

Maomao Zhang, Guangjie Zhao, Juan Guo, Alex C. Wiedenhoeft, Charles C. Liu and Yafang Yin*

Timber species identification from chemical fingerprints using direct analysis in real time (DART) coupled to Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS): comparison of wood samples subjected to different treatments

<https://doi.org/10.1515/hf-2018-0304>

Received December 22, 2018; accepted April 23, 2019; previously published online June 14, 2019

Abstract: Timber genus identification based on the anatomical features of wood is well established in botany. However, species-level wood identification is not always possible based on traditional wood morphology techniques alone. To compensate for the deficiencies of traditional methods, direct analysis in real time coupled to Fourier transform ion cyclotron resonance mass spectrometry (DART-FTICR-MS) was used to obtain the mass spectral fingerprints of different timber species. Using heartwood samples of two morphologically similar species, *Pterocarpus santalinus* and *Pterocarpus tinctorius*, subjected to different treatments, i.e. solvent extractions and powdered samples as well as air-dried samples and samples dried at low and high temperatures, we observed distinct chemical signatures for the wood samples from the two species, enabling rapid species-level identification when multivariate statistical analysis was adopted.

*Corresponding author: Yafang Yin, Department of Wood Anatomy and Utilization, Research Institute of Wood Industry, Chinese Academy of Forestry, Beijing 100091, China; and Wood Collections (WOODPEDIA), Chinese Academy of Forestry, Beijing 100091, China, e-mail: yafang@caf.ac.cn

Maomao Zhang: Department of Wood Anatomy and Utilization, Research Institute of Wood Industry, Chinese Academy of Forestry, Beijing 100091, China; College of Materials Science and Technology, Beijing Forestry University, Beijing 100083, China; and Wood Collections (WOODPEDIA), Chinese Academy of Forestry, Beijing 100091, China

Guangjie Zhao: College of Materials Science and Technology, Beijing Forestry University, Beijing 100083, China

Juan Guo: Department of Wood Anatomy and Utilization, Research Institute of Wood Industry, Chinese Academy of Forestry, Beijing 100091, China; and Wood Collections (WOODPEDIA), Chinese Academy of Forestry, Beijing 100091, China

Alex C. Wiedenhoeft: Center for Wood Anatomy Research, USDA Forest Service, Forest Products Laboratory, Madison, WI 53726, USA; and Department of Botany, University of Wisconsin, Madison, WI 53706, USA

Charles C. Liu: ASPEC Technologies Limited, Beijing 100101, China

The supervised orthogonal partial least squares discriminant analysis (OPLS-DA) models for samples subjected to different treatments all exhibited accurate differentiation performance of the explained fraction of variance of classes ($R^2Y=0.936-0.987$) and the cross-validated fraction of variance of classes ($Q^2=0.857-0.949$). Compared with solvent types and the physical form of the sample, the drying treatment method had a greater impact on the chemical fingerprint from DART-FTICR-MS. Air-dried wood chips were the optimal samples for the DART-FTICR-MS method coupled with statistical analysis.

Keywords: air-dried wood chips, DART-FTICR-MS, metabolomics analysis, OPLS-DA, species-level wood identification

Introduction

With the growing consumption of wood products, the international timber trade has developed rapidly, and illegal logging and trade driven by large profits have also increased (Scotland and Ludwig 2002; Brack 2003; Tacconi 2012). The prevalence of illegal logging has caused enormous pressure on ecological environments. Forensic wood identification can provide direct and strong evidence in the monitoring and legal enforcement of illegal logging and trade. The abilities of forensic wood analysis to achieve species-level identification have a substantial impact on the protection of forest resources. Timber genus identification based on the anatomical features of wood is well established in botany. However, traditional wood identification techniques cannot always reach a species level using wood morphology alone (Gasson 2011; Jiao et al. 2014; Dormontt et al. 2015; Hartvig et al. 2015; Yu et al. 2017; Jiao et al. 2018). Although DNA barcoding has recently been successfully used for forensic wood identification at the species level, its application is limited by the amount of time required, the difficulty of extracting DNA from wood, the professional experience necessary and the limited reference libraries established (Yu et al. 2017; Jiao et al. 2018).

Obviously, the development of a rapid and reliable system for identifying timber species is in high demand.

Differences in the chemical profiles of wood tissues are also outward manifestations of the genomic differences between different tree species. Each species should have unique metabolomics features regardless of variability in plant age, growth conditions or other factors (Giffen et al. 2017). Thus, wood metabolome profiling could potentially be used for timber species identification, wood fluorescence analysis and evaluation of the biological degradation of waterlogged or archeological wood samples. However, most previous studies on wood metabolome fingerprinting and profiling have focused on medicinal plants, especially agarwood (Espinoza et al. 2014). Due to their bioactive chemical components and potential as sources of important new drugs, studies on the metabolomes of medicinal plants have become common for the identification and quantification of bioactive metabolites from natural sources (Newman and Cragg 2012; De Combarieu et al. 2015), the quality control of medicinal plants and herbal medicines and combating counterfeit medicines (Wang et al. 2004; Li et al. 2015) and the discovery of metabolic biomarkers (Frédérich et al. 2010). Moreover, metabolomic profiles of wood tissues from medicinal plants have been used in forensic wood identification (Gao et al. 2014).

Mass spectrometry (MS) has attracted attention in analytical metabolomics due to its advantages of reduced sample consumption, rapid data acquisition and high sensitivity and specificity (Moco et al. 2007; Gowda and Djukovic 2013; Kim and Heyman 2018). However, most MS applications require a separation step before mass detection, typically gas chromatography (GC), liquid chromatography (LC) or capillary electrophoresis (CE). Although mass spectrometric methods are suitable for the detection of metabolites in complex mixtures when used in conjunction with a separation technique, the detection process can be tedious and time consuming. The sample preparation procedure, including solvent choice, sequential extractions, concentrations and even chemical derivatization, must usually be optimized to obtain satisfactory results. Furthermore, the profile of molecules detected in the final extract is largely dependent on the solvents chosen due to the differential uptakes of compounds in the selected solvent (Moco et al. 2007; Lesiak et al. 2015). The laborious sample preparation or derivatization procedures and lengthy chromatographic separations prior to mass analysis in the aforementioned traditional methods for metabolite profiling greatly constrain the development of metabolomics analyses of complex mixtures, such as wood.

Direct analysis in real time mass spectrometry (DART-MS) (Cody et al. 2005), which has been attracting increasing attention, can be used for the direct analysis of gases, liquids and solids with minimal or even no sample preparation in open air ambient conditions, which effectively overcomes the limitations of conventional analytical methods. Therefore, DART-MS is gradually being applied in an increasing number of fields such as the food (Kim et al. 2011; Farré et al. 2013; Gómez-Ríos et al. 2017), security (Laramee et al. 2008; Harris et al. 2011), environmental (Zhou et al. 2015; Gómez-Ríos et al. 2017), pharmaceutical and botanical (Ma et al. 2018; Wang et al. 2018) industries. Both polar and nonpolar small molecule metabolites in wood tissues can be detected by DART-MS. This technique is considered promising for the rapid classification of white oak and northern red oak (Cody et al. 2012), the analysis of selected *Dalbergia* and trade timber (Lancaster and Espinoza 2012a,b), evaluating and distinguishing agarwood products (Lancaster and Espinoza 2012a,b; Espinoza et al. 2014), the discrimination of selected Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES)-protected Araucariaceae (Evans et al. 2017) and source identification of western Oregon Douglas-fir wood cores (Finch et al. 2017).

In forensic wood identification to promote legal logging and timber trade, the size and physical form of the obtained wood sample often vary. Both powder and solid wood samples derived from illegal logs, furniture or wooden crafts are possible as forensic wood samples. In addition, wood drying is an essential step in wood processing to prepare wood products with suitable dimensional stability. Air drying and low- and high-temperature kiln drying are common drying methods used early on in wood treatment processes. To date, although a number of studies have investigated the efficacy of wood identification using DART-MS, very few studies have focused on the effect of wood samples subjected to different treatments on the accuracy of timber species identification using DART-MS methods, which is essential to further apply DART-MS for wood identification.

The *Pterocarpus* genus, from the Leguminosae family, consists of approximately 35 species (Mabberley 2017). *Pterocarpus* spp. plants are one of the most important taxa worldwide due to their high commercial value in furniture, crafts and dyes. In addition to its high utilization in the wood industry, *Pterocarpus* is also known for its medicinal properties. Earlier studies have also established that the genus *Pterocarpus* is a rich source of polyphenolic compounds such as isoflavonoids, pterocarpanes and isoflavones (Seshadri 1972). For these reasons, the *Pterocarpus* species are threatened by illegal logging, and the number

of wild species has decreased sharply. Therefore, *Pterocarpus santalinus* (L.f.) was listed in the CITES Appendix II in 1995, and it was listed as near threatened in the International Union for Conservation of Nature (IUCN) Red List (IUCN 2019). On the other hand, non-CITES listed *Pterocarpus tinctorius* (Welw) has very similar anatomical features and commonly appears as a finished product to that of *P. santalinus*. Therefore, distinguishing among species can be difficult, and these challenges have led to the misidentification of woods in the *Pterocarpus* genus both in the market and in inspections. The greatest challenge is confusion between *P. santalinus* and *P. tinctorius* due to the substantial difference in their values and trade restrictions.

Therefore, the goal of this study was the differentiation of two commercial *Pterocarpus* timber species from chemical fingerprints obtained using DART-FTICR-MS. Heartwood samples from two morphologically similar species, *P. santalinus* and *P. tinctorius*, were subjected to different treatment processes, i.e. solvent extractions and powdered samples as well as air-dried samples and samples dried at low (70°C) and high (120°C) temperatures: (1) to explore the effects of solvent types, sample physical form and drying treatment on the chemical fingerprint of the wood by DART-FTICR-MS, and (2) to identify marker compounds for the differentiation of *P. santalinus* and *P. tinctorius* wood using DART-FTICR-MS coupled with metabolomics.

Materials and methods

Collection of wood samples: Twelve specimens of *P. santalinus* heartwood and 17 specimens of *P. tinctorius* heartwood were collected from vouchered or validated xylarium collections (Wood Collections of Chinese Academy of Forestry, WOODPEDIA) (Table S1). Of these 29 specimens, nine specimens of *P. santalinus* and 11 specimens of *P. tinctorius* were randomly selected as the training set. The remaining nine heartwood specimens were used as the test set. Only seven specimens were used for the test set for the powdered and extracted samples as having only two specimens in the test set is insufficient for evaluating the effects of powdering and solvent extraction.

Chemicals: Ethanol absolute, ethanol (95%) and benzene were purchased from Beijing Chemical Works, Beijing, China. Ethyl acetate was bought from Fuchen Chemical Reagent Company, Tianjin, China.

Wood chip preparation: Three sets of wood chips less than 2 mm thickness were cut from each specimen, and the surface of each wood chip was planed by grinding with 100-grit sandpaper to remove the exposed surfaces, which may contain chemical contaminants. Triplicate wood chip samples were subsequently air dried, dried at low temperature (70°C for 20 days) and dried at high temperature (120°C for 10 days) and labeled as AD, LD and HD, respectively. Then, all the dried wood chips were conditioned at 25°C and 60% relative humidity (RH) for 30 days.

Solvent extractions: Wood tissue samples (200 mg) were obtained from each specimen and then frozen in liquid nitrogen and ground into a fine powder using a 6770 freezer/mill apparatus (2 min) from SpexSamplePrep (Metuchen, NJ, USA). The wood powder (5 mg) was extracted successively with 1 ml each of distilled water, 1:1 EtOH:H₂O (ethanol mixed with distilled water in a 1:1 ratio), ethyl acetate and benzene-ethanol (benzene mixed with ethanol in a 2:1 ratio). The mixtures of the powder and solvent were sonicated in an ultrasonic bath for 30 min at 25°C and then centrifuged at 750 × g for 2 min at 25°C (Sorvall ST 8R centrifuge, Thermo Scientific). The supernatant was removed and filtered through a 0.45-µm membrane filter for subsequent analysis.

DART-FTICR-MS experiments: For DART-FTICR-MS analyses, the system consisted of a DART-SVP ion source (IonSense, Saugus, MA, USA) and a 12 T Bruker solariX XR FTICR-MS (Bruker Daltonics, Bremen, Germany) in a positive ion mode. Helium was used at a flow rate of approximately 2.0 L/s. Standard settings were used for the other DART source parameters with the exception of positive ion mode and a gas beam temperature of 350°C. The MS analysis used a mass range of 50–1000 m/z and a resolution of 1 M full width at half maximum (FWHM).

A solid wood chip was placed in the gap between the ion source and the mass spectrometer inlet with tweezers, and the mass spectrum was then obtained. To analyze the extracts, a capillary was dipped in the extract, and the extract remaining on the surface of the capillary was analyzed by placing the capillary in the gap between the ion source and the mass spectrometer inlet. To analyze the powders, the surfaces of capillaries were coated in distilled water, which allowed powdered sample to stick to the surface of the capillary. Then, the powder remaining on the surface of the capillary was analyzed by DART-FTICR-MS.

The acquired MS data was processed using DataAnalysis™ (Bruker Daltonics, Billerica, MA, USA). The mass data generated by the DataAnalysis™ software were aligned with a mass tolerance of 0.025 m/z using MetaboAnalysis 4.0 (www.metaboanalyst.ca) (Chong et al. 2018) and then exported for subsequent multivariate statistical analysis.

Chemometric analysis including unsupervised principal component analysis (PCA), hierarchical cluster analysis (HCA) and supervised orthogonal partial least squares discriminant analysis (OPLS-DA) were used in this study. PCA and OPLS-DA were conducted using SIMCA-P 14.1 software (Umetrics, Umea, Sweden). HCA and heatmap analysis were conducted with R software (version 3.3.3). A paired Student t-test was used to determine the significance of the effects of the variables on the classification, and this test was carried out using SPSS 22.0 Software (SPSS, Chicago, IL, USA).

Results and discussion

DART-FTICR-MS analysis of *P. santalinus* and *P. tinctorius*

To determine the chemical fingerprints of heartwood from *P. santalinus* and *P. tinctorius*, representative samples were analyzed by DART-FTICR-MS. Representative

DART-FTICR-MS spectra of AD chips of *P. santalinus* and *P. tinctorius* are shown in Figure 1. The differences between the chemical signals detected in the mass spectra of *P. santalinus* and *P. tinctorius* are remarkable. The *P. santalinus* samples showed abundant ions in the range of 200–800 m/z, whereas *P. tinctorius* showed few high-abundance ions. The spectra of *P. santalinus* and *P. tinctorius* are quite similar, and they have a number of common peaks such as 221.19, 254.21, 257.11, 258.24, 272.22 and 273.11 m/z, but significant differences were observed in several specific peaks, such as 473.36 and 477.39 m/z, which were almost absent in the spectra of *P. tinctorius*.

Multivariate statistical analysis of the DART-FTICR-MS data for species identification

PCA is a common method for assessing unprocessed analytical data as it can reduce the number of variables through a linear combination of the original variables and then reconstruct them with the ordered principal components based on their contributions (Hur et al. 2009; Cajka et al. 2013). PCA was performed to evaluate the differences in the DART-FTICR-MS spectra of *P. tinctorius* and *P. santalinus* and provided an overview of the results. The score plot of the AD chips showed clustering behavior related to the wood species (Figure S1, panel a). Based on these mass data, all the vouchered or validated samples were divided into two independent groups. Group I was composed completely of *P. santalinus* samples, and group II contained only *P. tinctorius* samples. Similar results were also observed in the samples of the LD chips, the 1:1 EtOH:H₂O extract (EH), the ethyl acetate extract (EA), the

benzene-ethanol extract (BE) and the powder (Figure S1). Weak classification was observed in the score plots of samples from the HD chips and the distilled water extract (DWE) as the wood species overlapped in the two groups (Figure S1, panels c and d).

In addition, OPLS-DA is usually conducted to reveal the classification capacity and to identify differential metabolites that can serve as marker compounds. Eight classification models based on OPLS-DA using the samples from the training set were used to explore the effect of the physical form of the sample, the drying treatment and solvent extraction on the classification performance. The score plots from the OPLS-DA models based on the samples from the eight treatment conditions are shown in Figure 2. Regardless of the treatment conditions, all the samples were clustered into two groups according to their category. Analysis of the eight models provided $R^2X=0.488-0.713$, $R^2Y=0.936-0.987$ and $Q^2=0.857-0.949$ (Table 1), which indicated that these OPLS-DA models are highly predictive. When the models were further evaluated using the test set, AD chips and LD chips were identified with the highest accuracy, 100%, but the accuracy of the identification of the HD chips was only 66.67% (Table 1). The mass ions and their relative intensities in the spectra of the AD and LD chips were similar, and the base peaks of *P. santalinus* and *P. tinctorius* were 473.36 and 257.11 m/z, respectively. Although most of the peaks detected in the AD and LD chips were also present in the HD chips, the base peaks of *P. santalinus* samples varied from 473.36 to 257.11 m/z, which made them easy to confuse with samples of *P. tinctorius*, and this may be why the lowest accuracy was achieved with the HD chips. Notably, the test set of AD chips was identified correctly using the model built on the LD chips and *vice versa*, which further confirmed that the

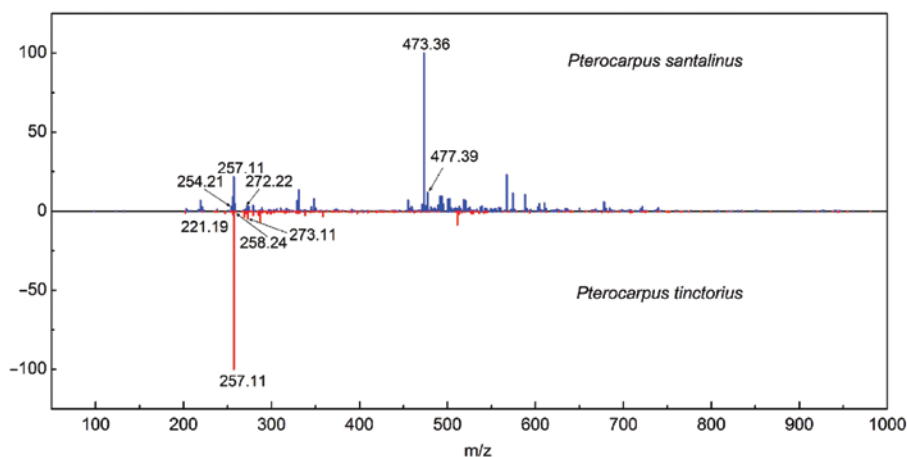


Figure 1: Representative DART-FTICR mass spectra of air-dried (AD) wood chips of *P. santalinus* (blue) and *P. tinctorius* (red).

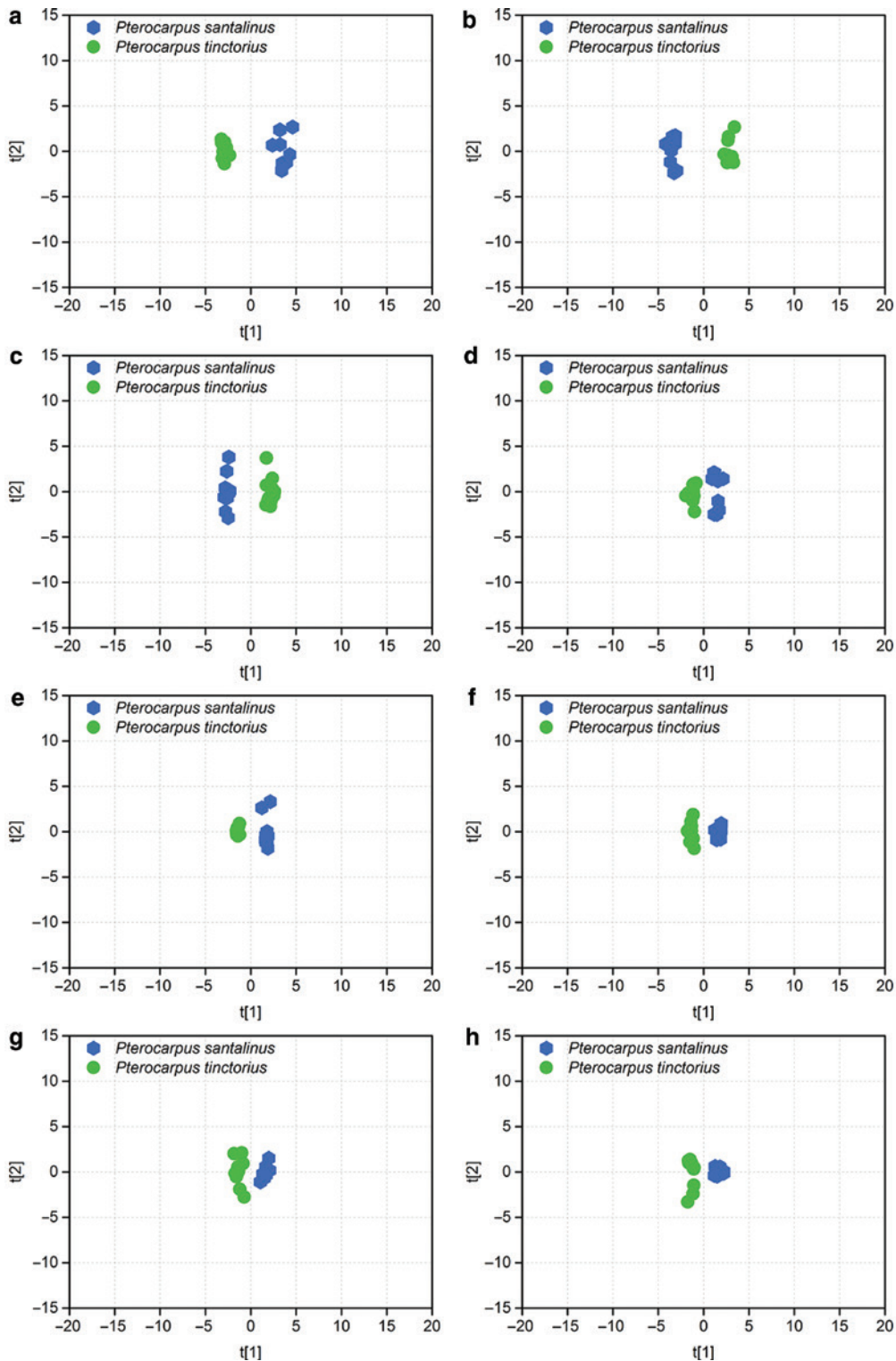


Figure 2: OPLS-DA score plots.

(a) Air-dried (AD) wood chips. (b) Low-temperature dried (LD) wood chips. (c) High-temperature dried (HD) wood chips. (d) Distilled water extract. (e) 1:1 EtOH:H₂O extract. (f) Ethyl acetate extract. (g) Benzene-ethanol extract. (h) Powdered samples.

low-temperature drying treatment had little impact on the classification performance. This is the first investigation of wood materials subjected to different drying treatments

using the DART-FTICR-MS. Similar results have been found between AD and LD chips, but the result obtained herein for HD chips is unique. These results elucidate the effect

Table 1: Classification capacity of different models based on OPLS-DA.

Models	R ² X	R ² Y	Q ²	Accuracy (%)
Air-dried wood chips	0.488	0.936	0.92	100.00 (9/9) ^a
Air-dried wood chips based on the five identified marker compounds	0.981	0.915	0.869	100 (9/9) ^a
Low-temperature dried wood chips	0.515	0.987	0.949	100.00 (9/9) ^a
High-temperature dried wood chips	0.694	0.987	0.87	66.67 (6/9) ^a
Distilled water extract	0.557	0.937	0.857	85.71 (6/7) ^a
1:1 EtOH:H ₂ O extract	0.713	0.987	0.941	85.71 (6/7) ^a
Ethyl acetate extract	0.544	0.979	0.891	100.00 (7/7) ^a
Benzene-ethanol extract	0.526	0.954	0.909	85.71 (6/7) ^a
Powdered samples	0.552	0.959	0.928	100.00 (7/7) ^a

^aNumber of samples correctly classified/number of samples in the test set.

of the drying treatment on the classification of *P. santalinus* and *P. tinctorius* by DART-FTICR-MS. These effects may make identifying wood species using DART-FTICR-MS more complicated as whether the wood samples have been dried or subjected to other heat treatments before analysis must be considered.

Previous reported analyses of plant seeds have demonstrated that extraction is not required because the DART-HR-TOF-MS spectrum of liquid extracts (using a 1:1:1 mixture v/v/v of ethyl acetate, ethanol and water as the extractant) was very similar to that obtained by direct analysis of the seeds (Lesiak et al. 2015). However, in those studies, the compounds of interest in the seeds were alkaloids, and it is unclear if this result is applicable to polyphenolic compounds in *Pterocarpus*. For this reason, comparing the chemical profile and the statistical results of wood chips to those of different wood extracts is of great importance. Several solvents commonly used for wood extraction including distilled water, ethyl acetate, benzene-ethanol and 1:1 EtOH:H₂O were selected for the extraction studies. The obtained chemical profiles are shown in Figures 3 and 4, and the OPLS-DA results are listed in Table 1. Unique chemical fingerprints were obtained from the four different extracts, and different trends were observed for *P. santalinus* and *P. tinctorius*. For *P. santalinus*, more high-abundance peaks were found in both the EH and EA. Ethyl acetate and benzene-ethanol proved to be more effective extraction solvents for *P. tinctorius* based on the large number of high-intensity peaks observed. Regardless of species, distilled water was the worst solvent as the fewest high-intensity peaks were detected by DART-FTICR-MS.

The predictive accuracy for EAs of the samples of the test set was highest (100%). However, for the DWEs, BEs and EHs, one sample was incorrectly classified in each set, and the accuracies were all 85.71%. This phenomenon may

be closely associated with the differences in the number of peaks and their relative abundance that may arise from the differential uptakes of compounds by the solvents. Ethyl acetate proved to be the most effective extraction solvent due to the greater number of peaks in the DART-FTICR-MS spectrum and the high predictive performance. Although different statistical results were obtained from the four extracts, the solvent type had a greater effect on the chemical profile than the classification performance based on OPLS-DA.

As an ambient ionization MS method, DART-MS has been applied to the analysis of gases, liquids and solids without requiring sample preparation (Cody et al. 2005). In addition to these applications, powdered samples can also be evaluated by placing them in the gap between the mass spectrometer and the ion source (Kim et al. 2011, 2015). Powdered *P. santalinus* wood is used medicinally to control hemorrhage, bleeding piles and inflammation (Arunakumara et al. 2011). A method for identifying *P. santalinus* powder is of great interest as abuse and adulteration in the industry and in the market place are increasing. Small amounts of *P. santalinus* and *P. tinctorius* powder were stuck to the surface of a capillary and then directly introduced into the ion source. The model established based on the powdered samples showed an accuracy of 100%, which is better than the models based on the DWE, BE and EH (Table 1). The number of ions detected in the powdered samples was greater than the number detected in the AD chips but less than that in the EA. Similar classification accuracies were obtained with the AD chips, EA and the powdered samples. Although noticeable differences were observed with different drying treatments, the solvent type and the physical form of the sample had little influence on the classification performance based on OPLS-DA. Considering the complexity of extract preparation and the toxicity associated with the chemical reagents, the analysis of

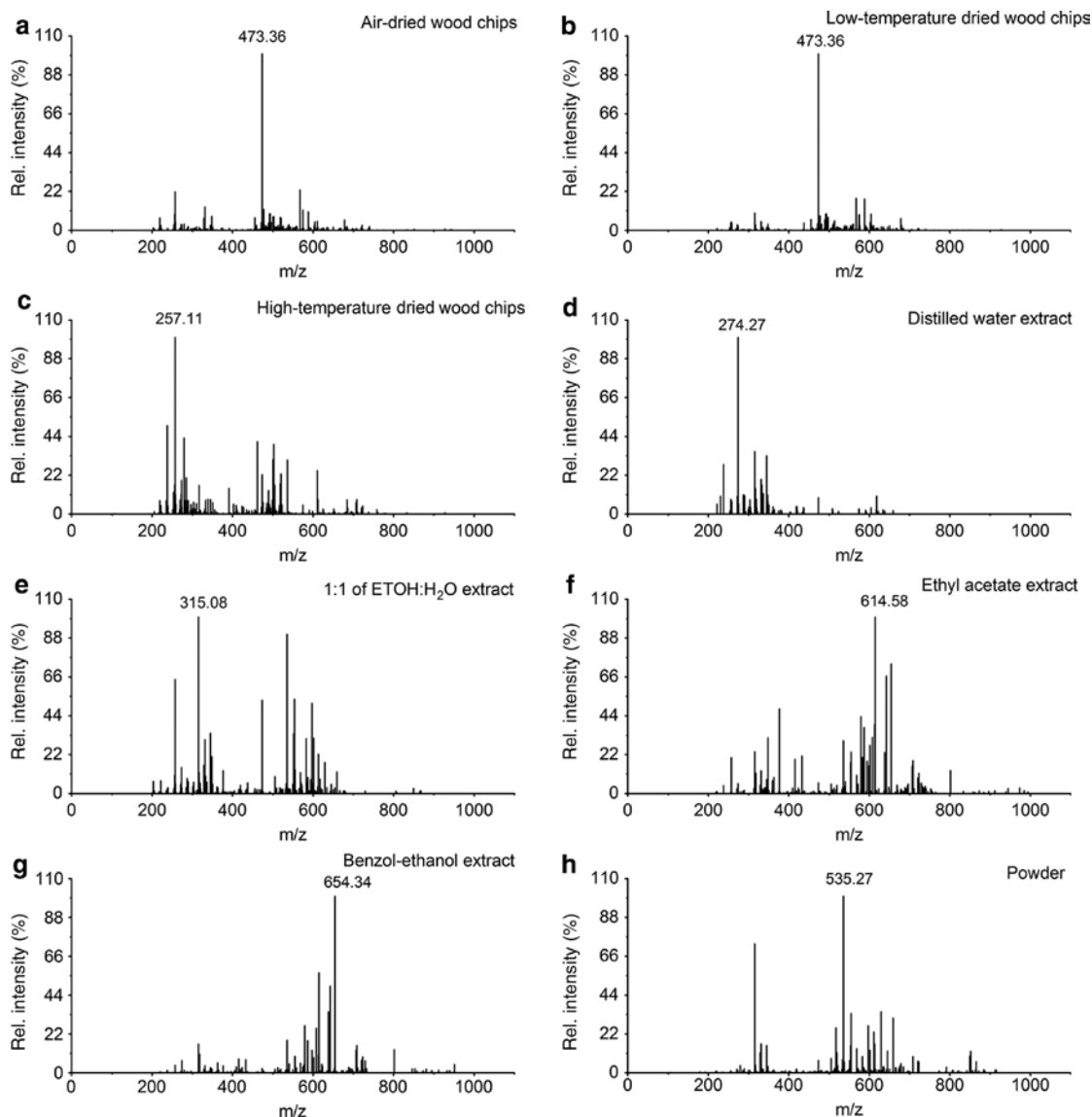


Figure 3: Representative DART-FTICR mass spectra of *P. santalinus*.

(a) Air-dried (AD) wood chips. (b) Low-temperature dried (LD) wood chips. (c) High-temperature dried (HD) wood chips. (d) Distilled water extract. (e) 1:1 EtOH:H₂O extract. (f) Ethyl acetate extract. (g) Benzene-ethanol extract. (h) Powdered samples.

wood extracts by DART-FTICR-MS is not recommended. Powdered samples can be used when wood chips of *P. santalinus* and *P. tinctorius* cannot be obtained.

Dramatic differences were observed in the DART-FTICR-derived fingerprints of the *P. santalinus* and *P. tinctorius* samples, and the high prediction accuracy (100%) demonstrated that DART-FTICR-MS can be used for rapid species-level identification of *P. santalinus* and *P. tinctorius*. Unique chemical profiles were obtained from the samples subjected to different treatment methods, illustrating the effects of the solvent type, the physical form of the sample and the drying treatment on the ions detected in the wood by DART-FTICR-MS.

Markers for the identification of *P. santalinus* and *P. tinctorius*

In addition to differentiating the two species based on model establishment and prediction, the most important advantage of OPLS-DA is that it can identify marker compounds for making this distinction. Variable influence on projection (VIP) reflects the importance of a variable in the model according to Y (in this study, Y is wood species) (Yang et al. 2006). The variable contributes substantially to Y when its VIP value is larger than 1 (Umetrics 2008). Variables with smaller covariance but high correlation to the score of the predictive OPLS component in the S-plot

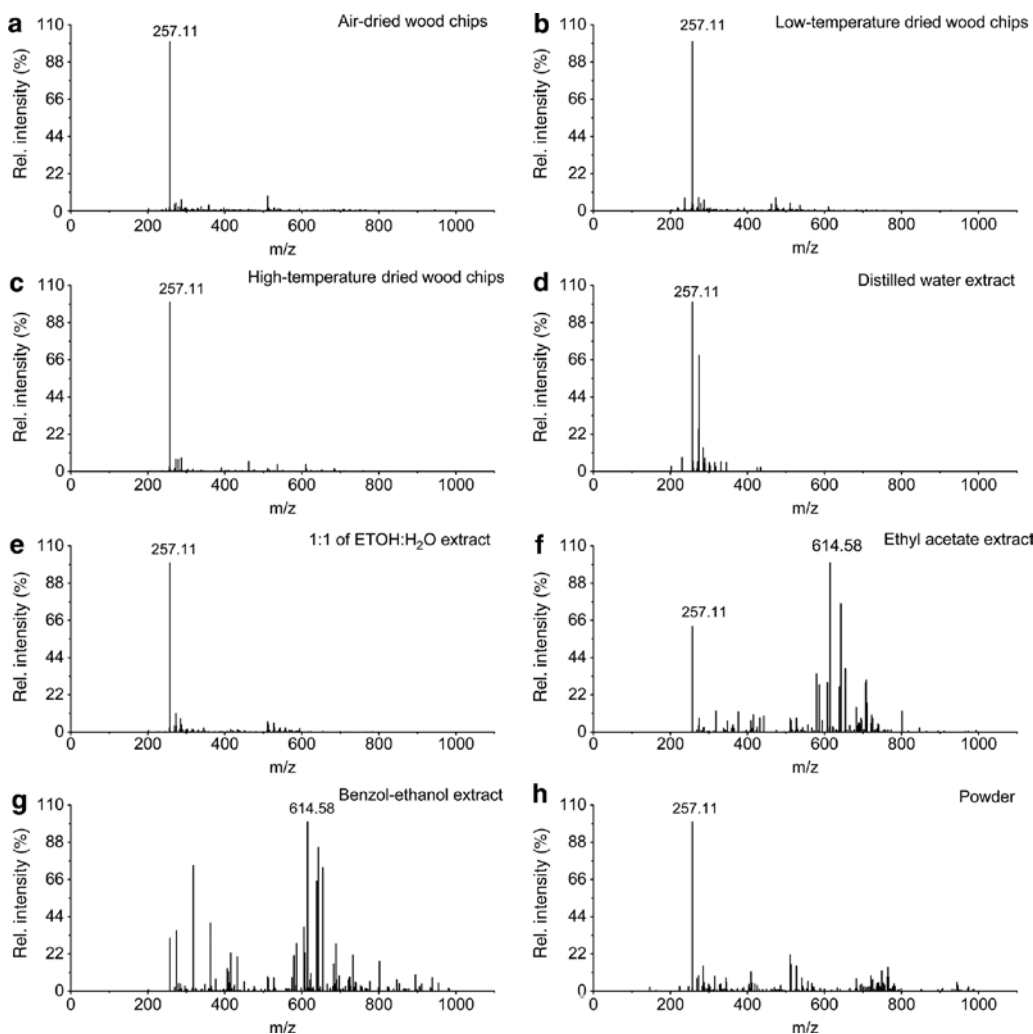


Figure 4: Representative DART-FTICR mass spectra of *P. tinctorius*.

(a) Air-dried (AD) wood chips. (b) Low-temperature dried (LD) wood chips. (c) High-temperature dried (HD) wood chips. (d) Distilled water extract. (e) 1:1 EtOH:H₂O extract. (f) Ethyl acetate extract. (g) Benzene-ethanol extract. (h) Powdered samples.

can be considered candidates for biomarkers (Rajalahti et al. 2009). Additionally, as another evaluation method, the t-test was performed to identify variables that significantly differentiate the two sets of data ($P < 0.05$). After comprehensively considering the S-plot in Figure S2, variables with VIP values larger than 1.6 (Figure S3) and P values less than 0.05 were selected as potential markers for the differentiation of *P. santalinus* and *P. tinctorius*. Based on these principles, five variables were considered as markers in the case of the AD chips (Table S2). Each variable, 477.39, 254.21, 473.36, 272.22 and 258.24 m/z, presented high VIP values and significant differences in intensity between the species.

To intuitively verify the capacity of these five potential markers to classify *P. santalinus* and *P. tinctorius*, mass data from the five potential markers in all the samples were

used to prepare a heatmap and conduct HCA based on ward.D (Figure S4). The HCA dendrogram shows that all the samples can be divided into two groups. All the samples from the training set and test set were classified correctly. When a new OPLS-DA model was built based on the mass data from the training set samples using the five potential markers, R^2X (0.981) was substantially improved and the prediction accuracy was also 100% (Table 1). These results suggested that these five variables, 477.39, 254.21, 473.36, 272.22 and 258.24 m/z, can be used as markers to classify AD chips of samples from *P. santalinus* and *P. tinctorius*. The high prediction accuracy obtained with the selected markers makes it possible to establish a model and prediction method that does not involve all of the variables, and this can significantly reduce time necessary for data analysis when evaluating a large number of samples.

To determine if the selected markers were applicable to other samples that had been subjected to different treatment procedures, the VIP and P values in the models derived from the seven treatments, namely, LD chips, HD chips, DWE, EH, EA, BE and powder, were calculated, and the results are listed in Table S3. The five marker ions were also detected in the LD and HD chips of the samples, and these markers can also be used as markers for LD chips due to the high VIP values ($VIP > 1.6$) and significant differences ($P < 0.05$). However, these five marker ions did not contribute substantially to the classification of the HD chips. Additionally, only the 473.36 m/z ion had an important role in distinguishing the wood species of extracts and powdered samples (with the exception of BE); the peaks at 477.39, 254.21, 272.22 and 258.24 m/z were poor markers, and they were not even detected in the EA and BE models. Significant differences in the relative peak intensities of the five marker ions were observed in the AD

chips of *P. santalinus* and *P. tinctorius* (Figure 5). Similar results can be seen in the case of the LD chips (Figure S5). Small differences were observed in the relative peak intensities in the other six models, and substantial deviations were observed in the HD chips, extracts and powdered samples (Figures S6–S11). These results demonstrated that the mass signals at 477.39, 254.21, 473.36, 272.22 and 258.24 m/z were suitable markers for wood identification using AD and LD chips. In addition, intraspecies variations were observed in the plot of relative peak intensity based on DART-FTICR-MS, which can be explained by the fact that the xylarium used in this work may have been cultivated in different regions with different growth conditions (Lesiak et al. 2015), and the process of milling and solvent extraction may further amplify the intraspecies differences. However, intraspecies variations did not affect the classification results by DART-FTICR-MS coupled with OPLS-DA.

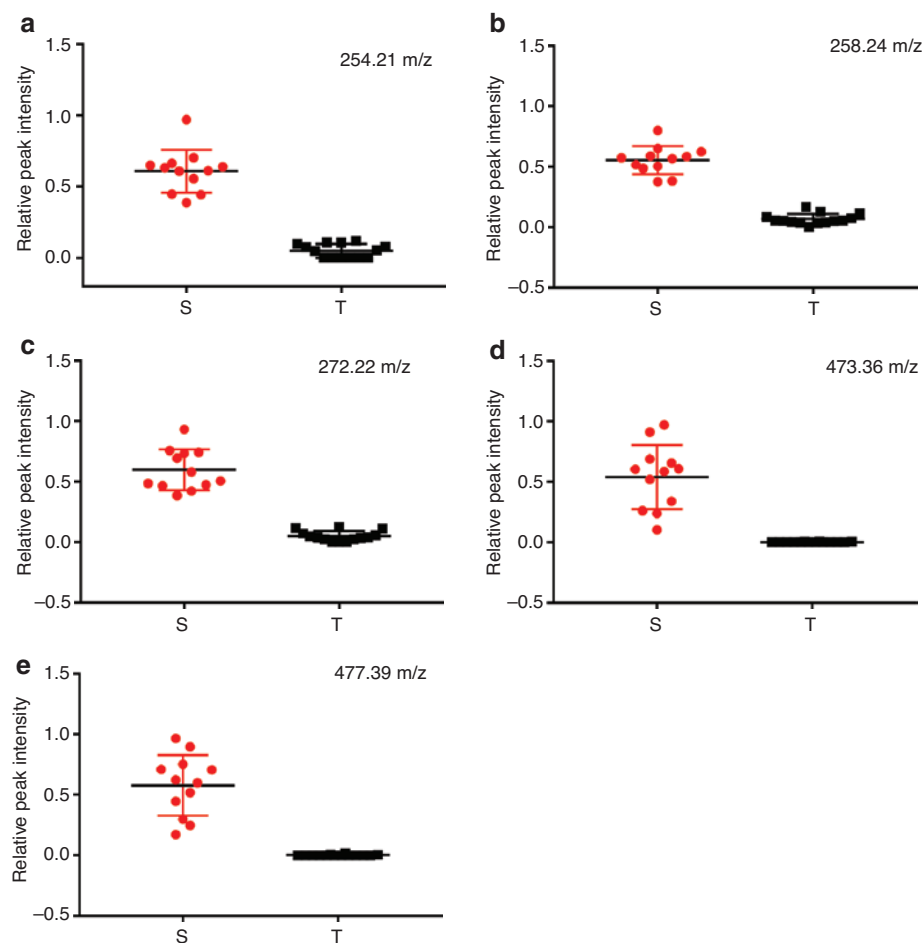


Figure 5: Relative peak intensities of five marker ions in the air-dried (AD) wood chips. (a) 254.21 m/z. (b) 258.24 m/z. (c) 272.22 m/z. (d) 473.36 m/z. (e) 477.39 m/z. (S) *P. santalinus*. (T) *P. tinctorius*.

Conclusions

The results shown in this study clearly demonstrate that DART-FTICR-MS coupled with multivariate statistical analysis is suitable for the rapid species-level identification of *P. santalinus* and *P. tinctorius* wood. Compared with the type of solvent and the physical form of the sample, the drying treatment had a greater effect on the chemical fingerprints obtained from DART-FTICR-MS. Air-dried wood chips are recommended as the optimal samples in the application of this direct analysis method. DART-FTICR-MS based on metabolomics is an effective method for identifying marker compounds for differentiating *P. santalinus* and *P. tinctorius* wood. Xylarium samples were used in this work, and they provided reliable information about wood species. The difficulties associated with collection of reliable (vouchered or validated) sample resulted in a limited number of samples of each species. The samples after drying treatments were not milled and extracted with solvent, which is a major drawback of this study. We hope that this can be done in the future when enough samples are available or using other species that are easier to obtain as an alternative.

Acknowledgments: We would like to express our gratitude to Professor IN-Gyu Choi of Seoul National University for his valuable suggestions. We recognize Professor Xinxiang Zhang, Professor Jiang Zhou and Dr. Wen Zhou of the Peking University for their selfless and constant assistance in using the DART-FTICR-MS instrument and its related software. The authors also wish to acknowledge Dr. Alex C. Wiedenhoeft of the Forest Product Laboratory, US, Professor Xiaomei Jiang, Dr. Bo Liu, Dr. Lichao Jiao and Mr. Tuo He of the Chinese Academy of Forestry and Mr. Di Xiao and Mr. Changyu Xu for their help in discussing the experiment and collecting the samples.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: This work was financially supported by the National High-level Talent for Special Support Program of China (the National “Ten Thousand” Talents Plan) (Grant No. W02020331) and the Fundamental Research Funds of the Chinese Academy of Forestry, Funder Id: <http://dx.doi.org/10.13039/501100004543> (Grant No. CAFYBB2017ZE003).

Employment or leadership: None declared.

Honorarium: None declared.

References

- Arunakumara, K.K.I.U., Walpola, B.C., Subasinghe, S., Yoon, M.-H. (2011) *Pterocarpus santalinus* Linn. f. (Rath handun): a review of its botany, uses, phytochemistry and pharmacology. *J. Korean Soc. Appl. Biol. Chem.* 54:495–500.
- Brack, D. (2003) Illegal logging and the illegal trade in forest and timber products. *Int. Forest. Rev.* 5:195–198.
- Cajka, T., Danhelova, H., Vavrecka, A., Riddellova, K., Kocourek, V., Vacha, F., Hajslova, J. (2013) Evaluation of direct analysis in real time ionization-mass spectrometry (DART-MS) in fish metabolomics aimed to assess the response to dietary supplementation. *Talanta* 115:263–270.
- Chong, J., Soufan, O., Li, C., Caraus, I., Li, S., Bourque, G., Wishart, D.S., Xia, J. (2018) MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* 46:W486–W494.
- Cody, R.B., Laramée, J.A., Durst, H.D. (2005) Versatile new ion source for the analysis of materials in open air under ambient conditions. *Anal. Chem.* 77:2297–2302.
- Cody, R.B., Dane, A.J., Dawson-Andoh, B., Adedipe, E.O., Nkansah, K. (2012) Rapid classification of White Oak (*Quercus alba*) and Northern Red Oak (*Quercus rubra*) by using pyrolysis direct analysis in real time (DART™) and time-of-flight mass spectrometry. *J. Anal. Appl. Pyrol.* 95:134–137.
- De Combarieu, E., Martinelli, E.M., Pace, R., Sardone, N. (2015) Metabolomics study of Saw palmetto extracts based on ¹H NMR spectroscopy. *Fitoterapia* 102:56–60.
- Dormontt, E.E., Boner, M., Braun, B., Breulmann, G., Degen, B., Espinoza, E., Gardner, S., Guillery, P., Hermanson, J.C., Koch, G. (2015) Forensic timber identification: it’s time to integrate disciplines to combat illegal logging. *Biol. Conserv.* 191: 790–798.
- Espinoza, E.O., Lancaster, C.A., Kreitals, N.M., Hata, M., Cody, R.B., Blanchette, R.A. (2014) Distinguishing wild from cultivated agarwood (*Aquilaria* spp.) using direct analysis in real time and time of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 28:281–289.
- Evans, P.D., Mundo, I.A., Wiemann, M.C., Chavarria, G.D., McClure, P.J., Voin, D., Espinoza, E.O. (2017) Identification of selected CITES-protected Araucariaceae using DART TOFMS. *IAWA J.* 38:266–S3.
- Farré, M., Picó, Y., Barceló, D. (2013) Direct peel monitoring of xenobiotics in fruit by direct analysis in real time coupled to a linear quadrupole ion trap-orbitrap mass spectrometer. *Anal. Chem.* 85:2638–2644.
- Finch, K., Espinoza, E., Jones, F.A., Cronn, R. (2017) Source identification of western Oregon Douglas-fir wood cores using mass spectrometry and random forest classification. *Appl. Plant Sci.* 5:1600158.
- Frédérich, M., Jansen, C., De Tullio, P., Tits, M., Demoulin, V., Angenot, L. (2010) Metabolomic analysis of Echinacea spp. by ¹H nuclear magnetic resonance spectrometry and multivariate data analysis technique. *Phytochem. Anal.* 21:61–65.
- Gao, X., Xie, M., Liu, S., Guo, X., Chen, X., Zhong, Z., Wang, L., Zhang, W. (2014) Chromatographic fingerprint analysis of metabolites in natural and artificial agarwood using gas chromatography-mass spectrometry combined with chemometric methods. *J. Chromatogr. B.* 967:264–273.

- Gasson, P. (2011) How precise can wood identification be? Wood anatomy's role in support of the legal timber trade, especially CITES. *IAWA J.* 32:137–154.
- Giffen, J.E., Lesiak, A.D., Dane, A.J., Cody, R.B., Musah, R.A. (2017) Rapid species-level identification of salvias by chemometric processing of ambient ionisation mass spectrometry-derived chemical profiles. *Phytochem. Anal.* 28:16–26.
- Gómez-Ríos, G.A., Gionfriddo, E., Poole, J., Pawliszyn, J. (2017) Ultrafast screening and quantitation of pesticides in food and environmental matrices by solid-phase microextraction-transmission mode (SPME-TM) and direct analysis in real time (DART). *Anal. Chem.* 89:7240–7248.
- Gowda, G.A., Djukovic, D. (2013) Overview of mass spectrometry-based metabolomics: opportunities and challenges. *Methods Mol. Biol.* 1198:3–12.
- Harris, G.A., Kwasnik, M., Fernández, F.M. (2011) Direct analysis in real time coupled to multiplexed drift tube ion mobility spectrometry for detecting toxic chemicals. *Anal. Chem.* 83:1908–1915.
- Hartvig, I., Czako, M., Kjær, E.D., Nielsen, L.R., Theilade, I. (2015) The use of DNA barcoding in identification and conservation of rosewood (*Dalbergia* spp.). *PLoS One* 10:e0138231.
- Hur, M., Yeo, I., Park, E., Kim, Y.H., Yoo, J., Kim, E., No, M.H., Koh, J., Kim, S. (2009) Combination of statistical methods and Fourier transform ion cyclotron resonance mass spectrometry for more comprehensive, molecular-level interpretations of petroleum samples. *Anal. Chem.* 82:211–218.
- IUCN. The IUCN Red List of Threatened Species. Version 2019-1. 2019.
- Jiao, L., Yin, Y., Cheng, Y., Jiang, X. (2014) DNA barcoding for identification of the endangered species *Aquilaria sinensis*: comparison of data from heated or aged wood samples. *Holz.* 68:487–494.
- Jiao, L., Yu, M., Wiedenhoef, A.C., He, T., Li, J., Liu, B., Jiang, X., Yin, Y. (2018) DNA barcode authentication and library development for the wood of six commercial *Pterocarpus* species: the critical role of *Xylarium* specimens. *Sci. Rep.* 8:1945.
- Kim, Y.M., Heyman, H.M. (2018) Mass spectrometry-based metabolomics. *Methods Mol. Biol.* 1775:107–118.
- Kim, H.J., Baek, W.S., Jang, Y.P. (2011) Identification of ambiguous cubeb fruit by DART-MS-based fingerprinting combined with principal component analysis. *Food Chem.* 129:1305–1310.
- Kim, H.J., Seo, Y.T., Park, S.I., Jeong, S.H., Kim, M.K., Jang, Y.P. (2015) DART-TOF-MS based metabolomics study for the discrimination analysis of geographical origin of *Angelica gigas* roots collected from Korea and China. *Metabolomics* 11:64–70.
- Lancaster, C., Espinoza, E. (2012a) Analysis of select *Dalbergia* and trade timber using direct analysis in real time and time-of-flight mass spectrometry for CITES enforcement. *Rapid Commun. Mass Spectrom.* 26:1147–1156.
- Lancaster, C., Espinoza, E. (2012b) Evaluating agarwood products for 2-(2-phenylethyl) chromones using direct analysis in real time time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 26:2649–2656.
- Laramee, J.A., Dupont, D.H., Connell, T.R., Nilles, J.M. (2008) Detection of chemical warfare agents on surfaces relevant to homeland security by direct analysis in real-time spectrometry. *Am. Lab.* 40:16–20.
- Lesiak, A.D., Cody, R.B., Dane, A.J., Musah, R.A. (2015) Plant seed species identification from chemical fingerprints: a high-throughput application of direct analysis in real time mass spectrometry. *Anal. Chem.* 87:8748–8757.
- Li, Z.Y., Zhang, Z.Z., Du, G.H., Qin, X.M. (2015) Comparative analysis of Danggui and European Danggui using nuclear magnetic resonance-based metabolic fingerprinting. *J. Pharm. Biomed. Anal.* 103:44–51.
- Ma, H., Jiang, Q., Dai, D., Li, H., Bi, W., Da Yong Chen, D. (2018) Direct analysis in real time mass spectrometry for characterization of large saccharides. *Anal. Chem.* 90:3628–3636.
- Mabberley, D.J. *Mabberley's Plant-Book: A Portable Classification of Plants, their Classification and Uses.* Cambridge University Press, Oxford, UK, 2017.
- Moco, S., Vervoort, J., Bino, R.J., De Vos, R.C., Bino, R. (2007) Metabolomics technologies and metabolite identification. *Trends Anal. Chem.* 26:855–866.
- Newman, D.J., Cragg, G.M. (2012) Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J. Nat. Prod.* 75:311–335.
- Rajalahti, T., Arneberg, R., Berven, F.S., Myhr, K.M., Ulvik, R.J., Kvalheim, O.M. (2009) Biomarker discovery in mass spectral profiles by means of selectivity ratio plot. *Chemom. Intell. Lab. Syst.* 95:35–48.
- Scotland, N., Ludwig, S. *Deforestation, the Timber Trade and Illegal Logging.* In EC Workshop on Forest Law, Enforcement, Governance and Trade, Brussels, 2002.
- Seshadri, T. (1972) Polyphenols of *Pterocarpus* and *Dalbergia* woods. *Phytochemistry* 11:881–898.
- Tacconi, L. *Illegal Logging: "Law Enforcement, Livelihoods and the Timber Trade"*. Earthscan, Routledge, 2012.
- Umetrics, A. *User Guide to SIMCA-P+ 12.* Umetrics Inc., Kinnelon, 2008.
- Wang, Y., Tang, H., Nicholson, J.K., Hylands, P.J., Sampson, J., Whitcombe, I., Stewart, C.G., Caiger, S., Oru, I., Holmes, E. (2004) Metabolomic strategy for the classification and quality control of phytomedicine: a case study of chamomile flower (*Matricaria recutita* L.). *Planta Med.* 70:250–255.
- Wang, J., Zhou, Y., Wang, M., Bi, W., Li, H., Chen, D.D.Y. (2018) High-throughput analysis for artemisinins with deep eutectic solvents mechanochemical extraction and direct analysis in real time mass spectrometry. *Anal. Chem.* 90:3109–3117.
- Yang, J., Zhao, X., Liu, X., Wang, C., Gao, P., Wang, J., Li, L., Gu, J., Yang, S., Xu, G. (2006) High performance liquid chromatography-mass spectrometry for metabolomics: potential biomarkers for acute deterioration of liver function in chronic hepatitis B. *J. Proteome Res.* 5:554–561.
- Yu, M., Jiao, L., Guo, J., Wiedenhoef, A.C., He, T., Jiang, X., Yin, Y. (2017) DNA barcoding of vouchered *xylarium* wood specimens of nine endangered *Dalbergia* species. *Planta* 246:1165–1176.
- Zhou, S., Forbes, M.W., Abbatt, J.P. (2015) Application of direct analysis in real time-mass spectrometry (DART-MS) to the study of gas-surface heterogeneous reactions: focus on ozone and PAHs. *Anal. Chem.* 87:4733–4740.

Supplementary Material: The online version of this article offers supplementary material (<https://doi.org/10.1515/hf-2018-0304>).