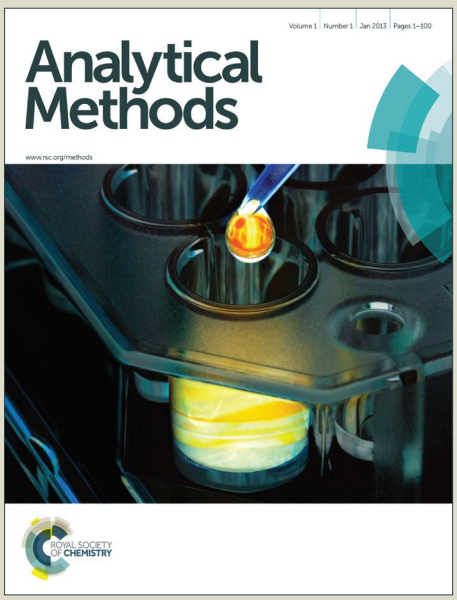
Analytical Methods

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

More Than Just Heat: Ambient Ionization Mass Spectrometry for Determination of the Species of Origin of Processed Commercial Products—Application to Psychoactive Pepper Supplements

Ashton D. Lesiak[a] and Rabi A. Musah[a*]

[a]Department of Chemistry, University at Albany-SUNY, 1400 Washington Ave, Albany, NY 12222

*Corresponding Author

ABSTRACT

The application of direct analysis in real time high resolution mass spectrometry (DART-HRMS) to the determination of the originating species of the plants from which processed commercially available plant-based products are derived is described. As a proof of principle, the method was employed for determination of the provenance of psychoactive pepper species, namely *Piper methysticum* (aka kava) and *P. betle* (aka betel). In addition to being of agricultural importance, these species are also of relevance in a forensics context. DART-HRMS spectra showed that extractions, heat treatment and other steps associated with the manufacture of these products result in significant differences in the mass spectral fingerprints observed. Nevertheless, the presence of key species-specific biomarkers such as kavalactones and chalcones in *P. methysticum*, and a variety of terpenes in *P. betle*, were retained. Chemometric processing by principal component analysis using a selection of feature masses that represented both compounds common to each of the species, and others that distinguished them, showed that the two *Piper* spp. could be readily identified, regardless of the manufacturing process used to create the product, with a leave-one-out cross validation result of 100%. Furthermore, unsupervised statistical analysis processing by hierarchical clustering not only enabled *P. methysticum* and *P. betle* products to be distinguished from one another, but also permitted further discrimination that was based on the processing method used to produce them. Advantages of the method over commonly used conventional protocols include minimal methods development; the capability of analyzing material in its native form without resorting to solvent extraction, derivatization or other sample preparation steps; speed; and the ability to detect and definitively identify biomarkers characteristic of a species.

The method has wide applicability and is particularly useful for analysis of products from plants whose genes have not been mapped, and which, as a consequence, cannot be subjected to DNA analysis to determine the originating plant species.

1. Introduction

Determination of the provenance of processed plant-derived commercially available products is an area of analytical chemistry that continues to present significant challenges. While plant species determination based on the anatomical and morphological features of plant parts can be accomplished when raw or dried plant segments of significant size are available, most of the processing steps used to convert the material to forms that are suitable for consumption as food, ingestion as dietary supplements, or use as additives in cosmetic products, result in loss of visual plant species identifiers. These treatments, which can include pulverization, heating, or blending with other substances, often make assessment of the authenticity of the ingredient profiles listed on product packaging very difficult, particularly when the final matrix is complex. The need and urgency to address these challenges have increased due to the exponential rise in the sale and distribution of adulterated products, and the marketing of unsafe herbal or dietary supplements that contain mind-altering plant ingredients or other toxic natural products. The crafting of legislation that could curtail the distribution, sale and abuse of such substances is severely hampered by the paucity of rapid and efficient analytical methods that can be applied to definitive determination of the species of plants of which a commercially available product is comprised.

Several conventional approaches including DNA analysis and methods involving chromatography, spectroscopy, spectrometry, or various combinations of these, have been successfully applied to the identification of plant species in matrices of differing levels of complexity. However, protocol development is usually time and resource intensive.[1] Furthermore, the developed methods are often specific for analysis of the product being

investigated, and they are therefore not easily applied to a broader range of substances. DNA has been used to differentiate between many species of plants including caraway and cumin in spices[2, 3] and for determination of the botanical origin of food products including nuts, fruit juice and coffee.[4] Twenty three genera of medicinal plants in the Araliaceae family[5] as well as common cash crops such as maize, wheat and legumes[6-9] have also been characterized by DNA analysis. However, routine screening for the purpose of plant species identification is impractical, costly, and limited to plant species whose genetic information is known. Spectroscopic techniques such as infrared (IR) spectroscopy and near-IR have been applied to discriminating between species of herbal medicines and flowering plants from the *Epilobium* and *Hypericum* genera[1, 10, 11] without complex sample preparation steps. However, in most reported cases, the output of these methods provides little information on the chemical constituents that are the basis of discrimination between the individual species. Hyphenated chromatographic and spectrometric techniques have also been used, for example, in *Actaea* (aka black cohosh)[12] and *Triticum* spp (aka wheat) differentiation.[13] The routine use of these methods has mainly been confined to cash crops, as the sample preparation steps are time consuming, the chromatographic run times are long, and the many replicates needed for the application of statistical analysis processing of the data require too much time and material resources (in the form of solvents, chromatographic columns etc.). Moreover, the subjection of the sample to solvent extraction and/or derivatization steps can adversely affect the ability to distinguish between plant species because of the bias exhibited by the solvent in selecting some biomarkers over others.

Direct analysis in real time high resolution mass spectrometry (DART-HRMS) is an ambient ionization mass spectrometry technique that can circumvent a number of the drawbacks often encountered when using some of the aforementioned methods. It can be used to analyze materials in their native form (solids, liquids or gases) and as such, sample preparation steps such as solvent extraction are often not required. Moreover, a wide variety of polar and non-polar compounds can be detected simultaneously. An added advantage is that a broader range of compounds can often be detected when no solvent is used, since differential solubilities of compounds by the solvent, which influences the profile of molecules observed, can be avoided. DART-HRMS derived data have previously been used to distinguish between various species of *Datura* seeds,[14] agarwood samples regulated under international trade agreements,[15] insect puparial casings,[16] and plant material of forensic relevance.[17] In these studies, the samples analyzed were in their natural form in that they had not been subjected to processing steps such as heating or pulverization. In this work, we investigated the hypothesis that DART-HRMS could be used to identify the species of plant from which material that had undergone various levels of processing had been derived. To demonstrate proof of principle, the technique was applied to psychoactive products from the *Piper* genus using chemical fingerprint signatures and statistical analysis processing.

P. methysticum and *P. betle*, commonly known and "kava" and "betel" respectively, are two economically important *Piper* species that have enjoyed centuries of use for their psychoactive and medicinal effects. The former is the most important psychoactive agent used in the islands of the South Pacific.[18] Its pharmacological effects are attributed to kavalactones found in the roots of the plant.[19, 20] While 18 kavalactones have been isolated and identified, six

(Fig. 1) are considered to be major constituents: yangonin, methysticin, dihydromethysticin, dihydrokavain, kavain, and desmethoxyyangonin (dehydrokavain).[19, 21-23] Other minor kavalactones as well as some chalcones have also been identified in *P. methysticum* plant material and products. These include tetrahydroyangonin and the chalcones flavokavain A and flavokavain B (Fig. 1). Kava ingestion has been implicated in poisonings and deaths, especially when used in conjunction with alcohol or other drugs, making its identification important in forensic and toxicological contexts.[24-27] *P. betle* is indigenous to Southeast Asia and is the third most widely used stimulant worldwide, with over 400 million users throughout Asia.[28] It contains a wide variety of constituents including hydroxychavicol, hydroxychavicol acetate, allypyrocatechol, chavibetol, eugenol, chavicol, methyl eugenol, allypyrocatechol acetate, allypyrocatechol diacetate, α-terpinene, and β-caryophyllene, some of which are shown in Fig. 2.[28-31] While the purported benefits of *P. betle* include anti-inflammatory, gastrointestinal, anti-fungal and anti-larval effects among others,[29, 30, 32] excessive use of betel and betel quid (areca nut wrapped in betel leaves) is associated with gum disease, mouth ulcers, oral cancers, and increased risk of cardiovascular disease.[28, 33]

Characterization of *P. methysticum* plant material and products has been limited to studies designed to determine the identities of their psychoactive constituents using conventional analytical methods, including spectroscopic, chromatographic, spectrometric and hyphenated techniques including HPTLC,[34] near-infrared reflectance spectroscopy,[35] NMR,[36, 37] LC-UV,[38] LC-MS,[38] and HPLC-MS[39] approaches, in addition to DNA analysis.[40] Similarly *P. betel* extracts, leaves and other goods have been analyzed by phytochemical color tests,[41] liquid-liquid and supercritical fluid extraction,[42] HPLC,[43] GC-MS,[44] and UPLC-MS/MS.[45] The above

mentioned methods require complex sample preparation steps including solubilization, extraction, filtration, concentration and derivatization, which greatly increase analysis time. Moreover, the wide variety of forms in which *P. methysticum* and *P. betle* are found, including extracts, powders, tinctures, fresh and dried leaf material, root bark and oils (Fig. 3) make streamlining the analysis difficult, since different sample preparation steps are needed for each physical form of the plant material.

While the application of multivariate statistical analysis to DART-HRMS derived data has been successfully applied to identify species,[14, 16, 17] the materials analyzed in earlier studies were in their natural form (e.g. seeds, leaves or bark) and relied on the presence of biomarkers that enabled differentiation of one species from the other. However, processing steps such as extraction and brewing can change the chemical composition of the sample by eliminating or modifying important biomarkers. This would in turn adversely affect the statistical analysis processing and prevent definitive identification of the parent plant.

Here, we show that a wide variety of processed kava and betel products including leaves, roots, powders, essential oils and a tincture can be rapidly identified and distinguished, even though the two species are from the same genus, by chemometric processing of DART-HRMS derived chemical fingerprint data. The results showed that this approach is amenable to being broadly applied to the determination of the species of origin of processed plant-derived material.

2. Experimental

2.1 Plant Materials

Kava 40% kavalactones powder was purchased from eBay (ebay.com, USA). Kava root powder, dried kava root and kava 70% kavalactones powder was purchased from Bouncing Bear Botanicals (Lawrence, KS, USA). Kava extract tincture was purchased from Herbal Island (herbal-island.com, USA). Betel leaf essential oil was purchased from Healing Solutions (Scottsdale, AZ, USA) and *Piper betle* plants were purchased from Logee's Plants (Danielson, CT, USA).

2.2 Standards

Yangonin was purchased from Cayman Chemical (Ann Arbor, MI, USA). Flavokavain A and flavokavain B were purchased from Abcam Biochemicals (Cambridge, MA, USA). Desmethoxyyangonin, methysticin, D,L-kavain, dihydromethysticin, and dihydrokavain standards were purchased from Cerilliant, Inc. (Round Rock, TX, USA). Isoeugenol, α-terpinene, eugenyl acetate and β-caryophyllene were purchased from Sigma Aldrich (St. Louis, MO, USA).

2.3 Acquisition of Mass Spectral Data

Soft ionization mass spectra of plant materials and standards were acquired using a DART-SVP ion source (IonSense, Saugus, MA, USA) coupled to a JEOL AccuTOF high resolution time-of-flight mass spectrometer (JEOL USA, Peabody, MA, USA) in positive-ion mode. The DART ion source parameters were: grid voltage, 250 V; and gas heater temperature, 350 °C. The mass spectrometer settings were: ring lens voltage, 5 V; orifice 1 voltage, 20 V; orifice 2 voltage, 5 V; and peak voltage, 600 V. Mass spectra were acquired over the *m/z* range 60-800 at a

spectral acquisition rate of 1 spectrum per sec. The helium flow rate for the DART ion source was 2.0 L s^{-1}. The resolving power of the mass spectrometer was 6000 FWHM.

For structural confirmation of compounds, in-source collision-induced dissociation (CID) was performed on plant material and standards through the use of "function switching" feature in which the orifice 1 voltage is varied from 20 V to 30, 60, and 90 V to induce increasing levels of compound fragmentation within a single analysis. The RF ion guide voltage ("Peaks voltage" in the Mass Center software) for CID analysis was set to 400 V, and the mass range was set to *m/z* 40-800. All other ion source and mass spectrometer parameters were as described above.

Kava powders were tested directly by dipping the closed end of a melting point capillary tube into the powder and presenting the coated surface of the tube to the 4 cm space between the DART ion source and the mass spectrometer inlet. The sample was held approximately 2 cm from the mass spectrometer inlet for 5 seconds for each analysis. The betel essential oil, kava extract tincture, and standards were analyzed in the same manner. The kava dried root was sampled by grasping the material with tweezers and suspending it between the ion source and the mass spectrometer inlet.

2.4 Data Processing

Mass spectral calibration, averaging, background subtraction, and peak centroiding were achieved using TSSPro3 software (Shrader Software Solutions, Detroit, MI). Polyethylene glycol (PEG 600) was used as the mass calibration standard for all samples. Mass Mountaineer (RBC Software, Portsmouth, NH, available from mass-spec-software.com) was used for mass spectral analysis, spectral elemental composition determination, isotope analysis and statistical

analysis, including principal component analysis, and heat map generation. Heat maps exported from Mass Mountaineer were imported into Cluster 3.0 and Java Treeview (Stanford University) for hierarchical clustering analysis. Nominal masses are listed throughout the text for clarity, but accurate masses are found on the figures and in the mass measurement tables.

3. Results and Discussion

3.1 High resolution masses consistent with those of compounds previously isolated from Piper spp. were detected by DART-HRMS

P. methysticum and the kava products derived from it are known to contain characteristic chemotaxonomic markers including kavalactones and chalcones that can be used for identification of plant material. To investigate the extent to which processing steps influenced the profile of compounds observed, five kava products that had been processed to varying extents were purchased from internet vendors. The kava products were analyzed by DART-HRMS and the results of the soft ionization analysis are shown in Fig. 4, with the corresponding mass measurement data presented in Table 1. Five averaged spectra are shown in each case, and each individual peak in the mass spectrum represents a unique protonated compound. The number of peaks in the mass spectra above a 1% abundance threshold varied from 40 in the kava tincture (Panel e) to 88 peaks in the kava 40% kavalactones powder (Panel a).

High resolution masses consistent with those of seven kavalactones that were previously isolated from *P. methysticum*[19, 22, 23, 37, 46] were present in all five of the kava samples, although the relative abundances varied between products (Table 1). A mass indicative of the

formula of protonated dehydrokavain ($[C_{14}H_{12}O_3]+H^+$ corresponding to *m/z* 229.0865) was detected in all of the samples. Masses corresponding to protonated kavain ($[C_{14}H_{14}O_3]+H^+$ at *m/z* 231.1021) and protonated dihydrokavain ($[C_{14}H_{16}O_3]+H^+$ at *m/z* 233.1178) were detected in varying relative amounts (23.7%-78.2% and 54.2%-88.7% respectively) in the five kava samples as well. Furthermore, a mass corresponding to protonated yangonin ($[C_{15}H_{14}O_4]+H^+$ at *m/z* 259.0970) was the base peak in three of the five kava samples [kava 70% powder, kava root powder and kava dried root] and it was 50.9% and 71.0% in the kava tincture and kava 40% powder respectively. Other tentatively assigned kavalactones detected included tetrahydroyangonin ($[C_{15}H_{18}O_4]+H^+$ corresponding to *m/z* 263.1283), methysticin ($[C_{15}H_{14}O_5]+H^+$ corresponding to *m/z* 275.0920), and dihydromethysticin ($[C_{15}H_{16}O_5]+H^+$ at *m/z* 277.1076). Masses consistent with two chalcones previously identified in kava were also detected. Tentative assignments for flavokavain A ($[C_{18}H_{18}O_5]+H^+$ corresponding to *m/z* 315.1233) and flavokavain B ($[C_{17}H_{16}O_4]+H^+$ corresponding to *m/z* 285.1127) were made for all five products. The peaks exhibited relative abundances in the range of 19.7% to 100.0%.

P. betle leaves and roots, and betel essential oil were also subjected to analysis by DART-HRMS, with the results presented in Fig. 5. Associated mass measurement data are listed in Table 2. Each panel shows an average of five spectra. The number of peaks corresponding to individual protonated compounds ranged from 72-75 using a 1% abundance threshold. Masses consistent with various natural products that have previously been isolated from *P. betle* were detected, including α-terpinene ($[C_{10}H_{16}]+H^+$ at *m/z* 137.1330) and isoeugenol ($[C_{10}H_{12}O_2]+H^+$ at *m/z* 165.0916). Methyl eugenol ($[C_{11}H_{14}O_2]+H^+$) was detected at *m/z* 179.1072 in the essential oil, as were chavicol ($[C_9H_{10}O]+H^+$ at *m/z* 135.0810) and eucalyptol ($[C_{10}H_{18}O]+H^+$ at *m/z*

155.1436). Masses corresponding to allyl pyrocatechol acetate ($[C_{11}H_{12}O_3]+H^+$ at *m/z* 193.0865), eugenyl acetate ($[C_{12}H_{15}O_3]+H^+$ at *m/z* 207.1021) allyl pyrocatechol diacetate ($[C_{13}H_{18}NO_4]+H^+$ at *m/z* 235.0940) and an unknown compound ($[C_{13}H_{14}O_4]+H^+$ at *m/z* 252.1236) were detected in all three betel samples in varying abundances (from 2.8%-100.0%).

3.2 In-source CID analysis confirmed the presence of kavalactones, chalcones, terpenes and other compounds in the kava and betel products

The soft-ionization spectra produced by DART-HRMS analysis allowed tentative assignments of compound formulas to be made based on the accurate masses. However, it has been shown previously that compound identities can be confirmed through comparison of in-source collision-induced dissociation (CID) spectra of plant parts with those of authentic standards obtained under identical conditions.[14, 17] The "function switching" feature of the AccuTOF mass spectrometer enables simultaneous acquisition of spectra at various orifice 1 voltages, with increasing voltages resulting in greater fragmentation. We applied the same approach to analysis of *P. methysticum* and *P. betle* products. The 90 V orifice 1 voltage consistently yielded the best results, in that both the protonated parent and fragment ions were detected. Representative spectra are illustrated in Fig. 6 as head to tail plots. In each of the panels a through g, the in-source CID spectrum of the kava product is on top, and that of the indicated standard appears on the bottom. The mass measurement data associated with the spectra are presented in Tables S1-S7.

The in-source CID spectrum of the kavain standard shown in Fig. 6a exhibited a protonated parent peak at nominal mass *m/z* 231, with fragments at *m/z* 155, 153, 115, and

129. These peaks were also present in the in-source CID spectrum of the kava product (in this case, the 40% powder), confirming the presence of kavain in the sample. In Panel b, the parent peak at *m/z* 233 and fragments at *m/z* 187, 155, 117 and 91 were present in both the dihydrokavain standard as well as the kava product. The presence of yangonin was confirmed as well (Panel c), as the peaks at *m/z* 259, 231, 171 and 161 that were observed in the standard were also detected in the kava powder. The bottom panel in Fig. 6d shows the in-source CID spectrum of methysticin with the protonated parent peak at *m/z* 275, and fragment peaks at *m/z* 159, 141 and 103. The top spectrum also had matching parent and fragment peaks at *m/z* 275, 159, 141 and 103, indicating that methysticin was present in the kava powder. The presence in the kava product of dihydromethysticin, flavokavain B and flavokavain A was similarly confirmed (Fig. 6e-g) through observation in its spectrum of the protonated parents (*m/z* 277, 285 and 315, respectively) and their corresponding fragments for each of the standards.

In-source CID experiments were also performed on *P. betle* products to confirm the identities of tentatively assigned peaks. The results are shown in Fig. 7 and the corresponding mass measurement data are presented in Tables S8-S11. In the bottom spectrum of Fig. 7a, the in-source CID spectrum of isoeugenol showed a protonated parent peak at nominal *m/z* 165 and fragments at *m/z* 165, 137, 105, 91 and 77. These were also present in the 90 V in-source CID spectrum of the betel leaf essential oil, thus confirming the presence of isoeugenol in the oil. Fig. 7b shows the head-to-tail plot of the betel essential oil and β-caryophyllene. The betel spectrum exhibited a parent peak at *m/z* 205 and matching fragments at 149, 119, 95 and 81, confirming the presence of β-caryophyllene in the oil. α-Terpinene (Panel c) and eugenyl

acetate (Panel d) were also both confirmed to be present in the essential oil, with parent peaks at *m/z* 137 and 207 respectively, along with their corresponding fragment peaks.

3.3 Statistical analysis of DART-HRMS data obtained from P. methysticum and P. betle products enabled species-level identification

In previous studies, DART-HRMS data have been used in conjunction with multivariate statistical analysis methods[14, 16, 17] to differentiate species and cultivars of botanical samples. Here we sought to ascertain whether *P. betle* and *P. methysticum processed* products could be identified based on chemometric processing of their MS-derived chemical fingerprints. To investigate this, a dataset of 70 averaged mass spectra (350 individual analyses) and 20 feature masses (Table 3) were used for principal component analysis (PCA) processing. The results are shown in Fig. 8. A total of 10 averaged spectra each from the kava tincture, kava 40% powder, kava 70% powder, kava dried root, kava root powder, betel essential oil and *P. betle* leaves were used in the classification training set, as these products have documented use as herbal products or in supplements. *P. betle* root was excluded from the classification set as there is no documented use of the root in a medicinal or illicit drug context. The two classes of data, *P. betle* (blue) and *P. methysticum* (red) were well resolved from one another, with three principal components (PCs) accounting for 70.92 % of the variance, and a leave-one-out cross validation (LOOCV) result of 100%. The tight clustering within each class and the resolution between the classes indicated that while the interspecies chemical profiles were unique, the intraspecies compositions were reproducible for different products that were derived from the same plant species.

Using the same dataset and feature masses, each of the total of eight kava and betel products was assigned to its own class in order to assess whether statistical analysis processing of the DART-HRMS data could be used to identify the specific type of psychoactive pepper product. The kava spectra were designated as kava 40% powder, kava 70% powder, kava dried root, kava root powder or kava tincture. The betel spectra were assigned to either the betel essential oil class or the *P. betle* leaf class. PCA was again applied to the DART-HRMS data of the five kava and two betel products. The results are illustrated in Fig. 9. Panel a shows the individual classes of products separated across the x-axis by species (*P. methysticum* or *P. betle*). Clustering was observed for the samples denoted in grey and yellow (both of which were *P. betle* products), while those represented by blue, green, turquoise, red and pink symbols (*P. methysticum* products) clustered together. The top down view of the same PCA plot (Panel b) showed that while the *P. betle* and *P. methysticum* products were resolved from one another, the individual products within each species were also resolved. For example, the yellow circles (betel leaf essential oil) and grey triangles (*P. betel* leaf) appeared in the same plane along the x-axis, showing that both were *P. betle* products. However, they were also resolved along the z-axis, indicating that the leaf and oil products could be distinguished based on differences in sample processing. Similarly, the *P. methysticum* products all appeared at > -2.40 along the x-axis, confirming their classification as *P. methysticum* products. However, the tincture (pink squares), kava root powder (red squares), kava 40% powder (turquoise circles), kava 70% powder (blue circles), and kava root (green triangles) were resolved from one another both along the x- and z-axes. Three principal components covered 70.92% of the variation, and the LOOCV was 95.72% with only three misclassifications.

Heat map renderings of the *P. betle* and *P. methysticum* mass spectra that visually illustrate the interspecies similarities and intraspecies differences are shown in Fig. 10. High intensity peaks are dark red in color and lower intensity peaks appear in lighter shades. The kava products all showed high intensity peaks corresponding to the various lactones identified in the soft ionization spectra (data presented in Fig. 4 and Table 1), whereas the *P. betle* samples showed higher intensity peaks corresponding to isoeugenol and eugenyl acetate, among other compounds (Fig. 5, Table 2). The heat maps were imported into open source hierarchical clustering software (Cluster 3.0) to assess whether product relatedness based on chemical fingerprint data could be determined. The resulting dendrogram is presented in Fig. 11. It has two main branches (enclosed with dashed lines) corresponding to the two species of psychoactive peppers analyzed in this study. The clear separation between *P. betle* and *P. methysticum* indicated that the two species could be easily distinguished based on their chemical fingerprints. The *P. betle* products were further separated by branching into distinct clades, indicating that the essential oil and the *P. betle* samples were distinct, despite being derived from the same species. Thus, the separation reflected differences in sample processing. The results also showed that although the kava products clustered together and were clearly distinguished from those of *P. betle*, there was overlap between some of the kava products processed by different methods. For example, although most of the kava tincture data clustered together, some of it clustered with the kava 40% powder data. Similarly, most of the kava 70% powder data appeared within the same clade (blue box), but a couple of the data points appeared to be more closely aligned with those of the dried whole root samples.

3.4 Implications

In this study, we investigated whether the originating species of plants that had been subjected to varying levels of processing could be determined, using the output of DART-HRMS experiments together with statistical analysis processing of the data. Commercially available *P. methysticum* and *P. betel* samples were used, and both supervised (PCA) and unsupervised (hierarchical clustering analysis) statistical tools were employed.

The results of the DART-MS experiments illustrated a number of trends. Not unexpectedly, it was observed that processing steps can dramatically change the chemical fingerprint of the sample relative to that of the starting material. For example, in the kava tincture (Fig. 4), there were 40 peaks detected above a 1% abundance threshold, compared to 57 peaks in the unprocessed whole root. This reduction in constituents was not surprising and was likely due to the differential solubilities of compounds in the alcohol extraction step performed by the manufacturer.[47] A decrease in the number of peaks in the kava 70% powder (88 peaks) relative to the kava 40% powder (59 peaks) was also noted, and was due to a reduction in the range of kavalactones and other compounds in the former vs. the latter.[48] Not only did the number of peaks change between products of the same species, but the relative intensities of the *m/z* values that were present in all of the kava samples also varied. For example, yangonin and methysticin were detected in each of the kava products (Fig. 4, Table 1), but the relative intensities varied from 50.9% to 100.0% in the case of yangonin, and 5.8%-39.6% in the case of methysticin. The analysis of the betel products provided similar results, where allyl pyrocatechol acetate, β-caryophyllene, and allyl pyrocatechol diacetate were detected in each of the betel products, but the intensities of each compound varied between the leaf samples and the oil (Fig. 5).

Due to the reduction of the number of diagnostic compounds and the variation in the relative intensities of the remaining peaks in the spectra of *P. methysticum* and *P. betle* products due to processing, we investigated whether the application of multivariate statistical analysis of the MS-derived data could be used for plant species identification of the products. For the creation of the training set, twenty feature masses which included biomarkers of *P. methysticum* and *P. betle* such as allyl pyrocatechol, eugenyl acetate, kavain, yangonin and flavokavains A and B, among others (Table 3) were chosen. The a priori selection of feature masses was undertaken with great care, as the peaks selected needed to correspond to diagnostic markers unaffected by processing techniques. The results of the PCA analyses (Figs. 8 and 9) demonstrated that despite the difference in the range of peaks from product to product, it could easily be confirmed that the samples were derived from either *P. betel* or *P. methysticum,* and the LOOCV was greater than 95% for both cases.

The dramatic variations in the number of observed peaks, a function of sample processing differences, was expected to adversely affect the degree to which sample provenance could be determined. Thus, we were surprised that the hierarchical clustering analysis performed as well as it did in the establishment of product relatedness, and in its ability to show differences based on changes in sample processing steps. The observed dendrogram further demonstrated the capacity of multivariate statistical analysis processing of DART-HRMS data to enable species identification, even without the selection of feature masses. The entire MS dataset covering a nominal mass range of *m/z* 50-570 (depending on the sample) was used in the analysis. The two main branches of the dendrogram defined the two species, with the *P. methysticum* data and the *P. betle* data clustering independently (Fig. 11). The betel

essential oil and the *P. betle* leaves were further differentiated. Of note was the observation that the kava 70% powder, kava root powder and dried kava whole root all clustered together and were all from the same vendor, Bouncing Bear Botanicals. Additionally, the majority of the kava extract tincture samples clustered within a single clade, indicating that the tincture product was very different from the powders or whole and dried root products analyzed. Thus, although there were some misclassifications of the types of kava products in the dendrogram, the results demonstrated that plant species identification is possible and that some information regarding the processing of samples or the vendor can be obtained, even without a priori selection of feature masses for differentiation.

4. Conclusion

DART-HRMS analysis of *P. methysticum* and *P. betle* products revealed unique chemical signatures with characteristic biomarkers for the species. In-source CID was used to confirm the presence of a variety of compounds in each species in a rapid fashion. Statistical analysis of the DART-HRMS derived data indicated that the two species of psychoactive pepper were clearly differentiated with 100% validation. Application of unsupervised hierarchical clustering analysis also allowed distinctions to be made between *P. methysticum* and *P. betle*, and gave indications of processing treatment steps and vendor identity. The results presented here demonstrate how chemometric processing of DART-HRMS data can be directed towards the rapid determination of the origin of commercial plant-derived products. The method is efficient, can be broadly applied, and circumvents a number of time and resource-intensive sample preparation steps that can complicate some conventional methods of analysis.

5. Acknowledgements

The authors thank Kristen L. Fowble for technical assistance. The support of the Presidential Initiatives Fund for Research and Scholarship in Forensic Science and Cybersecurity from the University at Albany-State University of New York to RAM is gratefully acknowledged.

REFERENCES

1. M. Kokalj, J. Kolar, T. Trafela and S. Kreft, *Phytochem. Anal.*, 2011, **22**, 541-546.
2. F. Focke, I. Haase and M. Fischer, *J. Agric. Food Chem.*, 2013, **61**, 2943-2949.
3. F. De Mattia, I. Bruni, A. Galimberti, F. Cattaneo, M. Casiraghi and M. Labra, *Food Res. Int.*, 2011, **44**, 693-702.
4. P. Madesis, I. Ganopoulos, I. Sakaridis, A. Argiriou and A. Tsaftaris, *Food Res. Int.*, 2014, **60**, 163-172.
5. Z. Liu, X. Zeng, D. Yang, G. Chu, Z. Yuan and S. Chen, *Gene*, 2012, **499**, 76-80.
6. P. L. McCarthy, J. L. Hansen, R. S. Zemetra and P. H. Berger, *Biotechniques*, 2002, **32**, 560,562-564.
7. B. Tar'an, T. Warkentin, D. J. Somers, D. Miranda, A. Vandenberg, S. Blade and D. Bing, *Euphytica*, 2004, **136**, 297-306.
8. P. Wang, H. Xia, Y. Zhang, S. Zhao, C. Zhao, L. Hou, C. Li, A. Li, C. Ma and X. Wang, *BMC Genomics*, 2015, **16**, 21/21-21/14, 14 pp.
9. J. K. P. Weder, *Lebensm Wiss Technol*, 2002, **35**, 504-511.
10. S. Strgulc Krajsek, P. Buh, A. Zega and S. Kreft, *Chem. Biodivers.*, 2008, **5**, 310-317.
11. A. Umek, S. Kreft, T. Kartnig and B. Heydel, *Planta Med.*, 1999, **65**, 388-390.
12. B. Jiang, C. Ma, T. Motley, F. Kronenberg and E. J. Kennelly, *Phytochem. Anal.*, 2011, **22**, 339-351.
13. S. B. Matthews, M. Santra, M. M. Mensack, P. Wolfe, P. F. Byrne and H. J. Thompson, *PLOS One*, 2012, **7**, e44179.
14. A. D. Lesiak, R. B. Cody, A. J. Dane and R. A. Musah, *Anal. Chem.*, 2015, DOI: 10.1021/acs.analchem.5b01611.
15. E. O. Espinoza, C. A. Lancaster, N. M. Kreitals, M. Hata, R. B. Cody and R. A. Blanchette, *Rapid Commun. Mass Spectrom.*, 2014, **28**, 281-289.
16. R. A. Musah, E. O. Espinoza, R. B. Cody, A. D. Lesiak, E. D. Christensen, H. E. Moore, S. Maleknia and F. P. Drijfhout, *Sci. Rep.*, 2015, **5**.
17. A. D. Lesiak, R. B. Cody, A. J. Dane and R. A. Musah, *Forensic Sci. Int.*, 2014, **242**, 210-218.
18. *P. betle and P. methysticum*, The Encyclopedia of Psychoactive Plants: Ethnopharmacology and its Applications, Park Street Press, 2005.
19. *Kava: from Ethnology to Pharmacology*, CRC Press, 2004.
20. A. Ligresti, R. A. Villano, M., I. Ujvary and V. Di Marzo, *Pharmacol. Res.*, 2012, **66**, 163-169.
21. US Drug Enfrcement Administration, Kava, http://www.deadiversion.usdoj.gov/drug_chem_info/kava.pdf, (accessed 09.02.2015).
22. S. M. Musser, *Kava (Piper methysticum)*, Encyclopedia of Dietary Supplements, Marcel Dekker, 2005.
23. Y. N. Singh, *J. Ethnopharmacol.*, 1992, **37**, 13-45.
24. J. C. Almeida and E. W. Grimsely, *Ann Intern Med*, 1996, **125**, 940-941.
25. J. E. Freshour, B. Odle, S. Rikhye and D. W. Stewart, *J. Diet. Suppl.*, 2012, **9**, 149-154.
26. F. Tarbah, Y. Barguil, W. Weinmann, C. Mueller, D. Duhet, P. Cabalion, B. Kardel and T. Daldrup, GTFCh (Society of Toxicological and Forensic Chemistry) Symposium, Mosbach, Germany, 2003.
27. A. M. Hartmann, D. K. Raap and C. R. Geist, *Reviews in Food and Nutrition Toxicity*, 2003, **1**, 1-15.
28. I. Ujvary, *Ann. Ist Super Sanita*, 2014, **50**, 12-27.
29. S. A. Bhalerao, D. R. Verma, R. V. Gavankar, N. C. Teli, Y. Y. Rane, V. S. Didwana and A. Trikannad, *J. Pharmacogn. Phytochem.*, 2013, **1**, 10-19.
30. F. Fazal, P. P. Mane, M. P. Rai, K. R. Thilakchand, H. P. Bhat, P. S. Kamble, P. L. Palatty and M. S. Baliga, *Chin. J. Integr. Med.*, 2014, 1-11.

31. V. Bajpai, D. Sharma, B. Kumar and K. P. Madhusudanan, *Biomed. Chromatogr.*, 2010, **24**, 1283-1286.
32. N. Kumar, P. Misra, A. Dube, S. Bhattacharya, M. Dikshit and S. Ranade, *Curr. Sci.*, 2010, **99**, 922-932.
33. C. H. Lee, A. M. Ko, C. F. Yen, K. S. Chu, Y. J. Gao, S. Warnakulasuriya, Sunarjo, S. O. Ibrahim, R. B. Zain, W. K. Patrick and Y. C. Ko, *Br. J. Psychiatry*, 2012, **201**, 383-391.
34. V. Lebot, T. K. Do and L. Legendre, *Food Chem.*, 2014, **151**, 554-560.
35. L. D. Gautz, P. Kaufusi, M. C. Jackson, H. C. Bittenbender and C. S. Tang, *J. Agric. Food Chem.*, 2006, **54**, 6147-6152.
36. A. R. Bilia, M. C. Bergonzi, D. Lazari and F. F. Vincieri, *J. Agric. Food Chem.*, 2002, **50**, 5016-5025.
37. H. R. Dharmaratne, N. P. Nanayakkara and I. A. Khan, *Phytochem.*, 2002, **59**, 429-433.
38. L. S. de Jager, G. A. Perfetti and G. W. Diachenko, *Food Addit. Contam.*, 2004, **21**, 921-934.
39. X. G. He, L. Z. Lin and L. Z. Lian, *Planta Med.*, 1997, **63**, 70-74.
40. H. Vandenbroucke, P. Mournet, R. Malapa, J. C. Glaszmann, H. Chair and V. Lebot, *Genome*, 2015, **58**, 1-11.
41. S. Shetty and K. K. Vijayalaxmi, *Int. J. Pharm. Biol. Sci.*, 2012, **3**, B-344-B-349.
42. N. Singtongratana, S. Vadhanasin and J. Singkhonrat, *Nat. Sci.*, 2013, **47**, 614-623.
43. F. Ferreres, A. P. Oliveira, A. Gil-Izquierdo, P. Valento and P. B. Andrade, *Phytochem. Anal.*, 2014, **25**, 453-460.
44. B. K. Dwivedi, S. Kumar, C. Nayak and B. K. Mehta, *J. Med. Plants Res.*, 2010, **4**, 2252-2255.
45. R. Pandey, P. Chandra, M. Srivastva, K. R. Arya, P. K. Shukla and B. Kumar, *Anal. Method*, 2014, **6**, 7349-7360.
46. P. A. Whitton, A. Lau, A. Salisbury, J. Whitehouse and C. S. Evans, *Phytochem.*, 2003, **64**, 673-679.
47. Kava Root Liquid Tincture, http://www.herbal-island.com/kava-root-liquid-tincture/, (accessed 09.02.2015).
48. Kava 70% Kavalactone Powder, http://www.bouncingbearbotanicals.com/kava-kavalactone-powder-p-964.html#.VecgXflVhZY, (accessed 09.02.2015).

Table 1. DART-HRMS data of *P. methysticum* products at 20 V. The corresponding spectra are shown in Fig. 4.

Product	Compound	Formula	Measured	Calculated	Diff.[†]	Rel. Int.[‡]
Kava 40% Powder	Dehydrokavain[a]	$C_{14}H_{12}O_3 + H^+$	229.0854	229.0865	1.1	29.8
	Kavain[b]	$C_{14}H_{14}O_3 + H^+$	231.1002	231.1021	1.9	66.1
	Dihydrokavain[b]	$C_{14}H_{16}O_3 + H^+$	233.1152	233.1178	2.6	84.8
	Yangonin[b]	$C_{15}H_{14}O_4 + H^+$	259.0964	259.0970	0.6	71.0
	Tetrahydroyangonin[a]	$C_{15}H_{18}O_4 + H^+$	263.1288	263.1283	-0.5	8.6
	Methysticin[b]	$C_{15}H_{14}O_5 + H^+$	275.0911	275.0920	0.9	39.6
	Dihydromethysticin[b]	$C_{15}H_{16}O_5 + H^+$	277.1042	277.1076	3.4	31.4
	Flavokavain B[b]	$C_{17}H_{16}O_4 + H^+$	285.1124	285.1127	0.3	100.0
	Flavokavain A[b]	$C_{18}H_{18}O_5 + H^+$	315.1233	315.1233	0.0	90.0
Kava 70% Powder	Dehydrokavain[a]	$C_{14}H_{12}O_3 + H^+$	229.0862	229.0865	0.3	36.2
	Kavain[b]	$C_{14}H_{14}O_3 + H^+$	231.1009	231.1021	1.2	42.8
	Dihydrokavain[b]	$C_{14}H_{16}O_3 + H^+$	233.1159	233.1178	1.9	88.7
	Yangonin[b]	$C_{15}H_{14}O_4 + H^+$	259.0975	259.0970	-0.7	100.0
	Tetrahydroyangonin[a]	$C_{15}H_{18}O_4 + H^+$	263.1297	263.1283	-1.4	13.3
	Methysticin[b]	$C_{15}H_{14}O_5 + H^+$	275.0926	275.0920	-0.7	17.0
	Dihydromethysticin[b]	$C_{15}H_{16}O_5 + H^+$	277.1059	277.1076	1.7	32.6
	Flavokavain B[b]	$C_{17}H_{16}O_4 + H^+$	285.1104	285.1127	2.3	32.1
	Flavokavain A[b]	$C_{18}H_{18}O_5 + H^+$	315.1249	315.1233	-1.7	21.9
Kava Root Powder	Dehydrokavain[a]	$C_{14}H_{12}O_3 + H^+$	229.0856	229.0865	0.9	20.8
	Kavain[b]	$C_{14}H_{14}O_3 + H^+$	231.1004	231.1021	1.7	23.7
	Dihydrokavain[b]	$C_{14}H_{16}O_3 + H^+$	233.1154	233.1178	2.4	54.2
	Yangonin[b]	$C_{15}H_{14}O_4 + H^+$	259.0973	259.0970	-0.3	100.0
	Tetrahydroyangonin[a]	$C_{15}H_{18}O_4 + H^+$	263.1296	263.1283	-1.3	7.5
	Methysticin[b]	$C_{15}H_{14}O_5 + H^+$	275.0923	275.0920	-0.3	19.5
	Dihydromethysticin[b]	$C_{15}H_{16}O_5 + H^+$	277.1055	277.1076	2.1	19.3
	Flavokavain B[b]	$C_{17}H_{16}O_4 + H^+$	285.1101	285.1127	2.6	58.4
	Flavokavain A[b]	$C_{18}H_{18}O_5 + H^+$	315.1248	315.1233	-1.5	67.7
Kava Root	Dehydrokavain[a]	$C_{14}H_{12}O_3 + H^+$	229.0859	229.0865	0.6	49.4
	Kavain[b]	$C_{14}H_{14}O_3 + H^+$	231.1008	231.1021	1.3	78.2
	Dihydrokavain[b]	$C_{14}H_{16}O_3 + H^+$	233.1157	233.1178	2.1	84.3
	Yangonin[b]	$C_{15}H_{14}O_4 + H^+$	259.0978	259.0970	-0.8	100.0
	Tetrahydroyangonin[a]	$C_{15}H_{18}O_4 + H^+$	263.1299	263.1283	-1.6	8.3
	Methysticin[b]	$C_{15}H_{14}O_5 + H^+$	275.0926	275.0920	0.7	18.4
	Dihydromethysticin[b]	$C_{15}H_{16}O_5 + H^+$	277.1060	277.1076	1.6	11.2
	Flavokavain B[b]	$C_{17}H_{16}O_4 + H^+$	285.1139	285.1127	-1.2	64.7
	Flavokavain A[b]	$C_{18}H_{18}O_5 + H^+$	315.1252	315.1233	-1.9	27.4
Kava Tincture	Glycerin	$C_3H_8O_3 + H^+$	93.0530	93.0552	2.2	100.0
	Dehydrokavain[a]	$C_{14}H_{12}O_3 + H^+$	229.0862	229.0865	0.3	29.0
	Kavain[b]	$C_{14}H_{14}O_3 + H^+$	231.1009	231.1021	1.2	29.3
	Dihydrokavain[b]	$C_{14}H_{16}O_3 + H^+$	233.1158	233.1178	2.0	73.0
	Yangonin[b]	$C_{15}H_{14}O_4 + H^+$	259.0978	259.0970	-0.7	50.9
	Tetrahydroyangonin[a]	$C_{15}H_{18}O_4 + H^+$	263.1271	263.1283	1.2	4.8
	Methysticin[b]	$C_{15}H_{14}O_5 + H^+$	275.0932	275.0920	-1.3	5.8
	Dihydromethysticin[b]	$C_{15}H_{16}O_5 + H^+$	277.1063	277.1076	1.3	8.8
	Flavokavain B[b]	$C_{17}H_{16}O_4 + H^+$	285.1106	285.1127	2.1	38.7
	Flavokavain A[b]	$C_{18}H_{18}O_5 + H^+$	315.1219	315.1233	1.4	19.7

[a]The corresponding mass was consistent with the formula of the indicated compound which has previously been isolated from *P. methysticum*.
[b]The presence of this compound was confirmed through comparison of the in-source collision-induced dissociation (CID) spectrum of the kava powder with the in-source CID spectrum of an authentic standard.
[†]Differences are reported in millimass units (mmu). Measured masses fell within 5 mmu of the calculated masses.
[‡]Relative intensities are reported in percent.

Table 2. Mass measurement data of *P. betle* products at 20 V (Figure X).

Product	Compound	Formula	Measured	Calculated	Diff.[†]	Rel. Int.[‡]
Betel Essential Oil	Chavicol	$C_9H_{10}O + H^+$	135.0845	135.0810	-3.5	24.2
	α-Terpinene[a]	$C_{10}H_{16} + H^+$	137.1322	137.1330	0.8	17.1
	Eucalyptol	$C_{10}H_{18}O + H^+$	155.1441	155.1436	-0.5	2.7
	Isoeugenol[a]	$C_{10}H_{12}O_2 + H^+$	165.0917	165.0916	-0.2	100.0
	Allyl pyrocatechol acetate[a]	$C_{11}H_{12}O_3 + H^+$	193.0897	193.0865	-3.2	2.7
	β-Caryophyllene[a]	$C_{15}H_{24} + H^+$	205.1931	205.1956	2.5	89.8
	Eugenyl acetate[a]	$C_{12}H_{15}O_3 + H^+$	207.0998	207.1021	2.3	50.1
	Allyl pyrocatechol diacetate[a]	$C_{13}H_{14}O_4 + H^+$	235.0970	235.0970	4.3	4.3
	Unknown	$C_{13}H_{17}NO_4 + H^+$	252.1197	252.1236	3.9	20.8
P. betle *Leaves*	α-Terpinene[a]	$C_{10}H_{16} + H^+$	137.1293	137.1330	3.7	5.6
	Hydroxychavicol[a]	$C_9H_{10}O_2 + H^+$	151.0751	151.0759	0.8	5.4
	Isoeugenol[a]	$C_{10}H_{12}O_2 + H^+$	165.0887	165.0916	2.9	37.9
	Allyl pyrocatechol acetate[a]	$C_{11}H_{12}O_3 + H^+$	193.0838	193.0865	2.7	75.1
	β-Caryophyllene[a]	$C_{15}H_{24} + H^+$	205.1907	205.1956	4.9	11.6
	Eugenyl acetate[a]	$C_{12}H_{15}O_3 + H^+$	207.0977	207.1021	4.4	74.0
	Allyl pyrocatechol diacetate[a]	$C_{13}H_{14}O_4 + H^+$	235.0940	235.0970	3.0	49.4
	Unknown	$C_{13}H_{17}NO_4 + H^+$	252.1211	252.1236	1.4	100.0
P. betle *Root*	Cyclohexadiene[a]	$C_6H_8 + H^+$	81.0701	81.0704	0.3	93.2
	Hydroxycatechol[a]	$C_6H_6O_3 + H^+$	127.0393	127.0395	0.2	4.4
	α-Terpinene[a]	$C_{10}H_{16} + H^+$	137.1354	137.1330	-2.4	100.0
	Hydroxychavicol[a]	$C_9H_{10}O_2 + H^+$	151.0758	151.0759	0.1	13.1
	Allyl pyrocatechol acetate[a]	$C_{11}H_{12}O_3 + H^+$	193.0838	193.0865	2.7	14.7
	β-Caryophyllene[a]	$C_{15}H_{24} + H^+$	205.1932	205.1956	2.4	61.9
	Allyl pyrocatechol diacetate[a]	$C_{13}H_{14}O_4 + H^+$	235.0957	235.0970	1.3	4.8
	Unknown	$C_{13}H_{17}NO_4 + H^+$	252.1232	252.1236	0.4	34.6

[a]The corresponding mass was consistent with the formula of the indicated compound which has previously been isolated from *P. betel.*

[†]Differences are reported in millimass units (mmu). Measured masses fell within 5 mmu of the calculated masses.

[‡]Relative intensities are reported in percent.

Table 3. Feature masses used for principal component analysis of *P. methysticum* and *P. betle* products. The spectra from which these masses were selected are presented in Fig. 4 and 5.

93.05300	207.09990	259.09640
151.07539	229.08540	275.09110
165.08881	231.10020	277.10419
187.11200	233.11520	285.11240
193.08360	235.09380	292.13351
202.12090	247.09590	315.12329
205.19310	252.12100	-

RSC Author Templates - ChemDraw (CDX) - Double Column Artwork

All text and images must be placed within the frame.

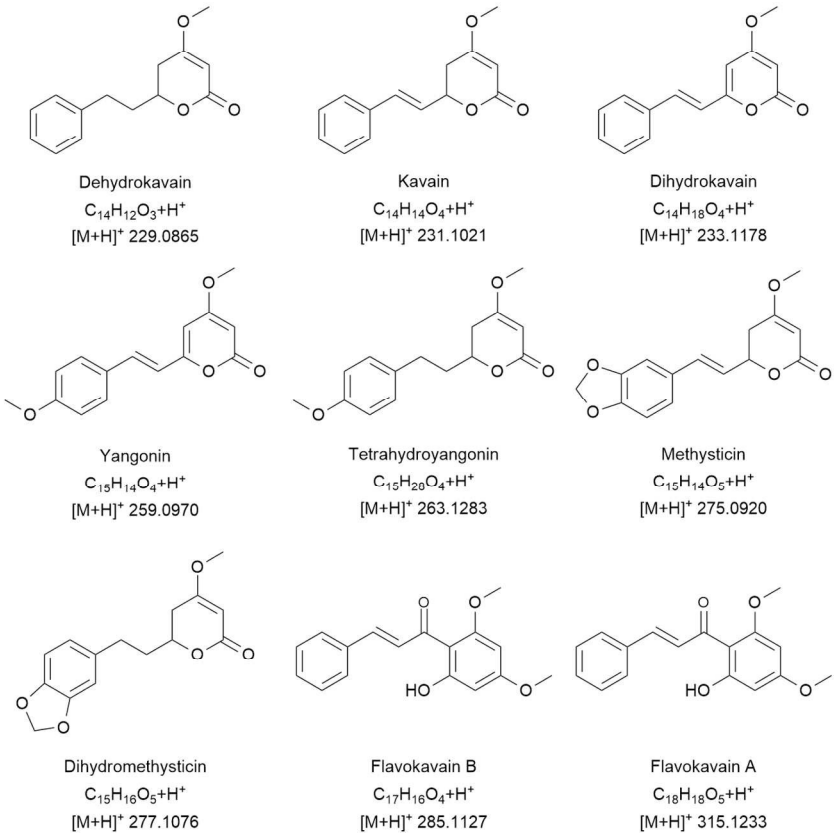


Fig. 1. Structures, formulas and protonated masses of seven kavalactones and two chalcones (flavokavains A and B) previously observed in *P. methysticum.*

RSC Author Templates - ChemDraw (CDX) - Single Column Artwork

All text and images must be placed within the frame.

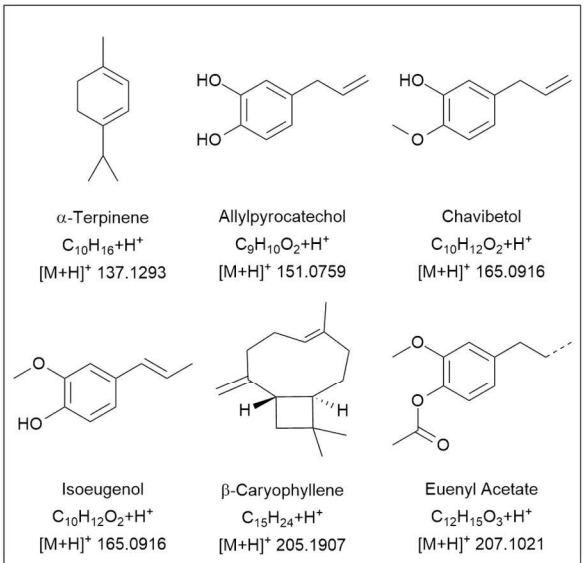


Fig. 2. Structures, formulas and protonated masses of five compounds previously observed in *P. betle*.

Fig. 3. *P. methysticum* and *P. betle* products analyzed in this study. Panel a: betel leaf essential oil; Panel b: kava root; Panel c: kava tincture; Panel d: *P. betle* plant; Panel e: kava 40% powder; Panel f: kava 70% powder; Panel g: kava root powder; and Panel h: *P. betle* root.

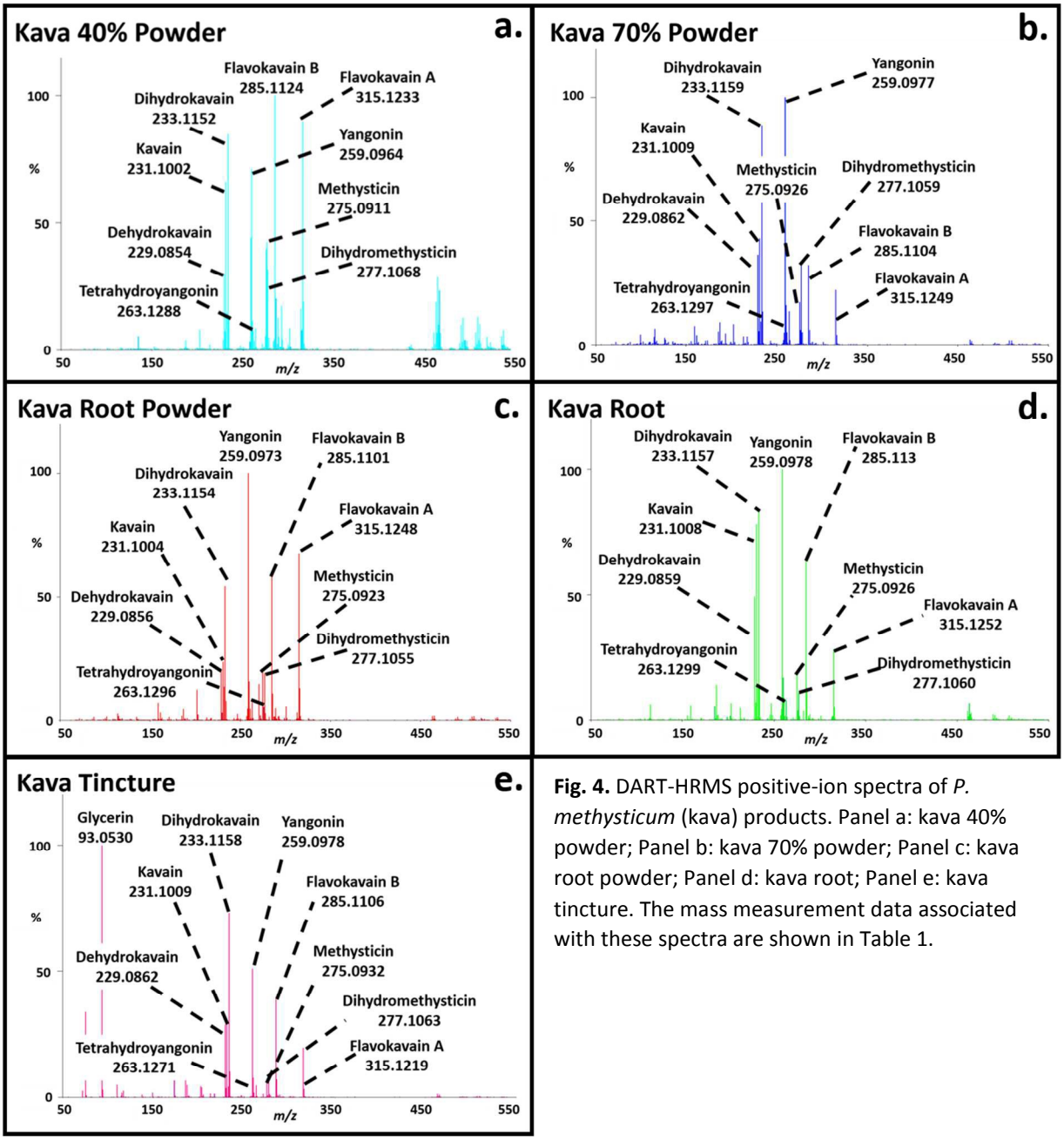


Fig. 4. DART-HRMS positive-ion spectra of *P. methysticum* (kava) products. Panel a: kava 40% powder; Panel b: kava 70% powder; Panel c: kava root powder; Panel d: kava root; Panel e: kava tincture. The mass measurement data associated with these spectra are shown in Table 1.

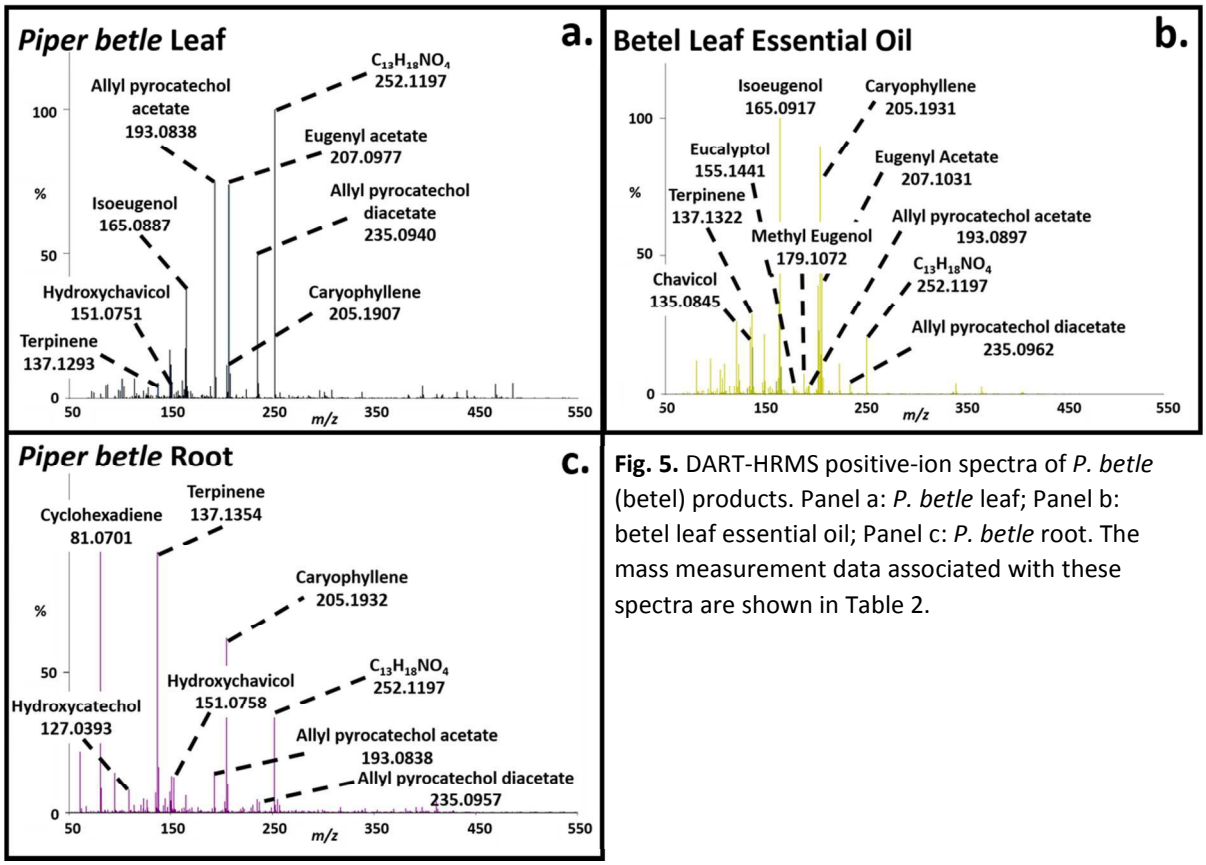


Fig. 5. DART-HRMS positive-ion spectra of *P. betle* (betel) products. Panel a: *P. betle* leaf; Panel b: betel leaf essential oil; Panel c: *P. betle* root. The mass measurement data associated with these spectra are shown in Table 2.

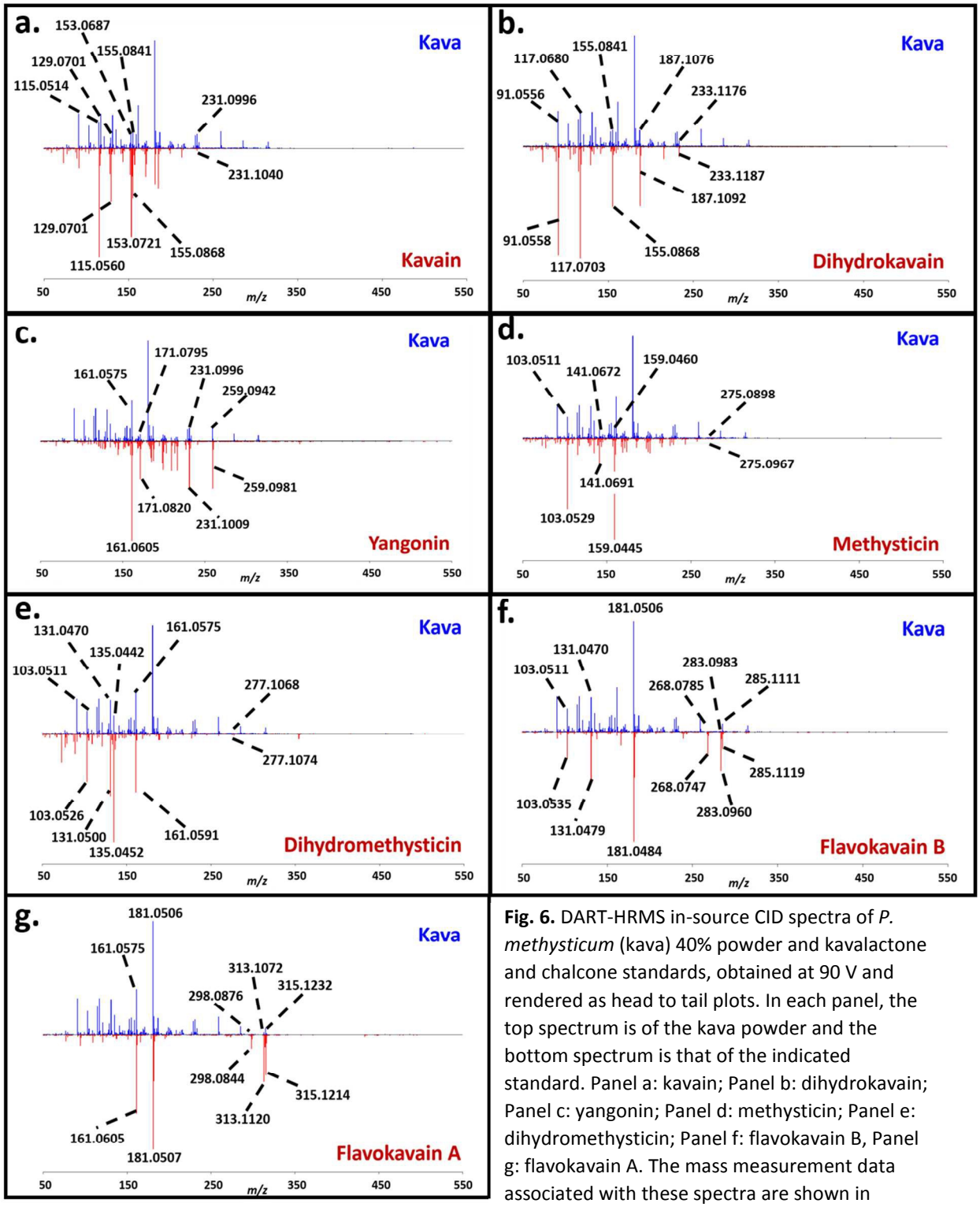


Fig. 6. DART-HRMS in-source CID spectra of *P. methysticum* (kava) 40% powder and kavalactone and chalcone standards, obtained at 90 V and rendered as head to tail plots. In each panel, the top spectrum is of the kava powder and the bottom spectrum is that of the indicated standard. Panel a: kavain; Panel b: dihydrokavain; Panel c: yangonin; Panel d: methysticin; Panel e: dihydromethysticin; Panel f: flavokavain B, Panel g: flavokavain A. The mass measurement data associated with these spectra are shown in Supplementary Tables 1-7.

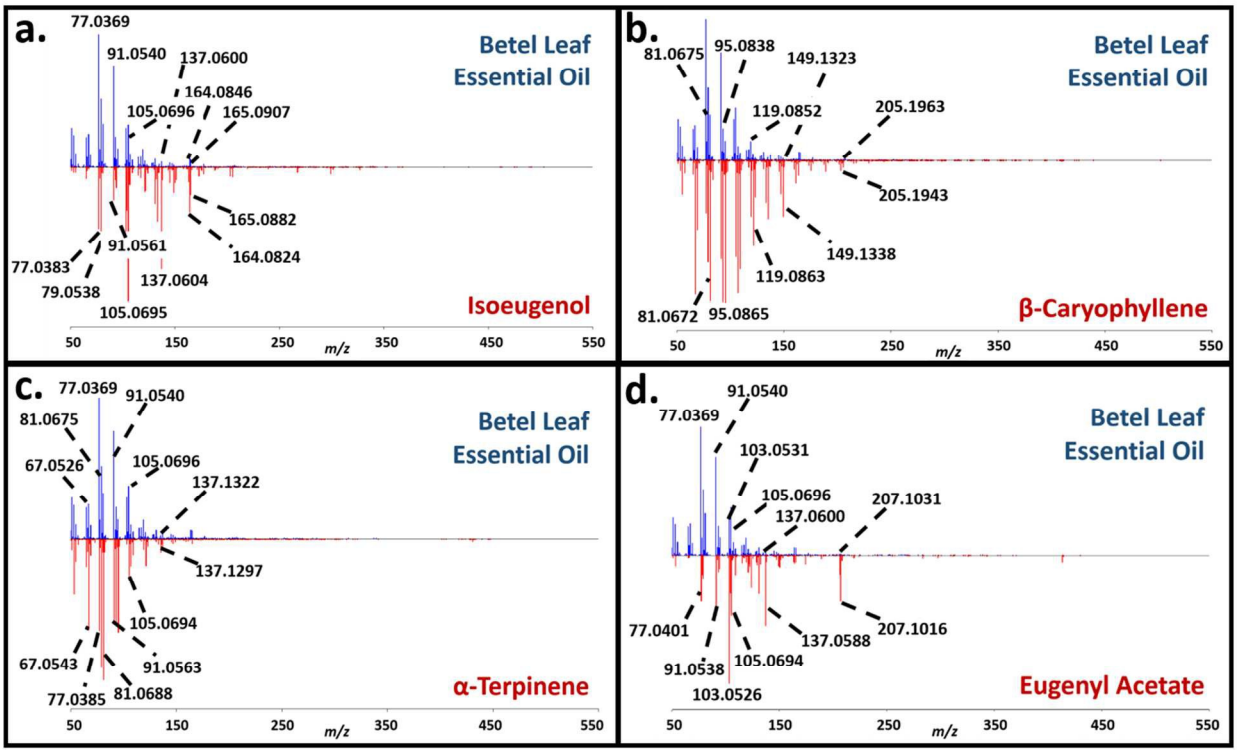


Fig. 7. DART-HRMS in-source CID spectra of Betel Leaf Essential Oil and various standards obtained at 90 V, rendered as head to tail plots. In each panel, the top spectrum is of the betel oil and the bottom spectrum is that of the indicated standard. Panel a: isoeugenol; Panel b: β-caryophyllene; Panel c: α-terpinene; Panel d: eugenyl acetate. The mass measurement data associated with these spectra are shown in Supplementary Tables 8-11.

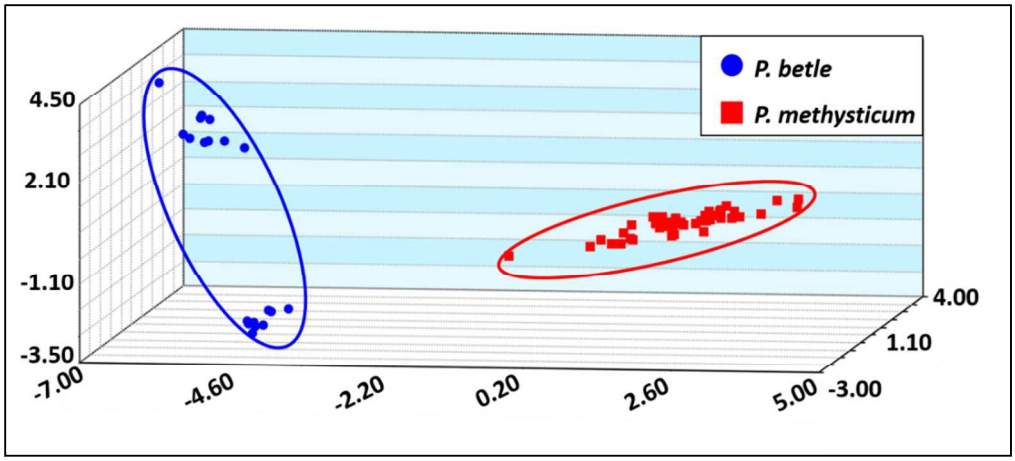


Fig. 8. Principal component analysis (PCA) plot of *P. methysticum* and *P. betle* products constructed using DART-HRMS derived data. Three principal components (PCs) accounted for 70.92% of the variation, and the leave-one-out cross validation (LOOCV) was 100%. The feature masses used for the PCA are listed in Table 3.

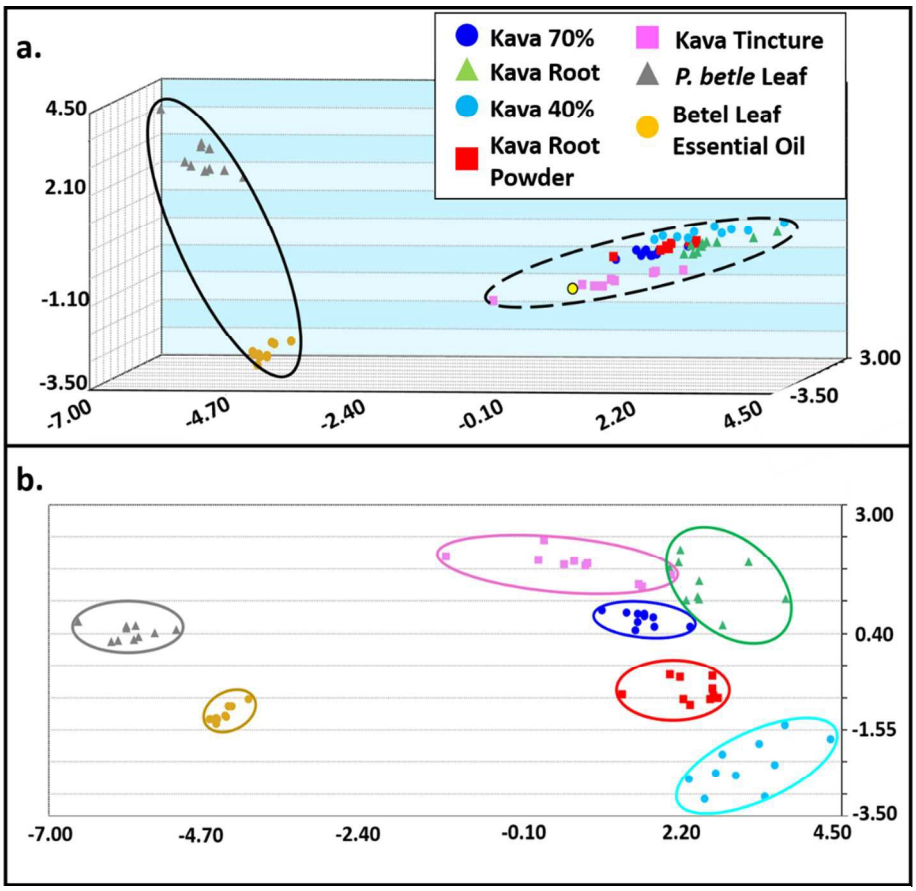


Fig. 9. Panel a: Principal component analysis (PCA) plot of *P. methysticum* and *P. betle* products using DART-HRMS data. Class assignments were based on differences in processing of plant materials. Panel b: Top-down view of the PCA plot of *P. methysticum* and *P. betle* products DART-HRMS data. Three principal components (PCs) accounted for 70.92% of the variation and the leave-one-out cross validation (LOOCV) was 95.72%. The feature masses used for PCA are listed in Table 3.

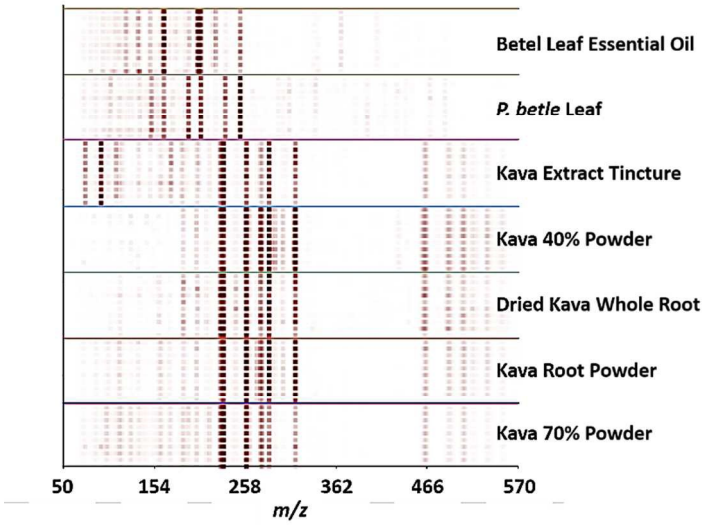


Fig. 10. Heat map renderings of the DART-HRMS spectra of *P. methysticum* and *P. betle* products. High intensity peaks are shown in dark red and lower intensity peaks are indicated by lighter shades.

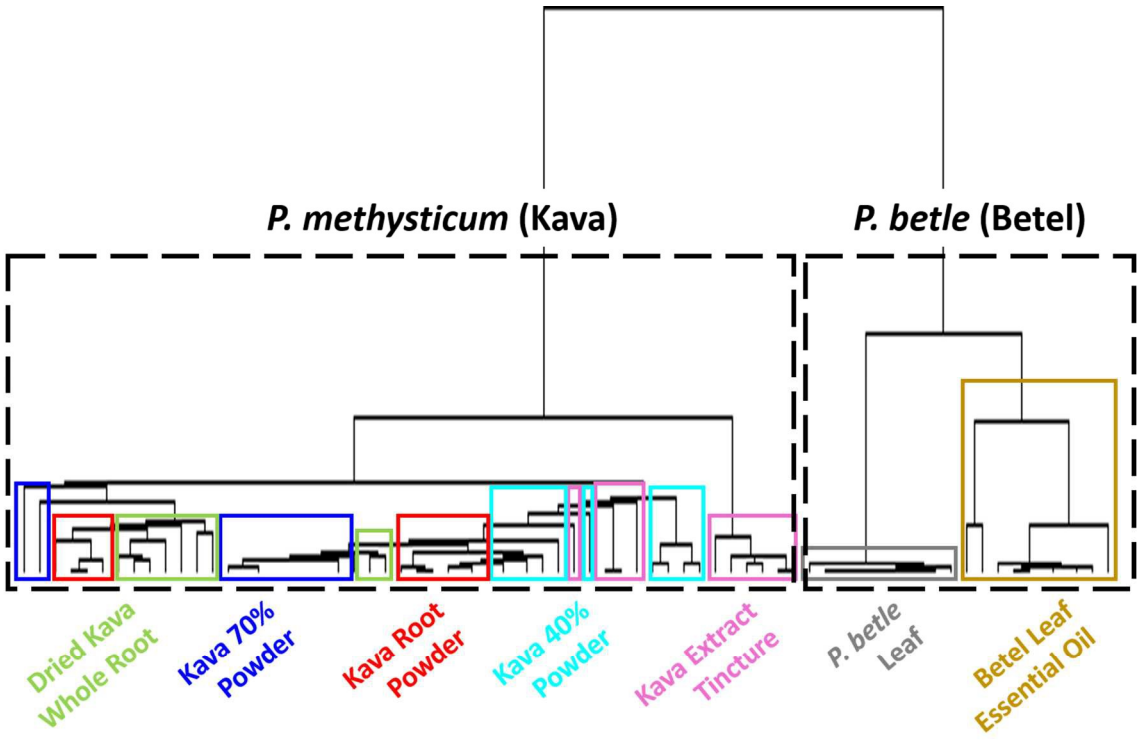
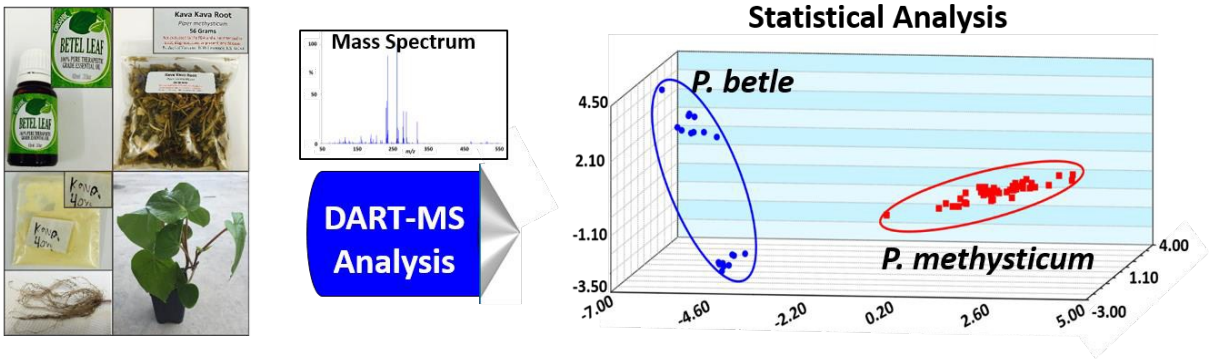


Fig. 11. Hierarchical clustering results obtained using the DART-HRMS derived data from analysis of *P. methysticum* and *P. betle* products. The species are clearly separated into two distinct branches in the dendrogram (enclosed in dashed-line boxes). The *P. betle* products are also resolved from one another. The *P. methysticum* products have some mis-groupings, but the species level classification remains correct.

Direct analysis in real time mass spectrometry coupled with statistical analysis tools were applied to deduce the plant species from which a variety of commercial products were derived.



Processed plant-derived products

Definitive species-level identification independent of product processing steps