Research Article

Rapid Communications in Mass Spectrometry

Received: 17 September 2013 Revised: 29 October 2013 Accepted: 31 October 2013 Published online in Wiley Online Library

Rapid Commun. Mass Spectrom. **2014**, *28*, 281–289
(wileyonlinelibrary.com) DOI: 10.1002/rcm.6779

Distinguishing wild from cultivated agarwood (*Aquilaria* spp.) using direct analysis in real time and time of-flight mass spectrometry

Edgard O. Espinoza[1]*, Cady A. Lancaster[1], Natasha M. Kreitals[2], Masataka Hata[3], Robert B. Cody[4] and Robert A. Blanchette[5]

[1]U.S. National Fish and Wildlife Forensics Laboratory, 1490 East Main Street, Ashland, OR 97520-1310, USA
[2]Centre for Forensic Science, University of Western Australia, Perth, Western Australia, Australia
[3]Shoyeido Incense Company, Karasuma Nijo Nakagyo-ku, Kyoto, Japan
[4]JEOL USA, Inc., 11 Dearborn Rd, Peabody, MA 01960, USA
[5]Department of Plant Pathology, University of Minnesota, St. Paul, MN, USA

RATIONALE: It is important for the enforcement of the CITES treaty to determine whether agarwood (a resinous wood produced in *Aquilaria* and *Gyrinops* species) seen in trade is from a plantation that was cultivated for sustainable production or was harvested from natural forests which is usually done illegally.
METHODS: We analyzed wood directly using Direct Analysis in Real Time (DART™) ionization coupled with Time-of-Flight Mass Spectrometry (TOFMS). Agarwood was obtained from five countries, and the collection contained over 150 samples. The spectra contained ions from agarwood-specific 5,6,7,8-tetrahydro-2-(2-phenylethyl)chromones as well as many other ions. The data was analyzed using either kernel discriminant analysis or kernel principal component analysis. Probability estimates of origin (wild vs cultivated) were assigned to unknown agarwood samples.
RESULTS: Analysis of the DART-TOFMS data shows that many of the chromones found in cultivated and wild agarwood samples are similar; however, there is a significant difference in particular chromones that can be used for differentiation. In certain instances, the analysis of these chromones also allows inferences to be made as to the country of origin. Mass Mountaineer™ software provides an estimate of the accuracy of the discriminate model, and an unknown sample can be classified as cultivated or wild. Eleven of the thirteen validation samples (85%) were correctly assigned to either cultivated or wild harvested for their respective geographic provenance. The accuracy of each classification can be estimated by probabilities based on Z scores.
CONCLUSIONS: The direct analysis of wood for the diagnostic chromones using DART-TOFMS followed by discriminant analysis is sufficiently robust to differentiate wild from cultivated agarwood and provides strong inference for the origin of the agarwood. Copyright © 2013 John Wiley & Sons, Ltd.

Agarwood is one of many names given to the dark resinous and aromatic wood harvested from species of *Aquilaria* and *Gyrinops* trees found throughout southeast Asia and the Malay Archipelago (Fig. 1).[1–3] It is widely accepted that agarwood forms as a response by the tree to various forms of injury and/or microbial attack and it is found very rarely in trees growing in natural forests.[2,4–6] Over the past several thousand years, the demand for agarwood has resulted in exploitation of these trees and overharvesting. Today, very few old growth *Aquilaria* trees exist in southeast Asian countries. Many of the trees that remain are located in national parks or other protected areas. In recent years, poaching and illegal harvesting have threatened these remaining trees. Recently, methods have been developed to induce agarwood production in young plantation trees[7] and products are now available that have been produced sustainably by cultivation of the agarwood. Due to the exploitation of this tree and because of the high value and commercial demand for agarwood, all *Aquilaria* and *Gyrinops* were listed on Appendix II of the Convention on the International Trade in Endangered Species (CITES) in 2004.[8] The International Union for Conservation of Nature (IUCN) Red List of Threatened Species lists one species as critically endangered (*A. crassna*), another as threatened (*A. rostrata*), and seven others as vulnerable (*A. banaensae*, *A. beccariana*, *A. cimingiana*, *A. hirta*, *A. malaccensis*, *A. microcarpa*, and *A. sinensis*).[9]

Agarwood trade occurs in the form of wood chips, incense sticks, resin balls, sawdust, perfumes, oils, and traditional medicines, as well as many other products. Since all *Aquilaria* species are listed in CITES Appendix II, all products in international trade, even if they contain small amounts of agarwood, must comply with CITES permit requirements, including essential oils (referred to as agarwood oil or oud oil) and traditional medicines.

* *Correspondence to:* E. O. Espinoza, U.S. National Fish and Wildlife Forensics Laboratory, 1490 East Main Street, Ashland, OR 97520-1310, USA.
E-mail: ed_espinoza@fws.gov

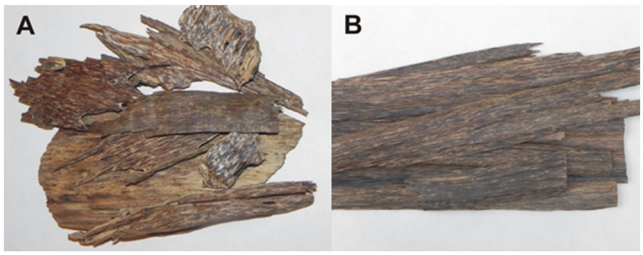

Figure 1. Examples of agarwood, a resin-filled wood that has been cut out of Aquilaria trees, from wild trees in natural forests (A) and cultivated trees grown in plantations (B).

Agarwood contains a complex mixture of many volatile aromatic compounds including agarofurans, cadinanes, eudesmanes, valencanes and eremophilanes, guaianes, prezizanes, vetispiranes, 2-(2-phenylethyl)chromones, tetrahydro-2-(2-phenylethyl)chromones, and many others.[2] Among 145 compounds, Naef described 16 highly oxidized 5,6,7,8-tetrahydro-2-(2-phenylethyl)chromones as compounds unique to agarwood.[2] In addition to these compounds, 23 other 2-(2-phenylethyl)chromones have been identified in differing qualities of agarwood products. It is claimed that these chromones are not found in the sound white wood of *Aquilaria* and *Gyrinops* and only occur in the resin-filled reaction wood that is unique to agarwood.[2,5] Lancaster *et al*. reported the use of the highly oxidized 5,6,7,8-tetrahydro-2-(2-phenylethyl)-chromones for agarwood identification and for verifying compliance of commercial products imported into the United States.[10] Differences in the aroma of agarwood, which occur when agarwood is grown in different regions of the world, are apparently due to varying types of compounds and the concentrations of the complex assortment of these compounds in the agarwood.

This paper presents research which facilitates legal agarwood trade and protects *Aquilaria* diversity that is in danger of extinction within natural forests. In this report, we show the potential of using Direct Analysis in Real Time Time-of-Flight Mass Spectrometry (DART-TOFMS) and statistical analysis to differentiate sustainably produced agarwood that has been cultivated in plantations from agarwood obtained from natural forests. CITES has traditionally encouraged the trade of cultivated exports if the cultivated type can be distinguished from the wild counterparts. The availability of cultivated agarwood eliminates the need for harvesting *Aquilaria* and *Gyrinops* from natural old growth forests and provides a new economy to people living in rural areas of Asia.

The samples analyzed in this research consisted of woodchips collected for research investigations by one of the authors (R.A.B.) and are considered *Aquilaria* voucher samples. All wood samples were analyzed directly using DART-TOFMS. DART-TOFMS uses an ambient atmospheric ionization source that provides rapid analysis and requires minimal sample preparation. Cody *et al*. have described the principal ionization mechanisms for DART-TOFMS.[11] DART-TOFMS has proven to be a useful tool for differentiating many different types of wood and has previously been used for the identification of *Dalbergia*,[12] *Aquilaria*,[10] *Quercus*[13] and *Eucalyptus*.[14] The objectives of our study were to determine if DART-TOFMS coupled with statistical analysis can be used to differentiate cultivated from wild agarwood of *Aquilaria* species (Fig. 1).

EXPERIMENTAL

Material

Agarwood from five countries (see Table 1) was analyzed. This reference set is composed of plantation-cultivated agarwood (*Aquilaria crassna*) from Thailand and Vietnam, plantation-cultivated agarwood (*Aquilaria beccariana*) from Malaysia, plantation-cultivated agarwood (*Aquilaria sinensis*) from Hainan Island, China, and wild-type agarwood from five countries. Because of the difficulty of obtaining field samples of agarwood, samples were obtained from merchants and the agarwood is presumed to have been harvested from the country indicated by the merchant. These merchant-obtained samples are denoted in Table 1 with an asterisk. In total, the analysis consisted of 78 samples from agarwood plantations and 74 samples of wild harvested agarwood (see Table 1). For Bornean samples, the species source of the wild agarwood is unknown. Reference samples of known origin that were used for validation purposes were obtained from the Shoyeido Incense Company (Kyoto, Japan) (see Table 2). Validation samples were not used for creating the classification models. All wild agarwood samples were obtained before CITES permits were required or had CITES permits issued by the exporting country.

Table 1. Agarwood reference samples

Geographic landscape	Country (Region)	Species	Method of growth	N
Borneo	Malaysia (Sabah)	*Aquilaria beccariana*	Cultivated	20
Borneo	Indonesia (Kalamantan)*	unknown	Wild Type	13
Hainan Island	China	*Aquilaria sinensis*	Cultivated	13
Hainan Island	China	*Aquilaria sinensis*	Wild	18
Thailand	Thailand	*Aquilaria crassna*	Cultivated	19
Thailand*	Thailand*	*Aquilaria crassna*	Wild Type	13
Vietnam	Vietnam	*Aquilaria crassna*	Cultivated	26
Vietnam*	Vietnam*	*Aquilaria crassna*	Wild Type	30

Rapid Communications in Mass Spectrometry

Table 2. Validation samples

Geographic landscape	Code	Country (Region)	Method of growth	Circa	Assignment
Vietnam	H2503VE01	Midland Viet Nam	Wild	1995	Wild (99.86%)
Vietnam	H2503VE02	Midland Viet Nam	Wild	1974	Wild (99.62%)
Vietnam	H2503VE03	Midland Viet Nam	Wild	1971	*"Cultivated"*
Vietnam	H2503VE04	Northern Viet Nam	Cultivated	2011	Cultivated (100%)
Vietnam	H2503VE05	Tinh Ha Tinh, Northern Viet Nam	Cultivated	2009	Cultivated (99.99%)
Vietnam	H2503VE06	Da Nang, Midland Viet Nam	Wild	2009	Wild (99.83%)
Vietnam	H2503VE07	Nha Trang. Southern Viet Nam	Wild	1970	Wild (97.52%)
Vietnam	H2503VE08	Nha Trang. Southern Viet Nam	Wild	1970	Wild (96.42%)
Borneo	H2503IN01	Midland Kalimantan, Indonesia	Wild	2008	Wild (91.97%)
Borneo	H2503AG02	Sandakan, Sabah, Malaysia	Wild	1970	Wild (97.70%)
Hainan Island	H2503AG05	Hainan Island , China	Wild	1960	Wild (98.98%)
Hainan Island	H2503AG06	Hainan Island, China	Cultivated	2011	Cultivated (99.99%)
Thailand	H2503AG03	Thailand	Wild	2011	*"Cultivated"*

Methods

All the samples were analyzed directly by DART-TOFMS by holding a wood sliver in the gas stream with no further sample preparation.[10,12] A mass calibration standard of poly(ethylene glycol) 600 (Ultra, Kingstown, RI, USA) was run between samples. Mass spectra were acquired using a DART-SVP ion source (IonSense, Saugus, MA, USA) coupled to a JEOL AccuTOF time-of-flight mass spectrometer (JEOL USA, Peabody, MA, USA) in positive ion mode. The DART source parameters were: electrode 1 voltage, 150 V; electrode 2 voltage, 250 V; and gas heater temperature, 450 °C. The mass spectrometer settings included: ring lens voltage, 5 V; orifice 1 voltage, 20 V; orifice 2 voltage, 5 V; cone temperature, 120 °C; peak voltage, 600 V; bias voltage, 28 V; focus voltage, –120 V; reflectron voltage, 870 V; pusher voltage, 778 V; pulling voltage, –778 V; suppression voltage, 0.00 V; flight tube voltage, –7000 V; and detector voltage, 2500 V. Spectra were obtained over the mass range of *m/z* 50 to 1100 at 1 scan per second. The helium flow rate for the DART source was 2.0 mL/s. The resolving power of the mass spectrometer, as stated by the manufacturer, was 2.0 mDa.

Diagnostic 2-(2-phenylethyl)chromones were used to construct a search library.[2,10] Due to isomeric configurations of the 39 chromones reported, there are 19 unique masses. Because of the complex naming convention for the chromones, we have used the naming system reported by Naef,[2] in which each 2-(2-phenylethyl)chromone is assigned a number (see Table 3). The 2-(2-phenylethyl)chromones are identified by their assigned numbers from 67 to 105. Further diagnostic ions were determined using a heat map (Fig. 2). A heat map is a graphical representation of every sample used in the statistical model, where each 'row' represents a different sample. The intensity (blot color) is a direct correlation to the relative intensity of that ion in the spectrum. The x axis shows the *m/z* values of all the ions detected.

In total 174 ions were used for the statistical analysis. For further discussion of the highly oxidized 2-(2-phenylethyl)-chromones, see Lancaster and Espinoza[10] and Naef.[2]

TSSPro3 (Shrader Analytical Labs, Detroit, MI, USA) data processing software was used to export the text files of the mass-calibrated, centroided mass spectra for elemental composition determination and further analysis. Statistical analysis of the datasets was conducted using the Mass Mountaineer software (RBC Software, Peabody, MA, USA). The classification capabilities of Mass Mountaineer were used to calculate the principal components of each data set. Kernel Discriminant Analysis (KDA) or Kernel Principal Component Analysis (KPCA) was performed using the 174 identified ions mentioned above. For all models, a tolerance of 5 mDa was applied. To assess model accuracy, leave-one-out cross-validation (LOOCV) was employed. The LOOCV is based on the distance from the cluster mean of each sample that is omitted. Essentially, each sample is successively omitted from the training set and placed as an unknown, thus subjecting each sample for comparison against the entire training set. The classification models were validated using the reference samples listed in Table 2. The Mass Mountaineer software assigns an estimate of probability to unclassified (e.g., unknown) spectra.

The unsupervised learning algorithms (PCA and KPCA) do not make use of the information about class membership; colors are assigned to points after the calculation. Therefore, these are good methods to determine whether there are real patterns in the data. The supervised learning algorithms (LDA and KLDA) use the class membership information to achieve the greatest separation between classes in the training set, and this information is then used to classify new data. KDA and KPCA are extensions of LDA and PCA that map features into a higher-dimensional space by using a typically nonlinear function.[15,16] This allows points that cannot be linearly separated in a two-dimensional space to be separated in higher dimensions.[17] Estimated probabilities are based on Z scores (distance divided by standard deviation) based on a normal distribution.

RESULTS

The reference collection for this study consisted of cultivated and wild samples from three different *Aquilaria* species obtained from four different geographic landscapes (Borneo, Hainan Island, Thailand, and Vietnam) (Table 1). This resulted in differences in the number and type of chromones or ions present in the sample. Kernel Discriminant Analysis (KDA)

Table 3. 2-(2-Phenylethyl)chromones reported in agarwood extracts.[2] Because of isomeric configurations, only 19 masses can be identified by DART-TOF-MS

Compound	Chromone type	Predicted (*m/z*)	Formula
67	2-(2-phenylethyl)chromone	251.1072	$C_{17}H_{15}O_2$ $[M+H]^+$
68	7-hydroxy-2-(2-phenylethyl)chromone	267.1021	$C_{17}H_{15}O_3$ [M+H]+
69	6-hydroxy-2-(2-phenylethyl)chromone		
70	6-methoxy-2-(2-phenylethyl)chromone	281.1178	$C_{18}H_{17}O_3$ $[M+H]^+$
71	2-[2-(4-methoxyphenyl)ethyl]chromone		
72	6-hydroxy-2-[2-(4′-hydoxyphenyl)ethyl]chromone	283.0970	$C_{17}H_{15}O_4$ $[M+H]^+$
73	6-hydroxy-2-[2-(2′-hydoxyphenyl)ethyl]chromone		
74	6,8-dihydroxy-2-(2-phenylethyl)chromone		
75	5,8-dihydroxy-2-(2-phenylethyl)chromone		
86	6-hydroxy-2-[(2*R*)-7′-hydroxy-2-phenylethyl]chromone		
103	oxidoagarochromone A		
76	6,7-dimethoxymethoxy-2-(2-phenylethyl)chromone	297.1127	$C_{18}H_{17}O_4$ $[M+H]^+$
77	6-hydroxy-2-(2-(4′-methoxyphenyl)ethyl)chromone		
78	6-hydroxy-7-methoxy-2-(2-phenylethyl)chromone		
79	6-methoxy-2-[2-(4′-methoxyphenyl)ethyl]chromone	311.1283	$C_{19}H_{19}O_4$ $[M+H]^+$
80	6-methoxy-2-[2-(3′-methoxyphenyl)ethyl]chromone		
81	6,7-dimethoxy-2-(2-phenylethyl)chromone		
82	5,8-dihydroxy-2-[2-(4′-methoxyphenyl)ethyl]chromone	313.1076	$C_{18}H_{17}O_5$ $[M+H]^+$
104	oxidoagarochromone B		
83	6-methoxy-2-[2-(3-methoxy-4-hydroxyphenyl)ethyl]chromone	327.1233	$C_{19}H_{19}O_5$ $[M+H]^+$
84	6,7-dimethoxy-2-[2-(4′-methoxyphenyl)ethyl]chromone	341.1389	$C_{20}H_{21}O_5$ $[M+H]^+$
85	7,8-dimethoxy-2-[2-(3′-acetoxyphenyl)ethyl]chromone	369.1338	$C_{21}H_{21}O_6$ $[M+H]^+$
87	5,6,7,8-tetrahydro-6β,7β-dihydroxy-2-(2-phenylethyl)chromone	287.1283	$C_{17}H_{19}O_4$ $[M+H]^+$
88	4*H*-1-benzopyran-4-one,5,6,7-tris(acetyloxy)-2-[2-[2-(acetyloxy)phenyl]ethyl]-5,6,7,8-tetrahydro-, [5S-(5α,6β,7α)]-	319.1182	$C_{17}H_{19}O_6$ $[M+H]^+$
91	agarotetrol		
92	isoagarotetrol		
93	5,6,7,8-tetrahydro-5β,6β,7α,8β-tetrahydroxy-2-(2-phenylethyl)chromone		
89	(5*S**,6*R**,7*S**)-5,6,7-trihydroxy-2-(3-hydroxy-4-methoxyphenethyl)-5,6,7,8-tetrahydro-4*H*-chromen-4-one	349.1287	$C_{18}H_{21}O_7$ $[M+H]^+$
90	(5*S**,6*R**,7*R**)-5,6,7-trihydroxy-2-(3-hydroxy-4-methoxyphenethyl)-5,6,7,8-tetrahydro-4*H*-chromen-4-one		
96	4*H*-1-benzopyran-4-one,5,6,7,8-tetrakis(acetyloxy)-5,6,7,8-tetrahydro-2-[2-(4-methoxyphenyl)ethyl]-, [5S-(5a,6b,7b,8a)]-		
97	4*H*-1-benzopyran-4-one,5,6,7,8-tetrahydro-5,6,7,8-tetrahydroxy-2-[2-(4-methoxyphenyl)ethyl]-, [5S-(5 α,6β,7α,8β)]-		
94	4*H*-1-benzopyran-4-one,5,6,7,8-tetrahydro-5,6,7,8-tetrahydroxy-2-[2-(2-hydroxyphenyl)ethyl]-,[5S-(5a,6β,7a,8β)]-	335.1131	$C_{17}H_{19}O_7$ $[M+H]^+$
95	5α,6β,7β,8α-tetrahydroxy-2-[2-(2-hydroxyphenyl)ethyl]5,6,7,8-tetrahydrochrome		
100A	4*H*-1-benzopyran-4-one,5,6,7,8-tetrahydro-5,6,7,8-tetrahydroxy-2-(2-hydroxy-2-phenylethyl)-, [5S-[2(S*),5α,6β,7α,8β]]-		
100B	4*H*-1-benzopyran-4-one,5,6,7,8-tetrahydro-5,6,7,8-tetrahydroxy-2-(2-hydroxy-2-phenylethyl)-, [5S-[2(R*),5α,6β,7α,8β]]-		
98	5,6,7,8-tetrahydroxy-2-(3-hydroxy-4-methoxyphenethyl)-5,6,7,8-tetrahydro-4*H*-chromen-4-one	365.1237	$C_{18}H_{21}O_8$ $[M+H]^+$
99	5α,6β,7β-trihydroxy-8α-methoxy-2-(2-phenylethyl)-5,6,7,8-tetrahydrochromone	333.1338	$C_{18}H_{21}O_6$ $[M+H]^+$
101	4*H*-1-benzopyran-4-one,8-chloro-5,6,7,8-tetrahydro-5,6,7-trihydroxy-2-(2-phenylethyl)-, (5R,6R,7R,8S)-rel-(+)-	337.0843	$C_{17}H_{18}ClO_5$ $[M+H]^+$
102	8-chloro-5,6,7-trihydroxy-2-(3-hydroxy-4-methoxyphenethyl)-5,6,7,8-tetrahydro-4*H*-chromen-4-one	382.0819	$C_{18}H_{19}ClO_7$ $[M+H]^+$
105	oxidoagarochromone C	329.1025	$C_{18}H_{17}O_6$ $[M+H]^+$

distinguished wild and cultivated samples from the same geographic region, in addition to separating samples based on geographic landscape of origin (see Fig. 3). The LOOCV of the KDA analysis was 88.82%, suggesting the need to investigate cultivated versus wild harvested differences for each region independently. With the exception of Vietnam, it was noted that the spectra of the cultivated samples show fewer ions than wild harvested types from the same region (see Fig. 2).

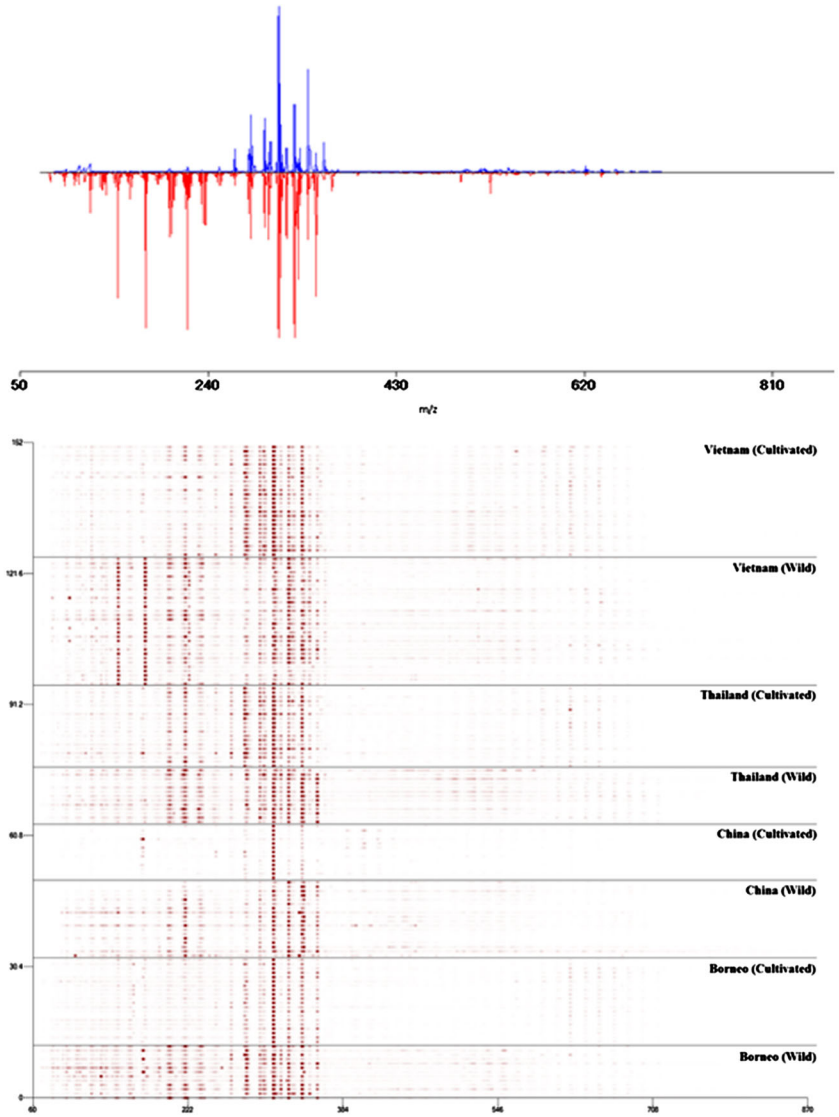

Figure 2. Top: DART-TOFMS spectrum of cultivated (blue) and wild (red) agarwood from Vietnam. Bottom: Heat map of ions present in cultivated and wild grown *Aquilaria* wood samples. The intensity of the heat map display correlates to the abundance of that ion in the spectrum.

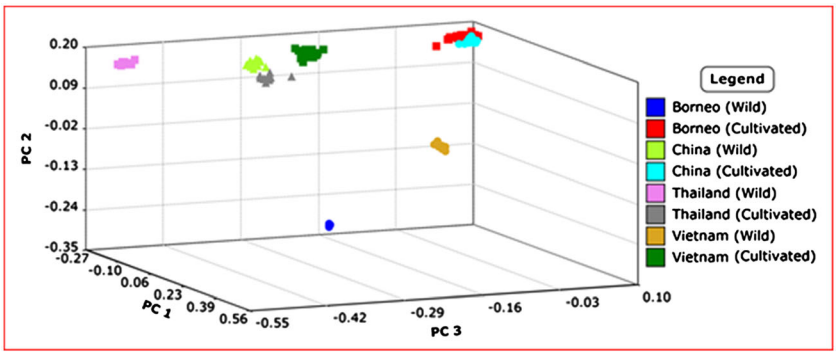

Figure 3. Kernel discriminant analysis of all ions detected in cultivated and wild grown *Aquilaria* wood samples (a 3D video of the model is available at http://www.fws.gov/lab/data-videos/agarwoodvideos/wild-vscultivated/figure3/).

Comparison of wild and cultivated *A. beccariana* from Borneo

The *A. beccariana* reference collection from Borneo consisted of 20 cultivated samples and 13 wild harvested samples (Table 1). Wild harvested samples were distinguishable from cultivated *A. beccariana* by an increase in the intensity of the ions present and in the greater number of detectable ions present (Fig. 2). The frequencies of *m/z* 85.029, 123.118, 143.084, 207.172, 313.111, 334.139 and 521.288 were more than 50% higher in wild than in cultivated samples and the frequencies of *m/z* 149.023, 178.060, 221.183, 282.121, 358.134 and 651.254 were more than 50% greater in cultivated samples. For ions that were present in *A. beccariana* at a frequency greater than 10%, *m/z* 133.101, 135.116, 203.184, 205.189, 281.109 and 333.140 were detected only in wild harvested samples, and *m/z* 165.091 and 667.248 were present only in cultivated samples.

Separation of cultivated and wild harvested *A. beccariana* was confirmed using statistical analysis. Figure 4 shows the KPCA for this analysis with a LOOCV of 100% achieved. The strength of the model was supported by the correct assignment of both validation samples to wild harvested with a probability of assignment of 91.97% and 97.70% (Table 2).

Comparison of wild and cultivated of *A. sinensis* from China

The *A. sinensis* reference collection from Hainan Island, China, consisted of 13 cultivated samples and 18 wild harvested samples (Table 1). Wild harvested samples were distinguishable from cultivated *A. sinensis* by an increase in the intensity of the ions present and in the greater number of detectable ions present (Fig. 2). The frequencies of *m/z* 95.085, 105.068, 107.085, 109.101, 123.118, 133.101, 135.116, 187.146, 189.162, 199.147, 207.172, 215.144, 253.180, 302.080, 313.111, 314.112, 330.107, 343.120, 344.121, 347.116, 350.133, 349.128, 365.123 and 521.288 were more than 50% higher in wild than in cultivated samples, and the frequencies of *m/z* 173.096, 191.105, 317.139, 334.139 and 622.258 were more than 50% greater in cultivated samples. For ions that were present in *A. sinensis* at a frequency greater than 10%, *m/z* 127.041, 165.091, 192.139, 205.189, 565.281, 583.298, 653.236, 667.248, 669.231, 688.226 and 713.259 were detected only in wild harvested samples, and *m/z* 333.140 was present only in cultivated samples.

Separation of cultivated and wild harvested *A. sinensis* was confirmed using statistical analysis. Figure 5 shows the KPCA for this analysis with a LOOCV of 100% achieved. The strength

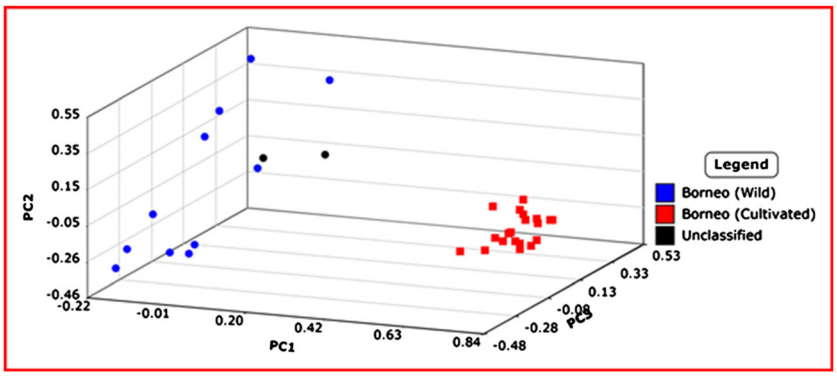

Figure 4. Kernel principal component analysis of ions detected in cultivated and wild grown *Aquilaria beccariana* from Borneo. Validation samples were correctly assigned to a wild origin (a 3D video of the model is available at http://www.fws.gov/lab/data-videos/agarwoodvideos/wild-vs-cultivated/figure4/).

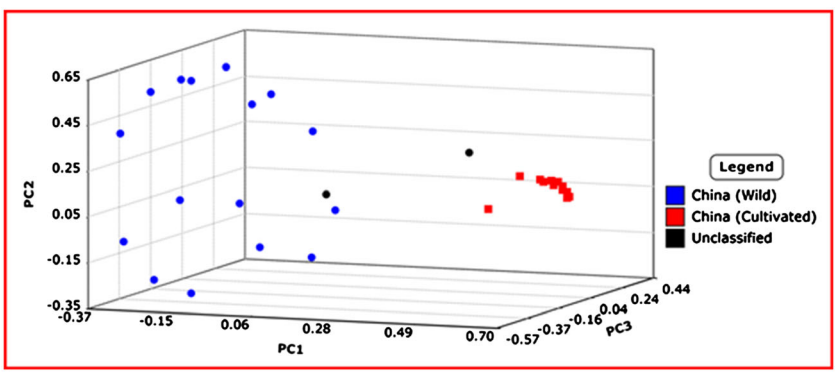

Figure 5. Kernel principal component analysis of ions detected in cultivated and wild grown *Aquilaria sinensis* from China. Validation samples correctly assigned sample H2503AG06 as cultivated and sample H2503AG05 as wild (a 3D video of the model is available at http://www.fws.gov/lab/data-videos/agarwoodvideos/wild-vscultivated/figure5/).

Rapid Communications in Mass Spectrometry

of the model was supported by the correct assignment of both validation samples with a probability of assignment of 99.99% for the cultivated sample and 98.99% for the wild harvest sample (Table 2).

Comparison of wild and cultivated *A. crassna* from Vietnam and Thailand

The *A. crassna* reference collection consisted of 26 cultivated and 30 wild harvested samples from Vietnam and 19 cultivated and 13 wild harvested samples from Thailand. Cultivated *A. crassna* from Vietnam exhibited unique ions not detected in the wild samples. For Thailand, wild harvested samples were distinguishable from cultivated *A. crassna* by an increase in the intensity and the greater number of detectable ions present (Fig. 2).

For Thailand, the frequencies of *m/z* 218.159, 221.183, 267.105, 299.096, 314.112, 344.122, 357.131 and 651.254 were more than 50% higher in wild than in cultivated samples, and the frequencies of *m/z* 312.135 and 317.139 were more than 50% greater in cultivated samples. For ions that were present in *A. crassna* at a frequency greater than 10%, *m/z* 287.120, 343.120 and 667.248 were detected only in wild harvested samples, and *m/z* 251.109, 252.111, 269.113 and 333.140 were present only in cultivated samples.

Distinguishing cultivated from wild harvested *A. crassna* from Thailand was demonstrated using statistical analysis. Figure 6 shows the KPCA for this analysis with a LOOCV of 93.75%. Validation analysis of the single sample from Thailand erroneously classified this sample as having a cultivated source, whereas the source indicated it had wild harvested origin (Table 2).

For Vietnam, the frequencies of *m/z* 149.086, 223.129 and 333.140 were more than 50% higher in wild than cultivated samples, and frequencies of *m/z* 85.029, 91.055, 95.085, 97.031, 105.068, 107.085, 115.038, 121.068, 127.041, 137.062, 143.084, 145.100, 161.128, 173.096, 175.112, 177.057, 187.146, 199.147, 201.164, 209.080, 213.124, 215.144, 217.158, 219.172, 220.177, 221.183, 231.136, 233.156, 237.183, 268.105, 283.097, 284.101, 285.114, 297.112, 298.117, 301.105, 302.110, 303.122,

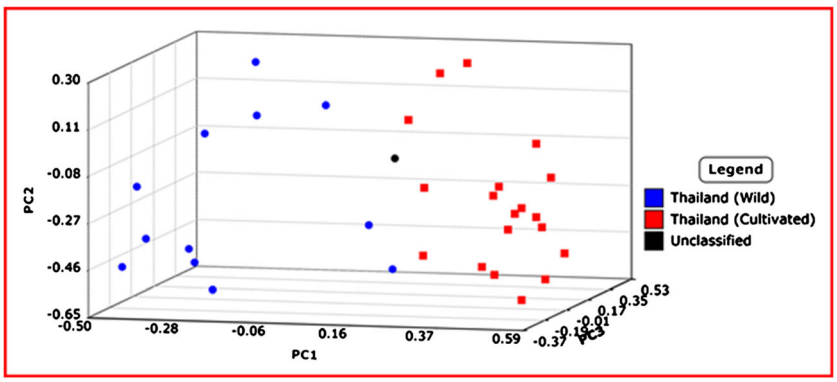

Figure 6. Kernel principal component analysis of ions detected in cultivated and wild grown *Aquilaria crassna* from Thailand. Validation sample incorrectly assigned sample H2503AG03 as cultivated (a 3D video of the model is available at http://www.fws.gov/lab/data-videos/agarwoodvideos/wild-vs-cultivated/figure6/).

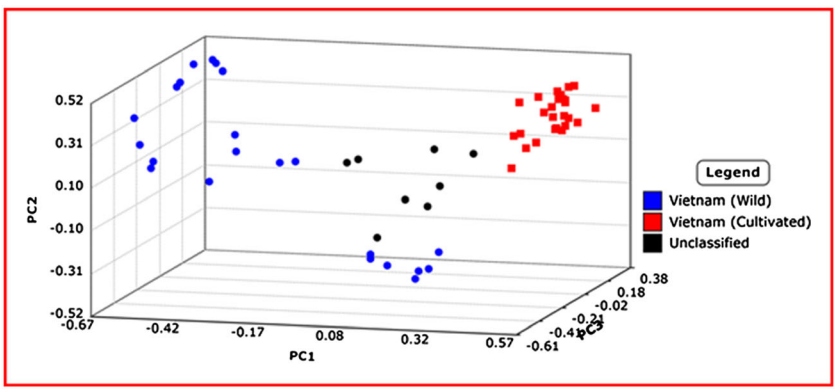

Figure 7. Kernel principal component analysis of ions detected in cultivated and wild grown *Aquilaria crassna* from Vietnam. Two validation samples were correctly classified as cultivated and five samples were correctly assigned as wild. Sample H2503VE03 was from a wild source and was misclassified (a 3D video of the model is available at http://www.fws.gov/lab/data-videos/agarwoodvideos/wild-vs-cultivated/figure7/).

309.113, 311.127, 312.135, 315.120, 317.139, 319.120, 320.123, 325.107, 327.122, 328.126, 331.119, 341.136, 342.141, 349.128 and 621.247 were more than 50% greater in cultivated samples. For ions that were present in *A. crassna* at a frequency greater than 10%, *m/z* 149.023, 177.100, 201.204, 203.215, 217.193, 219.208, 525.178, 565.281, 583.298 and 669.231 were detected only in wild harvested samples, and *m/z* 163.146, 251.109, 252.111, 269.113 and 281.109 were present only in cultivated samples.

Distinguishing cultivated from wild harvested *A. crassna* was demonstrated using statistical analysis. Figure 7 shows the KPCA for this analysis with a LOOCV of 100% achieved. Seven of the validation samples in the group were assigned to the respective wild or cultivated origin, with probabilities of assignment exceeding 96% for all samples (Table 2). KPCA also demonstrated two clear groups within the wild harvested samples. Interestingly, the wild harvested Midland Vietnam samples were assigned to one cluster, while the Southern Vietnam samples were assigned to a second cluster (Fig. 7).

Two samples in the group (H2503AG03 from Thailand; H2503VE03 from Vietnam) were incorrectly assigned. In both cases, these were wild samples but were classified as cultivated by the analysis. This misclassification may be because the spectra showed very few and low concentration of chromones. Empirically, we have observed that samples that exhibit little agarwood resin also reveal few of the diagnostic chromones. Therefore, we conclude that the misclassification is due to the very low amount of resin present and suggests that a sample with exceedingly low resin content will not be able to be differentiated into cultivated or wild.

DISCUSSION AND CONCLUSIONS

DART-TOFMS is a powerful tool that allows for rapid analysis of samples. When analyzing timber and wood products, DART-TOFMS spectra can be obtained with virtually no sample preparation. From a traditional mass spectrometry point of view, it is almost inconceivable to be able to analyze a solid material, like wood, without extracting the chemical compounds of interest. The DART-TOFMS system does this well.

The analysis of the highly oxidized 5,6,7,8-tetrahydro-2-(2-phenylethyl)chromones has been very useful in the identification of agarwood in commercial products consisting of mixtures.[10] Table 3 lists the chromones that have been reported in agarwood.[2] In this study, the analysis of 2-(2-phenylethyl)chromones was used to distinguish cultivated from wild harvested agarwood, and the results are promising. These chromones make up 41% of the components of agarwood.[18] However, not all ions detected are chromones. For the many ions detected whose identity is unknown, one can posit they belong to sesquiterpenes, a group which makes up 52% of the constituents of agarwood, or they are the products of 2-(2-phenylethyl)chromone fragmentation, as in the case of *m/z* 107.085 and 161.128.[18]

The Mass Mountaineer software allows for data mining and statistical analysis in a rapid and efficient manner. The power of invoking KPCA and KDA of spectral data lies in the fact that one can analyze an unknown sample and compare it against a population, and not merely against a single reference standard. Since the country of origin and cultivation state (wild vs cultivated) of the agarwood voucher samples were known, the spectral data could be examined in multiple ways. This advantage is demonstrated in the analysis of the validation (Shoyeido) samples and by comparing the spectra of these samples against the standard reference collection. Eleven of the thirteen validation samples (85%) were correctly assigned to either cultivated or wild harvested for their respective geographic provenance.

Two of the 13 validation samples (15%) yielded incorrect statistical assignment, and we conclude that the misclassification is due to the very low amount of resin present. This suggests that a sample with exceedingly low resin content will not be able to be differentiated into cultivated or wild. However, it is very unlikely that any unknown sample of protected wild agarwood will have low amounts of resin, and this method will be able to classify such samples correctly.

The cultivated reference samples of agarwood from Vietnam were dark in appearance and rich in resin and the analysis showed a great diversity of ions. This suggests that as the resin content increases, the diversity of chromones may also increase. It is interesting to note that there were chromones in the cultivated agarwood that did not occur in wild agarwood samples and some distinguishing chromones in wild agarwood that were not found in the cultivated samples. Additional investigations are needed to determine what specific chromones and/or sesquiterpenes can be used as signature compounds for cultivated vs wild agarwood. This also could provide new information on the chemical signature of agarwood from different species and geographical locations. A recent study has shown that the species of *Aquilaria* in Vietnam may be different from the south to midland regions of the country, and a new species, *A. rugosa*, was recently identified.[19] Our analysis of unknown samples of agarwood from Vietnam showed two clusters, midland and south Vietnam, and this may reflect differences in agarwood produced by *A. crassna* and *A. rugosa*.

The limitations of making statistical inferences lie in the robustness of the reference database. When analyzing unknown samples, statistical tools make an assignment to the closest higher-dimensional space of the reference dataset. Therefore, the conclusions that are inferred are dependent on the size and validity of the reference sample database. The interpretation of wild vs cultivated is therefore empirical.

Lastly, DART-TOFMS coupled with statistical analysis is a reliable tool for chemo-taxonomic inferences and is probably the best tool today for species inference and geographic assignment of agarwood when morphology or DNA is absent.

Acknowledgements

The authors thank Doina Voin for her laborious persistence in procuring wood samples and to Henry Heuveling van Beek and Joel Jurgens for their help in obtaining research samples used in this study. We are thankful to Barry Baker and Mark Kirms for their critical review. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the U.S. Fish and Wildlife Service.

Rapid Communications in Mass Spectrometry

REFERENCES

[1] A. Barden, N. A. Anak, T. Mulliken, M. Song. *Heart of the matter: agarwood use and trade in CITES implementation for Aquilaria malaccensis*. Traffic International, Cambridge, **2000**.

[2] R. Naef. The volatile and semi-volatile constituents of agarwood, the infected heartwood of *Aquilaria* species: A review. *Flavour Fragr. J.* **2011**, *26*, 73.

[3] T. Soehartono, A. C. Newton. The gaharu trade in Indonesia: is it sustainable? *Econ. Bot.* **2002**, *56*, 271.

[4] H. Dai, J. Liu, Y. Zeng, Z. Han, H. Wang, W. Mei. Two new 2-(2-phenylethyl)chromones from Chinese eaglewood. *J. Asian Nat. Prod. Res.* **2010**, *12*, 134.

[5] H. Q. Chen, J. H. Wei, J. S. Yang, Z. Zhang, Y. Yang, Z. H. Gao, C. Sui, B. Gong. Chemical constituents of agarwood originating from the endemic genus *Aquilaria* plants. *Chem. Biodiv.* **2012**, *9*, 236.

[6] G. D. Paoli, D. R. Peart, M. Leighton, I. Samsoedin. An ecological and economic assessment of the nontimber forest product gaharu wood in Gunung Palung National Park, West Kalimantan, Indonesia. *Cons. Biol.* **2001**, *15*, 1721.

[7] R. A. Blanchette, H. Heuveling van Beek. Cultivated agarwood. U.S. 6,848, 211 6,848, 211, **2005**.

[8] Convention on International Trade in Endangered Species of Wild Fauna and Flora Appendices I, II, and III. UNEP-WCMC. Available: http://www.cites.org/eng/app/appendices.php (accessed March 29, 2012).

[9] IUCN. IUCN Red List of Threatened Species, version 2011.2, **2011**. Available: www.iucnredlist.org (accessed March 29, 2012).

[10] C. A. Lancaster, E. O. Espinoza. Evaluating agarwood products for 2-(2-phenylethyl)chromones using direct analysis in real time time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **2012**, *26*, 2649.

[11] R. B. Cody, J. A. Laramee, H. D. Durst. Versatile new ion source for the analysis of materials in open air under ambient conditions. *Anal. Chem.* **2005**, *77*, 2297.

[12] C. A. Lancaster, E. O. Espinoza. Analysis of select *Dalbergia* and trade timber using direct analysis in real time, time of flight mass spectrometry for CITES enforcement. *Rapid Commun. Mass Spectrom.* **2012**, *26*, 1147.

[13] R. B. Cody, A. J. Dane, B. Dawson-Andoh, E. O. Adedipe, K. Nkansah. 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.* **2012**, *95*, 134.

[14] S. D. Maleknia, T. M. Vail, R. B. Cody, D. O. Sparkman, T. L. Bell, M. A. Adams. Temperature-dependent release of volatile organic compounds of eucalypts by direct analysis in real time (DART) mass spectrometry. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 2241.

[15] S. Mika, G. Ratsch, J. Weston, B. Schölkopf, K. Müller. Neural networks for signal processing IX, 1999. *Proc. 1999 IEEE Signal Processing Society Workshop*, **1999**.

[16] B. Schölkopf, A. Smola, K.-R. Müller. Kernel Principal Component Analysis. Available: http://www1.cs.columbia.edu/~cleslie/cs4761/papers/scholkopf_kernel.pdf (accessed September 10, 2013).

[17] C. Souza. Kernel Principal Component Analysis in C#. Available: http://crsouza.blogspot.com/2010/01/kernel-principal-component-analysis-in.html (accessed September 10, 2013).

[18] B. Xia, J. Li, W. Mei, L. Ding, H. Xu, Y. Zhou. Tandem mass spectrometry fragmentation of the protonated 2-(2-phenylethyl)chromones from agarwood: radical ions versus non-radical ions. *J. Mass Spectrom.* **2013**, *48*, 979.

[19] K. Le Cong, P. J. A. Kessler, M. Eurlings. A new species of *Aquilaria* Thymelaeaceae from Vietnam. *Blumea* **2005**, *50*, 135.

