Non-targeted analysis by LC–MS of major metabolite changes during the oolong tea manufacturing in New Zealand

Fraser, Karl; Lane, Geoff A.; Otter, Don E.; Harrison, Scott James; Quek, Siew-Young; Hemar, Yacine; Rasmussen, Susanne

Published in:
Food Chemistry

Link to article, DOI:
10.1016/j.foodchem.2013.11.054

Publication date:
2014

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Non-targeted analysis by LC–MS of major metabolite changes during the oolong tea manufacturing in New Zealand

Karl Fraser a,b,⇑, Geoff A. Lane a, Don E. Otter a, Scott J. Harrison a,1, Siew-Young Quek b, Yacine Hemar b, Susanne Rasmussen a

aAgResearch, Private Bag 11008, Tennent Drive, Palmerston North 4410, New Zealand
bSchool of Chemical Sciences, University of Auckland, Auckland 1142, New Zealand

Article history:
Received 1 August 2013
Received in revised form 4 October 2013
Accepted 11 November 2013
Available online 18 November 2013

Abstract
Oolong tea is a semi-fermented tea that is partially oxidised during the manufacturing process to create a product unique in composition. In this study, we investigated the potential of non-targeted LC–MS with two complementary chromatographic modes to provide a “comprehensive and unbiased” view of biochemical compositional changes occurring during oolong tea manufacturing in New Zealand. Tea leaf samples from throughout the manufacturing/fermentation process during three different harvest periods (spring, summer and autumn) were analysed by four different LC–MS streams. Principal component analysis revealed the de-greening stage of the manufacturing process was responsible for major changes in the biochemical profile, with the methodology detecting changes in a wide range of metabolites of differing polarities, such as flavonoids, nucleosides and primeverosides. Changes during the fermentation phase of the manufacturing process were less marked, however significant increases in levels of free amino acids, a hydroxyjasmonic acid and related metabolites were observed.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Oolong tea is a semi-fermented/oxidised tea made from a hot water infusion of the leaves of partially oxidised (10–70%) Camellia sinensis. It is a traditional Chinese tea, dating back hundreds of years, that is now becoming more popular outside Asia due to its distinctive attributes, such as its aroma and flavour, which are generated during the manufacturing process (Chen et al., 2011). Many of these major attributes of tea are controlled, to a large extent, by the amount of post-harvest fermentation/oxidation allowed during manufacturing, and this typically characterises the type of tea produced, i.e. green tea (minimal oxidation) and black tea (fully oxidised) (Harbowy, Balentine, Davies, & Cai, 1997). Along with the well-known variables, such as cultivar, region, climate and soil type (Chen et al., 2011; Lee et al., 2010), the less strictly defined oxidation level for oolong tea gives rise to a range of oolong products which can vary considerably in composition.

The fermentation/oxidation process is catalysed by enzymes present in the leaf. For example, the dark colour of black tea is partly due to polyphenolic theaflavins and thearubigins formed from flavanols by enzymes, such as polyphenol oxidase (Wang & Ho, 2009). Many other reactions occur during the fermentation period, such as protein breakdown and lipid oxidation, which also contribute to product attributes, such as aroma, colour and taste (Chen et al., 2011; Tomlins & Mashingaidze, 1997). To control these reactions the post-harvest fermentation is monitored and stopped at the appropriate level by heating the leaves rapidly to 300 °C for a few min (“de-greening” step). As the composition of the final product is directly influenced by the timing of the de-greening step, any significant variation in the level of oxidation will affect the reproducibility and quality of the tea produced. Thus, the decision as to when to halt fermentation requires expert knowledge and is made by an experienced tea maker, or ‘tea master’, by frequently observing, smelling and tasting infusions of the leaves during the fermentation process (Bhattacharyya et al., 2007).

There are considerable changes in metabolite levels/profiles during the fermentation processes of tea production and as such there is a need for improved rapid and comprehensive analytical methodologies which will monitor and remove subjectivity from the process. The effect of fermentation during the manufacturing process of the less fermented oolong tea is not well documented. A recent study followed a single batch of leaves through different relative levels of fermentation (based on measuring the area of observed colour changes occurring to the leaf) to represent green, oo-
long and black tea (Kim, Goodner, Park, Choi, & Talcott, 2011). They reported decreases in the levels of flavanols, flavonoid glycosides and caffeine, particularly once fermentation levels achieved 40% or greater; and increases in thearubigins, theaflavins and most volatiles from the 20% fermented level and above. Few studies like Kim et al. (2011) have followed a single batch of plant material throughout the controlled oolong fermentation process, a critical requirement to understand the effects of the amount of processing and manufacturing techniques used during the processing and, fermentation parameters such as time and temperature on the overall metabolite profile.

The area of comprehensive non-targeted metabolite analysis has been applied to a wide variety of biological sample types, including foods and beverages (Cevallos-Cevallos, Reyes-De-Corcuera, Dtextendashier, Danyluk, & Rodrick, 2009). In recent years there has been an increase in the number of non-targeted studies, enabled by significant improvements in the analytical instrumentation that are capable of detecting an increasingly wide array of components. Of the instrumental analysis techniques employed for profiling tea, the most frequently used are nuclear magnetic resonance (NMR) (Lee et al., 2010, 2011) or mass spectrometry (MS) coupled to liquid chromatography (LC) (Fraser et al., 2012; Ku, Kim, Park, Liu, & Lee, 2010; van der Hooft et al., 2012), or gas chromatography (GC) (Frijdeevenh & Machan, 2011), or combining data from both GC–MS and LC–MS analyses (Kim et al., 2011; Ku et al., 2010; Yang et al., 2012). Compared to NMR, the MS based techniques are more sensitive, generally have a wider dynamic range, and due to readily available automation, can be set up to run hundreds of samples in a rapid manner. To date, most LC–MS based studies employ either reversed-phase (RP) for semi-nondesirable compounds or hydrophilic interaction chromatography (HILIC) for polar compounds, coupled to electrospray ionisation (ESI) interfaces on the MS, but not both on the same experimental sample set. Given the large amount of data generated, these broad non-targeted analyses require complex data processing software and multivariate chemometric tools to process the multidimensional data and mine biologically significant features (Theodoridis, Gika, & Wilson, 2011).

In this study, we have investigated the potential of comprehensive non-targeted analyses using two complementary LC phases (RP and HILIC) coupled with high resolution MS detection in positive and negative ESI to resolve and detect a wide polarity range of metabolites. We have then applied statistical data mining tools to detect metabolic changes in tea samples collected throughout a typical 36 h oolong tea manufacturing process (from freshly picked leaves, throughout the fermentation process and to the final product). The LC–MS data were processed with univariate and two multivariate chemometric methods, principal component analysis (PCA) and partial least squared discriminant analysis (PLS-DA), to monitor changes in metabolite composition during the manufacturing process. Samples were collected from the Zealong Tea Estate in Hamilton, New Zealand, at three harvest points/manufacturing dates over the season. The data analysis was carried out in two stages: (1) comparison of the pre and post de-greening step to examine effects of the high-temperature processing step on metabolite profiles and (2) comparison of samples throughout the fermentation stage (pre de-greening step) to identify key metabolic changes during the fermentation process. Distinctive ions altering in abundance during the manufacturing process have been further investigated and characteristic information acquired on significant metabolites, including tentative identifications where possible. The findings from this study have confirmed several previously reported metabolic changes observed during tea fermentation. In addition, we have detected several novel metabolic changes not previously reported in oolong tea manufacturing and been able to associate the observed changes with particular stages of manufacturing.

2. Materials and methods

2.1. Samples and reagents

Samples from the Zealong Tea Estate (Hamilton, New Zealand) taken throughout the oolong tea manufacturing process for a single batch of tea, from freshly harvested leaves to final product. Samples were collected on three separate harvest dates covering the southern hemisphere late-spring (November 2011 – early season harvest), mid-summer (January 2012 – mid-season harvest) and early autumn (March 2012 – late season harvest). On each of the three sampling occasions, 18 sequential samples were collected during the 36 h oolong production process. Table 1 summarises the sampling time points and key manufacturing steps occurring at these time points. At each sampling, the samples collected were immediately placed on dry ice to limit any further compositional change and once all 18 samples were collected they were transported on dry ice to the laboratory, freeze-dried to eliminate differences in sample weights due to moisture content, and stored at -20 °C until analysis.

The mobile phase modifiers, formic acid and ammonium formate; the internal standards d3-tyrosine and dichlorofluoroscein; and the other phenolics, nucelosides and amino acid standards listed below were purchased from Sigma–Aldrich Chemicals Co. (St Louis, MO). Ultrapure water was obtained from a Milli-Q® system (Millipore, Bedford, MA); Acetonitrile of Optima LC–MS grade was purchased from Thermo Fisher Scientific (Auckland, New Zealand).

2.2. Sample preparation

Triplicate tea infusions were prepared by adding 5 ml of boiling water to 50 mg (±0.5 mg) of ground tea (ground to fine particles of approximately tea bag grade) in a 15 ml tube as previously described (Fraser et al., 2012). The samples were spiked with 50 µl of internal standard solution containing 230 µg/ml d3-tyrosine and 230 µg/ml dichlorofluoroscein (in 50:50 acetonitrile–water (v/v)). A 0.5 ml aliquot was filtered through a 0.22 µm stainless steel filter into a vial and the filtrate was diluted with 0.5 ml ace- tonitrile and analysed by LC–MS.

2.3. Analytical procedure and MS conditions

The HILIC and RP analytical conditions have been previously reported in detail (Fraser et al., 2012, 2013). Briefly, the LC–MS instrument utilised was a Thermo Accela 1250 UHPLC system connected to a Thermo Exacta mass spectrometer with electrospray ionisation. Injection volumes of 2 µl of the tea extract and a column oven temperature of 25 °C were utilised for both chromatographic analyses. RP LC–MS analysis was performed using an Agilent RRHD SB-C18 column (150 mm × 2.1 mm, 1.8 µm) with a gradient elution of water containing 0.1% formic acid and acetoni- trile containing 0.1% formic acid at a 400 µl/min flow rate. HILIC LC–MS analysis was performed using a Merck polymeric bead based ZIC-pHILIC column (100 mm × 2.1 mm, 5 µm) with a grad- ient elution of acetonitrile containing 0.1% formic acid and water containing 16 mM ammonium formate at 250 µl/min flow rate. Samples were run in both positive and negative ionisation mode as separate chromatographic runs.

2.4. Data extraction and multivariate statistical analysis

Components eluting between 3 and 14 min for the RP analysis and 3 and 18 min for the HILIC analysis were extracted and aligned from the LC–MS data using PhenoAnalyzer (SpectralWorks Ltd,
3. Results and discussion

3.1. Non-targeted profiling and peak detection

In order to ensure the acquisition of information-rich non-targeted chromatographic profiles and to maximise the number of metabolites detected, minimal sample preparation and generic instrument parameter settings were employed. We then utilised previously reported RP (Fraser et al., 2013) and HILIC (Fraser et al., 2012) chromatography methods, along with positive and negative ESI ionisation modes, to give four separate LC–MS analysis streams (RP +ve, RP –ve, HILIC +ve and HILIC –ve). This ensured that we were covering a diverse range of physico-chemical properties of metabolites extracted and increasing the range of tea metabolites potentially detected in these samples. After combining chromatographic data from all three harvests and performing chromatographic peak detection, alignment and de-isotoping, the four LC–MS streams (RP +ve, RP –ve, HILIC +ve and HILIC –ve) contained 1179, 1012, 1016 and 768 mass/retention time features, respectively, which were reduced to 707, 607, 609 and 575 features, respectively after interquartile range filtering.

3.2. Statistical analysis of the pre and post de-greening data

To eliminate confounding effects from the different LC–MS streams, such as higher ‘analytical noise’ in one stream obscuring significant differences, the statistical analysis of the four LC–MS streams was performed separately. As samples were harvested during different seasons of the year (during the southern hemisphere late spring, midsummer and early autumn), it was likely that metabolite levels in planta would differ between the three harvests. Thus, the data for the three harvest dates were analysed separately, generating 12 datasets (4 streams × 3 harvests) for statistical analysis. Initial multivariate analysis by PCA, without a priori knowledge of the sample time point, consistently revealed two groupings within each of the 12 datasets, corresponding to the samples prior to (pre) and after (post) the de-greening stage of the manufacturing process. Fig. 1 demonstrates these groupings in PCA plots for the RP analysis streams (+ve and –ve ionisation) across the three harvests, with similar trends observed in both HILIC streams (Supplementary Fig. 1S). These observed differences within the sample groupings indicated that the 300 °C de-greening treatment had a major effect on the composition of the samples.

3.3. Identification of differentiating features between pre and post de-greening

Data from the two groupings (pre and post de-greening) were examined further using univariate tests such as t-test and fold change analysis, as well as the loadings from the PCA analysis. The t-test results, after Bonferroni correction (with adjusted P-value < 0.05/total number of features detected), revealed 69, 80, 84 and 30 feature ions from the RP +ve, RP –ve, HILIC +ve and HILIC –ve streams, respectively, that were significantly different between pre- and post-de-greening samples across all three harvest dates. The peak area data of these feature ions were assessed graphically and the ‘noisier’ or less significant features discarded to leave 18, 18, 33 and 12 features, respectively. A complete list of these remaining mass features detected in the four analytical streams and the magnitude of those changes are reported as heatmaps in the supplementary information (Supplementary Figs. 3S–6S). Many of these significant feature ions also contributed high loadings in the associated PCA score plots. The feature ion lists were compiled by chromatographic mode (i.e. RP or HILIC) and sorted by retention time to identify candidate source adducts or

---

Table 1
Manufacturing process and sampling time points for oolong tea.

<table>
<thead>
<tr>
<th>Sample #</th>
<th>Time from harvest (hr)</th>
<th>Manufacturing phase</th>
<th>Sample Grouping</th>
<th>DG status*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>Harvest</td>
<td>Harvest + Wilting</td>
<td>Pre</td>
</tr>
<tr>
<td>2</td>
<td>0.75</td>
<td>Wilting</td>
<td>Harvest + Wilting</td>
<td>Pre</td>
</tr>
<tr>
<td>3</td>
<td>1.63</td>
<td>Wilting</td>
<td>Harvest + Wilting</td>
<td>Pre</td>
</tr>
<tr>
<td>4</td>
<td>2.33</td>
<td>Fermentation</td>
<td>Early Fermentation</td>
<td>Pre</td>
</tr>
<tr>
<td>5</td>
<td>4.33</td>
<td>Fermentation</td>
<td>Early Fermentation</td>
<td>Pre</td>
</tr>
<tr>
<td>6</td>
<td>6.33</td>
<td>Fermentation</td>
<td>Mid Fermentation</td>
<td>Pre</td>
</tr>
<tr>
<td>7</td>
<td>8.33</td>
<td>Fermentation</td>
<td>Mid Fermentation</td>
<td>Pre</td>
</tr>
<tr>
<td>8</td>
<td>10.66</td>
<td>Fermentation</td>
<td>Mid Fermentation</td>
<td>Pre</td>
</tr>
<tr>
<td>9</td>
<td>11.5</td>
<td>Fermentation</td>
<td>Late Fermentation</td>
<td>Pre</td>
</tr>
<tr>
<td>10</td>
<td>13</td>
<td>Fermentation</td>
<td>Late Fermentation</td>
<td>Pre</td>
</tr>
<tr>
<td>11</td>
<td>14.16</td>
<td>Fermentation (end)</td>
<td>Fermentation</td>
<td>Pre</td>
</tr>
<tr>
<td>12</td>
<td>14.33</td>
<td>De-greening (after)</td>
<td>Early Drying</td>
<td>Post</td>
</tr>
<tr>
<td>13</td>
<td>14.58</td>
<td>Rolling</td>
<td>Early Drying</td>
<td>Post</td>
</tr>
<tr>
<td>14</td>
<td>15.03</td>
<td>Drying (first)</td>
<td>Early Drying</td>
<td>Post</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>Drying</td>
<td>Mid Drying</td>
<td>Post</td>
</tr>
<tr>
<td>16</td>
<td>28</td>
<td>Drying</td>
<td>Mid Drying</td>
<td>Post</td>
</tr>
<tr>
<td>17</td>
<td>32</td>
<td>Drying</td>
<td>Mid Drying</td>
<td>Post</td>
</tr>
<tr>
<td>18</td>
<td>36</td>
<td>Final product</td>
<td>Final Product</td>
<td>Post</td>
</tr>
</tbody>
</table>

\* De-greening status.
source induced dissociation (SID) fragment ions, which were highly correlated and co-eluted with the parent ion. Components were often detected in both ionisation modes and some were also detected in complementary chromatographic streams. The increasing number of online databases and their continued curation and expansion, are aiding the identification of unknown metabolites.

Fig. 1. Principal component analysis score plots (first two components) of normalised peak areas from (a) RP +ve spring harvest, (b) RP +ve summer harvest, (c) RP +ve autumn harvest, (d) RP −ve spring harvest, (e) RP −ve summer harvest, (f) RP −ve autumn harvest, showing differentiation between pre (●) and post (○) de-greening samples.
however it is important that searches be limited to the most relevant domain databases to avoid assigning incorrect identifications. Accurate mass information can provide likely molecular formulae to investigate further, once the molecular formulae generated from the accurate mass have been subjected to chemically intelligent scrutiny, such as the ‘Seven Golden Rules’ for heuristic filtering (Kind & Fiehn, 2007). The formulae may aid in at least speculating a likely chemical class of the unknown, or at least provide a placeholder of information pertaining to its identity. To identify these key differentiating components, the accurate mass m/z data was searched against an in-house library (which contains both accurate mass and retention time of standard compounds previously run on each analysis stream) and two LC–MS-based online databases (accurate mass only): METLIN (http://metlin.scripps.edu/) and HMDB (http://www.hmdb.ca/). Table 2 lists data for many of the most highly significant feature ions including confirmed identified metabolites (co-elution with a standard), tentatively identified metabolites (matches from online databases and a component which has either been observed in tea or a modification of components observed in tea), and ions from unidentified components. Suggested molecular formulae for these detected components are provided in the supplementary information (Supplementary information, Table 1S).

Guanosine, adenine and cyclic guanosine monophosphate (cGMP) were identified by a combination of co-elution with an authentic standard, accurate mass match and the observation of common source induced (SID) fragments. Several studies have shown that the common MS fragment observed for nucleotides/nucleosides is in each case the respective nucleic base (Kammerer et al., 2005; Liu, Ye, Qiang, Liao, & Zhao, 2008). Both the cGMP peak and the guanosine peak detected in HILIC +ve mode exhibited the expected 152.0566 m/z SID fragment characteristic of the nucleic base guanine, and these fragments are also reported in the METLIN database. The expected negative ion mode SID fragment of 150.0416 m/z was also observed in both the cGMP and guanosine HILIC –ve mode peaks. This consistent SID fragmentation pattern led to the tentative assignment of three other highly significant differentiating ions in the HILIC streams as nucleotides, based on accurate mass matches and appropriate nucleic base SID fragments (also reported in the METLIN database). These three further nucleotides were tentatively identified: adenosine monophosphate (AMP), based on a co-eluting weak 136.0611 m/z SID fragment matching the weak 136.0623 m/z adenine fragment; uridine 2',3'-cyclic phosphate or 2'-deoxyuridine 5-monophosphate, which have the same molecular formulae for the mass and contained the correct SID fragment of 113.0345 m/z for the uridine base (which both suggested nucleotides should yield); and cyclic adenosine monophosphate, which gave the 136.0613 m/z adenine SID fragment. The levels of these 4 nucleotides, guanosine (a nucleoside) and the nucleic base adenine, decreased 4-fold as a consequence of the de-greening step (Fig. 2). Nucleotides/nucleosides have been previously reported in fresh and manufactured (green and black tea) tea leaves (Koshiishi, Crozier, & Ashihara, 2001), with the highest levels observed in fresh leaves, followed by decreasing amounts in green and black tea samples, respectively. In the current study, the decline in nucleotide/nucleoside levels occurred largely in the de-greening step and can be attributed to the thermal instability of the N-glycosyl bond (White, 1984) rather than enzyme catalysed degradation as suggested by Koshiishi et al. (2001). Further investigation into the effects of different de-greening parameters on the flavour of the final product is warranted, as some of the nucleotides have been suggested to contribute to the desired ‘umami’ flavour of high quality teas (Koshiishi et al., 2001).

Flavonoids are commonly observed in tea (Wang & Ho, 2009) and in this experiment several flavonoids were observed to be affected by the de-greening stage. Myricetin and apigenin were positively identified, based on co-elution with authentic standards and accurate mass MS data. Two flavonoid-C-glycosides (Table 2) were tentatively classified based on accurate mass matches to flavonoid formulae in the databases, and on the observation of a number of highly correlated candidate SID fragment ions. In positive ionisation mode, two sequential water loss fragments were observed for both compounds ([M+H–H2O]+ and [M+H–2H2O]+) and in the negative ionisation mode, these two components exhibited co-eluting SID fragments of [M–H–C6H6O3]+, [M–H–C6H5O4]+, [M–H–C7H6O4]+ and [M–H–C8H7O4]+ consistent with fragmentations reported for flavonoid-C-glycosides in ESI-MS (Guo et al.,

### Table 2

<table>
<thead>
<tr>
<th>Parent</th>
<th>Mode</th>
<th>Rt (mins)</th>
<th>Identification</th>
<th>ppm</th>
<th>Error</th>
<th>Reference</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>594.1535</td>
<td>CN</td>
<td>8.51</td>
<td>Vitisin 2'-O-β-D-glucoside (C27H32O15)</td>
<td>2.5</td>
<td>HMDB, METLIN</td>
<td>Pre &gt;&gt; post</td>
<td></td>
</tr>
<tr>
<td>582.1602</td>
<td>CP, CN</td>
<td>3.82, 3.85</td>
<td>Flavonoid-C-glycoside (C26H30O14)</td>
<td>2.5–3.4</td>
<td>HMDB, METLIN</td>
<td>Pre &gt;&gt; post</td>
<td></td>
</tr>
<tr>
<td>552.1472</td>
<td>CP</td>
<td>4.08</td>
<td>Flavonoid-C-glycoside (C27H32O15)</td>
<td>0</td>
<td>HMDB, METLIN</td>
<td>Pre &gt; post</td>
<td></td>
</tr>
<tr>
<td>464.2284</td>
<td>CN</td>
<td>4.79</td>
<td>Linalool 3.6 (or 3.7)-oxide primaryerose</td>
<td>3</td>
<td>HMDB, METLIN</td>
<td>Post &gt;&gt; pre</td>
<td></td>
</tr>
<tr>
<td>448.2335</td>
<td>CN</td>
<td>5.97</td>
<td>Geranyl primaryrose</td>
<td>1</td>
<td>HMDB, METLIN</td>
<td>Post &gt;&gt; pre</td>
<td></td>
</tr>
<tr>
<td>446.0872</td>
<td>CN</td>
<td>5.23</td>
<td>Glucorhein (C21H18O11), Baicalin or Genistein glucuronide</td>
<td>3.5</td>
<td>HMDB</td>
<td>Pre &gt; post</td>
<td></td>
</tr>
<tr>
<td>448.0572</td>
<td>CN</td>
<td>5.78</td>
<td>Rothmiltin (C22H20O15)</td>
<td>0.4</td>
<td>HMDB</td>
<td>Pre &gt; post</td>
<td></td>
</tr>
<tr>
<td>416.1706</td>
<td>CN</td>
<td>4.39</td>
<td>Phenylethyl primaryrose</td>
<td>3</td>
<td>HMDB, METLIN</td>
<td>Post &gt; pre</td>
<td></td>
</tr>
<tr>
<td>416.1535</td>
<td>HP, CN</td>
<td>14.14, 14.04</td>
<td>Unknown*</td>
<td>1</td>
<td>HMDB, METLIN</td>
<td>Pre &gt;&gt; post</td>
<td></td>
</tr>
<tr>
<td>394.1856</td>
<td>CN</td>
<td>4.4</td>
<td>cis-3-Hexenyl b-primaryrose</td>
<td>2</td>
<td>HMDB, METLIN</td>
<td>Post &gt; pre</td>
<td></td>
</tr>
<tr>
<td>350.0616</td>
<td>HP</td>
<td>13.45</td>
<td>Unknown*</td>
<td>0.5</td>
<td>HMDB, METLIN</td>
<td>Pre &gt;&gt; post</td>
<td></td>
</tr>
<tr>
<td>347.0626</td>
<td>HP</td>
<td>12.06</td>
<td>Adenosine mono-phosphate</td>
<td>0.7</td>
<td>HMDB, METLIN</td>
<td>Pre &gt; post</td>
<td></td>
</tr>
<tr>
<td>345.0677</td>
<td>HP, CN</td>
<td>12.88, 12.72</td>
<td>Cyclic guanosine monophosphate</td>
<td>2.0–2.4</td>
<td>Standard, HMDB, METLIN</td>
<td>Pre &gt; post</td>
<td></td>
</tr>
<tr>
<td>334.0875</td>
<td>HP</td>
<td>13.47</td>
<td>Unknown*</td>
<td>2.7</td>
<td>Standard, HMDB, METLIN</td>
<td>Pre &gt; post</td>
<td></td>
</tr>
<tr>
<td>329.052</td>
<td>HP, CN</td>
<td>10.59, 10.52</td>
<td>Cyclic adenosine monophosphate</td>
<td>0</td>
<td>HMDB, METLIN</td>
<td>Pre &gt; post</td>
<td></td>
</tr>
<tr>
<td>318.0384</td>
<td>CN, CP</td>
<td>5.74, 5.78</td>
<td>Myricetin</td>
<td>2.1–3.2</td>
<td>Standard, HMDB, METLIN</td>
<td>Pre &gt; post</td>
<td></td>
</tr>
<tr>
<td>306.247</td>
<td>HP, CN</td>
<td>11.58, 11.46</td>
<td>Cyclic uridine monophosphate</td>
<td>0.6–3.0</td>
<td>HMDB, METLIN</td>
<td>Pre &gt; post</td>
<td></td>
</tr>
<tr>
<td>283.0911</td>
<td>HP, CN</td>
<td>10.60, 10.48</td>
<td>Guanosine</td>
<td>0.3–1.9</td>
<td>Standard, HMDB, METLIN</td>
<td>Pre &gt; post</td>
<td></td>
</tr>
<tr>
<td>270.0543</td>
<td>CN</td>
<td>7.18</td>
<td>Apigenin (C15H10O5)</td>
<td>1.4</td>
<td>Standard, HMDB, METLIN</td>
<td>Pre &gt; post</td>
<td></td>
</tr>
<tr>
<td>162.0323</td>
<td>CN</td>
<td>5.3</td>
<td>Umbelliferone (C15H10O5)</td>
<td>0.7</td>
<td>Standard, HMDB, METLIN</td>
<td>Post &gt; pre</td>
<td></td>
</tr>
<tr>
<td>135.0543</td>
<td>CN</td>
<td>8.2</td>
<td>Adenine</td>
<td>3.1</td>
<td>Standard, HMDB, METLIN</td>
<td>Pre &gt; post</td>
<td></td>
</tr>
</tbody>
</table>

Key: mode refers to both chromatographic and ionisation mode; CP = reversed phase +ve ionisation, CN = reversed phase –ve ionisation, HP = HILIC +ve ionisation, HN = HILIC –ve ionisation; component Rt (retention time) values are listed in the same order as the mode of detection; reference refers to the databases (HMDB = http://www.hmdb.ca/; METLIN = http://metlin.scripps.edu/) or if the compound co-eluted with an authentic standard; change refers to the trend of the differences, with the greater the magnitude of the change described by >>,* Further data provided in Supplementary information.
The remaining two flavonoids, glucorhein (or other likely database hits) and rothindin, were tentatively assigned based solely on accurate mass matches for the flavonoid formulae in the databases. All of the above flavonoids were observed to decrease after the de-greening stage. Further targeted investigation by extracting accurate mass peak areas for signals corresponding to other flavonoid glycosides commonly observed in tea extracts, such as kaempferol and quercetin glycosides (Dou, Lee, Tzen, & Lee, 2007; Lin, Chen, & Harnly, 2008; van der Hooft et al., 2012), revealed that many (but not all) other flavonoid glycosides were reduced in abundance as a consequence of de-greening, but all to a lesser degree than free myricetin (Supplementary Fig. 2S).

Although this process has not been previously reported during the tea manufacturing process, flavonoid glycosides are known to be thermally unstable, depending on the type and position of the glycoside (Rohn, Buchner, Driemel, Rauser, & Kroh, 2007). The flavonoids are valued in the product for their purported health benefits, so any decrease during the manufacturing stages may affect the perceived quality of the final product consumed.

The tentative assignments of four primeverosides (Table 2) are based on accurate mass matches to mono-terpene glycoside formulae in the databases and previous reports of the occurrence of primeverosides in tea (Wang, Yoshimura, Kubota, & Kobayashi, 2000). The four primeverosides were observed to increase in abundance significantly after the de-greening step. Higher concentrations of primeverosides in the final oolong tea product compared to fresh and withered leaves have previously been reported (Wang, Kubota, Kobayashi, & Juan, 2001). The hydrolysis of these primeverosides by β-primeverosidases to liberate various aroma compounds (e.g. mono-terpenes) which contribute to the overall flavour, is well known (Wang et al., 2001), particularly for black tea where substantial cell disruption and fermentation processes occur. It has previously been suggested that the primeverosides are stored in the vacuole (Mizutani et al., 2002), thus we suggest that since the increase occurs after the de-greening step (300 °C for 6 min) the intact primeverosides are likely thermally released from the vacuole and therefore more available for extraction after de-greening.

Fig. 2 illustrates the magnitude, direction and reproducibility of the trends observed for nine example metabolites from Table 2 across the three harvest periods, demonstrating the consistent and highly significant differences between pre- and post-de-greening. Some of the unknown components increasing in abundance after the de-greening treatment which we have not been able to identify from current natural products databases such as METLIN and HMDB, may be products of thermally catalysed reactions of plant components.

3.4. Statistical investigation of pre-de-greening fermentation data

To further investigate the wilting and fermentation stages of the manufacturing process, the pre- de-greening data were reprocessed separately. The eleven pre-de-greening sampling points were segmented into four groups, based on the samples being in a similar ‘fermentation phase’ (harvest + wilting, early fermentation, mid fermentation and late fermentation) to simplify data classification and interpretation. Initial investigations by PCA revealed some potential groupings based on fermentation phase, and the supervised PLS-DA statistical technique was applied to provide...
an indication of key components responsible for the separation. The PLS-DA method was generally able to group the samples within each of the four fermentation stages together over all four analytical streams (Supplementary Figs. 7S and 8S). Given that there were now the four assigned sample ‘groups’ an ANOVA rather than a t-test was required to generate a list of significant differentiating ions to compare with the key differentiating ions generated by PLS-DA.

3.5. Identification of key features varying with fermentation

Components consistently differing significantly by ANOVA across the four fermentation stages i.e. passing the significance test (adjusted \( P < 0.05 \) with Bonferroni correction) (as mentioned in Section 3.3) in each individual harvest and each analytical stream and detected in all three harvests were selected for further examination. Normalised peak area data for these components were then visually examined for consistent trends across the four fermentation stages. This univariate approach revealed a preliminary list of 46 mass features from the four streams, showing a consistent pattern of change for all three harvests which passed these criteria (16 from RP +ve; 10 from RP –ve; 17 from Hilic +ve; 3 from Hilic –ve). The fermentation group differentiation observed in the PLS-DA plots was mostly in the component 1 axis of the 2D score plots and, as such, the 60 highest ranking ‘variable in projection’ (VIP) scores (approximately 5% of the total mass features detected per analysis) for component 1 for each harvest and each stream were chosen for further scrutiny. These were further filtered for consistency and within each stream, VIP features that were highly ranked (within the top 60) at all three harvests were selected. This filtering of PLS-DA VIP scores for component 1 yielded a total of 45 mass features (4 from RP +ve; 12 from RP –ve; 19 from Hilic +ve; 10 from Hilic –ve). Finally, to aid de-replication, the significant features detected by either or both ANOVA and PLS-DA from the four analytical streams were merged into a single table (Table 3).

Investigation of the key components changing during fermentation revealed several amino acids which increased during the process. Tyrosine, tryptophan, phenylalanine, asparagine and proline (and their many respective SID fragment ions) all appeared as key ions from the statistical analysis above and their identifications were confirmed with co-elution and accurate mass matches to authentic standards (Table 3). We have previously shown that amino acids are readily detected with this analytical approach, and that their levels can vary significantly in teas of differing tea type and origin (Fraser et al., 2012). Fig. 3 shows example data for two free amino acids which accumulated during the fermentation phase, with relative levels increasing 7-fold for tryptophan and 10-fold for tyrosine. Further targeted investigation into other amino acids revealed increasing levels of leucine, isoleucine, serine, histidine, threonine and lysine, while minor decreases in glutamic acid, glutamine and threonine levels were also observed (Supplementary Table 2S). Amino acids that did not change significantly in abundance were valine, alanine, arginine and aspartic acid. An increase of most free amino acid levels during withering for black tea manufacturing for up to 20 h (equivalent to the total ‘gentle’ fermentation process used for the oolong manufacturing process monitored here) have been previously observed, with the most likely source of these amino acids being release from protein breakdown occurring during the fermentation (Dev Choudhury & Bajaj, 1980; Roberts & Sanderson, 1966; Sanderson & Graham, 1973; Tomlins & Mashingaidze, 1997). This withering process clearly influences the abundance and composition of free amino acids released and is affected by duration, temperature and moisture content (Roberts & Sanderson, 1966).

The highly significant ion of neutral mass 226.1204 observed in 3 analytical streams matched that for hydroxyjasmonic acid (HO-JA) in the METLIN database (0.8 ppm error). Two highly correlated (and significant) co-eluting SID fragment ions attributable to consecutive water losses in positive mode provided supporting evidence for this identification, as these are likely neutral losses from a HO-JA structure with two free hydroxyl groups. HO-JA, together with other jasmonate metabolites, have previously been reported to accumulate in both tomato and arabidopsis leaf tissue as part of a response to wounding (Glauser, Boccard, Rudaz, & Wolffender, 2010; Miersch, Neumerkel, Dippe, Stenzel, & Wasternack, 2008). A search of our data revealed metabolite peaks exhibiting similar trends to HO-JA which could be assigned to jasmonic acid (JA) and hydroxyjasmonoyl isoleucine (HO-JA-Ile), respectively, on the basis of accurate mass matches and relative retention order (Glauser et al., 2010), with JA levels peaking slightly before both HO-JA and HO-JA-Ile. In the case of JA, identification was subsequently confirmed with an authentic standard. Given the tea leaves are hand-picked and then hand-tossed several times during the wilting phase, the detection and observed trend of compounds likely to have been induced as part of a wound response mechanism is not unexpected, although this is the first time this has been reported in tea fermentation studies. It is likely there are other jasmonate metabolites formed at lower levels by this mechanism, as observed in other plants, but monitoring these might require more targeted methodology.

Two candidate metabolites related to lipid degradation were also observed to significantly change during the fermentation process. Phosphocholine was tentatively identified by accurate mass matches to METLIN and HMDB and the observation of highly correlated co-eluting significant SID fragment ions matching the MS² spectrum in the METLIN database. It was observed to increase in abundance during fermentation and this could be accounted for by the release of the phosphocholine head group during lipid degradation. The other metabolite was assigned as a putative lipid (m/z 314,2466) as it was an excellent match for dixo-C18 lipids in the databases, eluted later off the RP column (as can be expected for lipids) and followed the trend of declining levels due to potential lipid degradation (Fig. 3).

As a final note, the levels of phenolic monomers such as epigallocatechin and epigallocatechin gallate extracted from the RP –ve stream, did not change significantly during the fermentation process in this study. This suggests that the fermentation process used in this study was very mild, as in other studies the levels of these phenolic monomers have been found to decrease markedly during tea fermentation (Harbowy et al., 1997; Wang & Ho, 2009). A recent report (Kim et al., 2011) showed a constant decline in phenolic monomer concentrations across the 0–80% fermentation range monitored, however only minor changes occurred in phenolic levels between the 0% and 20% fermented leaves. Oolong tea from this estate has been monitored in our earlier studies (Fraser et al., 2012, 2013) using non-targeted analysis and multivariate analysis, and has been observed by PCA to cluster with both non-fermented green teas and lightly fermented oolong teas, consistent with the relatively ‘mild’ fermentation of the manufacturing process at the Zealong Estate.

4. Conclusions

We have presented here the first comprehensive non-targeted analysis of the oolong tea fermentation/manufacturing process. This study has revealed several novel metabolite changes occurring
during the tea manufacturing process and has identified particular stages of tea manufacturing where these occur. We have shown that the de-greening step is responsible for the most significant biochemical changes during the manufacturing process, causing a significant decrease in the levels of nucleotides/nucleosides which had previously been assumed to be due to enzymatic rather than...
thermal processes. This de-greening step also caused a decrease in the levels of the health beneficial flavonoids, and increased levels of primeverosides, all of which could have implications for the flavour/taste of the final product. Investigations into a sub-set of the data revealed the more subtle changes that occur during the fermentation process, such as an increase in jasmonic acid metabolites not previously reported during tea manufacture. We have also confirmed previous reported biochemical changes, such as the increase of free amino acids during fermentation, without the need to perform separate targeted analyses.

The methodology presented demonstrated that a non-targeted LC–MS approach to collecting as much non-biased information on as many metabolites as possible in a single sample, combined with modern statistical techniques, can rapidly and comprehensively monitor compositional changes during a fermentation and manufacturing process. We have also shown that monitoring a range of metabolites covering both a wide polarity range and abundance can contribute to a broader understanding of biochemical changes. The approach is useful for generating a list of candidate markers to be further investigated and identified and may reveal processes occurring that have not been previously observed, such as the release of jasmonic acid metabolites during fermentation, and/or provide validation and monitoring of already known processes. The ‘conventional’ analytical approach is to start with a list of known target metabolites (and standards) to measure, rather than detection of mass features followed by identification. The challenge of this non-targeted approach is then to identify the significantly differing unknown components (Dunn et al., 2013), which can be more difficult where the components differing are products formed from complex thermal reactions and are unlikely to be recorded in databases of naturally-occurring compounds.

Finally, a cautious approach to the use of appropriate statistical tools for the large non-targeted data matrices should be adhered to, to limit false positives in the first instance. For example, while the Bonferroni correction is a conservative measure of significance driven by the large number of analytes measured, it is generally recognised as a starting point (Broadhurst & Kell, 2006). By utilising more than one statistical tool to generate leads, and performing visualisation of the significant raw data, the analyst can safeguard the quality of the data. If the data analysis fails to reveal significant features, then re-analysis with less stringent statistical parameters may take longer to examine but could still yield genuine results.

The results of this study suggest there are two directions in which this research could be developed to provide useful tools for the tea manufacturing industry. Monitoring the processes of tea fermentation/manufacturing by rapid targeted LC–MS methodology could ensure feature molecules that provide important quality attributes are maintained at optimum levels. Non-targeted LC–MS methodology may be useful for the detection and identification of other/improved indicators for assisting tea manufacturers to produce a more consistent/reproducible final product, and to provide indications of sample contamination or manufacturing faults, minimising valuable brand damage by detecting problems before the product reaches the consumer.

Acknowledgements

We wish to thank Jeff Howell, Nicole Wang and Vincent Chen, from Zealong, for providing access to the Zealong tea estate to collect the samples throughout the manufacturing process and for many helpful discussions on tea manufacturing processes. We wish to thank Leo Liu of AgResearch for his assistance in the laboratory, and Dr Mingshu Cao of AgResearch for his assistance with de-isotyping the raw data matrix and many informative discussions and advice on statistical analyses.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.foodchem.2013.11.054.

References


