchi, C.; Cordero, C. Strategies for Accurate Quantitation of Volatiles from Foods and Plant-Origin Materials: A Challenging Task. *J. Agric. Food Chem.* **2019**, *67*, 1619–1630, doi:10.1021/acs.jafc.8b06601.
36. Schieberle, P.; Hofmann, T. Mapping the combinatorial code of food flavors by means of molecular sensory science approach. In *Food Flavors: Chemical, Sensory and Technological Properties*; 2011; pp. 413–438 ISBN 9781439814925.
37. McLafferty, F.W. Wiley Registry of Mass Spectral Data, 12th Edition 2020.
38. NIST/EPA/NIH Mass Spectral Library with Search Program Data Version: NIST v17.
39. Sanderson, G.W.; Selvendran, R.R. The organic acids in tea plants. A study of the non-volatile organic acids separated on silica gel. *J. Sci. Food Agric.* **1965**, *16*, 251–258, doi:10.1002/jsfa.2740160504.
40. Sai, V.; Chaturvedula, P.; Prakash, I. The aroma , taste , color and bioactive constituents of tea. *J. Med. Plants Res.* **2011**, *5*, 2110–2124, doi:10.1016/j.ijcard.2016.09.092.
41. Preininger, M.; Grosch, W. Evaluation of Key Odorants of the Neutral Volatiles of Emmentaler Cheese by the Calculation of Odour Activity Values. *Leb. und-Technologie* **1994**, *27*, 237–244, doi:10.1006/fstl.1994.1048.
42. Hofmann, T. Identification of the Astringent Taste Compounds in Black Tea Infusions by Combining Instrumental Analysis and Human Bioresponse Identification of the Astringent Taste Compounds in Black Tea Infusions by Combining Instrumental Analysis and Human. *J. Agric. Food Chem.* **2016**, *52*, 3498–3508, doi:10.1021/jf049802u.
43. Eurachem *Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics.*; 2014; ISBN 0-94948926-12-0.
44. Commission, E. Commission Implementing Regulation (EU) 2021/808 of 22 March 2021 on the performance of analytical methods for residues of pharmacologically active substances used in food-producing animals and on the interpretation of results as well as on the methods to. *Off. J. Eur. Union* **2021**, *180*, 84–109.
45. Kinlin, T.E.; Muralidhara, R.; Pittet, A.O.; Sanderson, A.; Walradt, J.P. Volatile Components of Roasted Filberts. *J. Agric. Food Chem.* **1972**, *20*, 1021–1028.
46. Stilo, F.; Tredici, G.; Bicchi, C.; Robbat, A.; Morimoto, J.; Cordero, C. Climate and Processing Effects on Tea (Camellia sinensis L. Kuntze) Metabolome: Accurate Profiling and Fingerprinting by Comprehensive Two-Dimensional Gas Chromatography/Time-of-Flight Mass Spectrometry. *Molecules* **2020**, *25*, 2447, doi:10.3390/molecules25102447.