Published in final edited form as:
Phytochemistry. 2013 July ; 91: 187–197. doi:10.1016/j.phytochem.2013.02.018.

Accurate mass - time tag library for LC/MS-based metabolite profiling of medicinal plants

Daniel J. Cuthbertson[a,1], **Sean R. Johnson**[a], **Jasenka Piljac-Žegarac**[a,b], **Julia Kappel**[a,c], **Sarah Schäfer**[a,c], **Matthias Wüst**[c], **Raymond E. B. Ketchum**[a,2], **Rodney B. Croteau**[a], **Joaquim V. Marques**[a], **Laurence B. Davin**[a], **Norman G. Lewis**[a], **Megan Rolf**[d], **Toni M. Kutchan**[d], **D. Doel Soejarto**[e,f], and **B. Markus Lange**[a,*]

[a]Institute of Biological Chemistry and M.J. Murdock Metabolomics Laboratory, Washington State University, Pullman, WA 99164-6340, USA

[b]Ruđer Bošković Institute, Bijenička cesta 54, HR-10000 Zagreb, Croatia

[c]Institute of Nutrition and Food Sciences, University of Bonn, Endenicher Allee 11-13, 53115 Bonn, Germany

[d]Donald Danforth Plant Science Center, 975 N. Warson Road, St. Louis, MO 63132, USA

[e]Department of Medicinal Chemistry and Pharmacognosy, College of Pharmacy, University of Illinois at Chicago, 833 S. Wood St. (M/C 781), Chicago, IL 60612, USA

[f]Botany Department, Field Museum, 1400 S. Lake Shore Drive, Chicago, IL 60605-2496, USA

Abstract

We report the development and testing of an accurate mass – time (AMT) tag approach for the LC/MS-based identification of plant natural products (PNPs) in complex extracts. An AMT tag library was developed for approximately 500 PNPs with diverse chemical structures, detected in electrospray and atmospheric pressure chemical ionization modes (both positive and negative polarities). In addition, to enable peak annotations with high confidence, MS/MS spectra were acquired with three different fragmentation energies. The LC/MS and MS/MS data sets were integrated into online spectral search tools and repositories (Spektraris and MassBank), thus allowing users to interrogate their own data sets for the potential presence of PNPs. The utility of the AMT tag library approach is demonstrated by the detection and annotation of active principles in 27 different medicinal plant species with diverse chemical constituents.

Keywords

AMT tag library; mass spectrometry; metabolomics; natural product

*Corresponding author: Tel.: 509-335-3794; fax: 509-335-7643. lange-m@wsu.edu (B.M. Lange).
[1]Present address: Metabolomics Laboratory Application Group, Agilent Technologies, 5301 Stevens Creek Blvd., Santa Clara, CA 95051, USA
[2]Present address: Diana Plant Sciences, 18167 S.W. Boones Ferry Rd., Portland, OR 97224, USA

1. Introduction

Throughout human history plant extracts have been used in traditional medicines to treat a multitude of different ailments. These extracts often contain complex mixtures of structurally diverse small molecules generally referred to as plant natural products (PNPs), secondary (or specialized) metabolites, or phytochemicals (Li and Vederas, 2009). Isolated, pharmacologically active, PNPs have been the source of many successful drugs and their structures have inspired medicinal chemists to synthesize derivatives in order to further develop specific desirable properties. Examples of well-known PNP-derived anticancer drugs include paclitaxel (taxol®) (**1**) from Pacific yew (*Taxus brevifolia*), vincristine (**2**) from Madagascar periwinkle (*Catharanthus roseus*), camptothecin (**3**) (topotecan (**4**)) from the Happy Tree (*Camptotheca acuminata*, and podophyllotoxin (**5**) (etoposide (**6**)) from the Himalayan mayapple (*Podophyllum hexandrum*) (Newman and Cragg, 2012) (Fig. 1). Other important uses of PNPs, among others, are in pain management (e.g., morphine (**7**) from opium poppy [*Papaver somniferum*]) and as antimalarial treatments (e.g., quinine (**8**) from the genus *Cinchona* or artemisinin (**9**) from Sweet wormwood [*Artemisia annua*]) (Newman and Cragg, 2012) (Fig. 1). From the mid 1990s to mid 2000s, there was a decline of PNP research in the pharmaceutical industry, in part because of a lack of compatibility of plant extract libraries with high-throughput screening technologies (Koehn and Carter, 2005). However, recent innovations in plant extract screening, combined with the realization that structurally complex and diverse libraries are difficult to synthesize using high-throughput approaches, have led to a renewed interest in PNP-based drug discovery (Li and Vederas, 2009).

The plant kingdom is a source of extraordinary chemical diversity (Buckingham, 2011). It is thus essential to develop fast and reliable methods for the chemical screening of plant extracts. Liquid chromatography (LC) is the most common method used for the separation of PNPs. The LC system can be interfaced with various different detectors, which are often used in series. Diode array, fluorescence and refraction index detectors provide excellent sensitivity and dynamic range, but the extractable structural information is limited. High resolution mass spectrometers (HRMS) are exceptionally sensitive and, in particular in MS/MS mode, contribute highly valuable structural information, but the dynamic range achievable with these instruments can be limiting, when one considers that the concentration range of metabolites in plants can exceed 10^6 between the most and least abundant metabolites (Moco et al., 2007). Nuclear magnetic resonance spectroscopy is immensely important for structural elucidation, particularly of complex molecules with (multiple) chiral center(s), but its limited sensitivity has placed constraints on its online use with LC (Koehn, 2008). A full elucidation of novel chemical structures requires, however, the use of several complementary chromatographic and spectroscopic techniques. Yet, if authentic standards are available, a combination of (ultra) high performance LC with HRMS is a very rapid and powerful approach for the structural identification of PNPs in complex plant matrices. HRMS analyses can be used to deduce empirical formulas from accurate mass-over-charge (m/z) determinations, which is of great importance for metabolite profiling, extract authentication and dereplication in screening efforts (Kind and Fiehn, 2007). However, it is crucial to perform these analyses using various different ionization modes and polarities to assess if protonated or deprotonated ions are formed and to which extent adducts or fragments need to be considered. The confidence level for peak annotations is further increased when such data sets are combined with information about chromatographic behavior (such as (relative) retention time on a specific stationary phase under controlled LC conditions) (Korfmacher, 2005). Additional complementary information can be acquired in MS/MS or MS^n mode, which can be used to determine the decoration patterns of glycosylated or prenylated PNPs, the presence of certain structural cores or, if a large

number of structurally related compounds is available, to establish fragmentation rules for structure prediction (Kind and Fiehn, 2010; Sawada et al., 2012).

A particularly powerful strategy, termed accurate mass – time tag (AMT) analysis, was originally developed for the integration of orthogonal LC and HRMS data sets, and enabled a higher throughput for quantitative shotgun proteomics by eliminating (or at least drastically reducing) the necessity of repetitive MS/MS analyses (Bruce et al., 1999; Shen et al., 2001). Further standardization has facilitated the development of AMT tag databases for the proteomics community (Tarasova et al., 2009). Online resources for MS/MS spectral searches of PNPs have been developed recently (METLIN, Smith et al., 2005; MassBank, Horai et al., 2010; ReSpect, Sawada et al., 2012). However, the integration of LC, HRMS and MS/MS data in metabolomics experiments has only recently been reported for human lipids (Ding et al., 2008; Sorensen et al., 2010). Herein reported is the successful adaptation of the AMT tag approach for PNP discovery in medicinal plant extracts. Our PNP spectral library uses the MassBank output format and can be integrated into and interrogated by various online resources, including our own Spektraris tool (http://langelabtools.wsu.edu/amt/).

2. Results and discussion

2.1 AMT tag library design

To ensure consistent chromatographic results, a normalized LC flow path was used in combination with a rapid resolution LC method suitable for the reversed phase separation of a wide range of bioactive PNP analytes. A mid-polarity compound was then sought that was not present in plant extracts and would ionize under all relevant ionization conditions, thus enabling the calculation of relative retention times. Among several tested candidate compounds, anthracene-9-carboxylic acid (**10**) was selected as an internal standard with the desired properties (Sana et al., 2008a). Authentic standards of PNPs were arranged in bins of 8–16 compounds each (52 bins total) and mixed with the internal standard before injection onto the LC system.

HRMS data were acquired with standardized methods using electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI) sources in both positive (pos) and negative (neg) polarity modes, thus generating 4 separate result files for each compound bin (Fig. 2A). ESI is generally a softer ionization method compared to APCI, which means that fewer fragments are formed (Wolfender, 2009), and this is also reflected in our data sets (Supplementary Table 1). Protonation in positive mode (generating the $[M+H]^+$ pseudo-molecular ion) tends to work best with compounds that contain basic functional groups that form stable cations, such as amines. An example would be the alkaloid skimmianine (**11**) (peak 5 in the chromatograms reproduced in Fig. 2A), which ionizes well in ESI(pos) and APCI(pos), but is difficult to detect in ESI(neg) or APCI(neg). Deprotonation in the negative mode (generating the $[M-H]^-$ pseudo-molecular ion) is most useful with acidic species, including hydroxyl and carboxyl groups. The multiply oxygenated quinoline alkaloid foliosidine (**12**) (peak 8 in the chromatograms reproduced in Fig. 2A) would be an example of a metabolite that is readily detected in ESI(neg), while the diterpene forskolin (**13**) ionizes well in all modes and polarities. Cationization in positive mode (formation of adducts most commonly containing alkali or ammonium ions) or anionization in negative mode (formation of adducts most commonly containing acetate or formate ions), and the formation of multimers can lead to the occurrence of additional ions in the mass spectrum. Thus, the m/z values of all ion species detected by HRMS using a QTOF mass spectrometer (excluding fragments) were recorded and interpreted (Supplementary Table 1). Agilent's molecular formula generation (MFG) algorithm was then used to calculate and score empirical molecular formulas (Sana et al., 2008b) (Fig. 3). HRMS data were acquired for a

total of 740 compounds, of which 492 gave satisfactory spectral results when highly stringent selection criteria were employed (100 % match between empirical and actual molecular formula; > 95 % match of isotopic abundance and spacing patterns; minimum relative retention time of 0.05). For each detected compound, the accurate mass in all modes and polarities (mass tags), relative retention time (time tag), and relevant meta data (e.g., trivial name, IUPAC name, and CAS number) were compiled using a custom script (MS^1 Extractor) and stored in an AMT tag database that uses the MassBank record format (Horai et al., 2010) (Fig. 3). The meta data set for the AMT tag library is available for download as Supplementary Table 1.

MS/MS data were acquired at three different collision energies in separate analytical runs (Fig. 2). This was done to sample at least 5 times across each mass peak, thus ensuring accurate mass measurements. With newer, faster scanning QTOFs, such analyses could potentially be performed in individual runs (Kueger et al., 2012). The patterns generated in MS/MS mode by fragmentation of compounds in the QTOF collision cell depend on the energy applied and are indicative of structural features. For example, when **11** (peak 5 in the chromatograms reproduced in Fig. 2A) was analyzed in MS/MS mode, the $[M+H]^+$ peak was very prominent at a collision energy of 10 eV, while the $[M-CH_3-H_2O+H]^+$ peak was the most abundant at 20 eV, and several fragments in the lower mass range were detected at 40 eV (Fig. 2B). A custom script (MS^2 Extractor) was used to associate MS/MS data files with relevant meta information, and the resulting data files were integrated into the AMT tag library in MassBank record format (Horai et al., 2010) (Fig. 3). All MS and MS/MS records are available for download on the Spektraris websites (http://langelabtools.wsu.edu/amt/) and MS/MS data files have also been submitted to MassBank (www.massbank.jp).

2.2 AMT tag library searching

Raw data files obtained in LC-MS runs with plant extracts were first processed using Agilent's molecular feature extraction (MFE) algorithm within the MassProfiler software (Sana et al., 2008b) (Fig. 4). This algorithm enables (1) the grouping of multiple ions generated from the same compound (e.g., molecular ion, adducts and dimers) to one molecular feature and (2) the automated listing of all molecular features (compounds) in a chromatographic run. From the output of the MFE analysis in .csv (comma-separated values) format, the following information was extracted: (1) unique compound IDs, (2) retention times or relative retention times (time tags), and (3) m/z values (mass tags), monoisotopic masses or empirical formulas. These data sets were then copied into the appropriate data entry field of our Spektraris online resource (http://langelabtools.wsu.edu/amt/). An online search against Spektraris generates result files that include the full meta data for each compound, including alternative names, CAS number, chemical structure (.mol), SMILES (Simplified Molecular-Input Line-Entry System) representation, relative retention time, and other relevant information (Fig. 5). Our MS/MS data sets can also be accessed via the online version of the MassBank community resource (Horai et al., 2010) (www.massbank.jp). Alternatively, a distributed copy of the MassBank database can be downloaded on a local computer and searches against our (and other) data sets can be performed.

2.3 Testing the robustness of the accurate mass - time tag approach

As a first test the reproducibility of the retention time and HRMS data over longer periods of time was determined. A standard mix was run on LC columns from different production lots between 2009 and 2012. The relative standard deviation for the retention times of 5 individual components ranged between 0.35 and 1.52 % (n=11) (Supplementary Table 2), and all analytes were properly identified by a Spektraris search, thus providing evidence for the utility of the approach when used with columns of different ages and use patterns, and a

mass spectrometer that was used for various different applications between replicate analyses. It is, however, important to note that regular system maintenance is a prerequisite for reliable peak annotation. When the LC and QTOF systems had been in use for other projects and not been cleaned thoroughly, an increased tendency for the formation of adducts was observed, thus complicating spectral pattern matching. In general, peak annotations based on MS/MS fragmentation patterns confirmed those made by the AMT tag approach. However, examples were also documented when MS/MS data acquisition proved essential for high confidence peak annotation. This was particularly true for isobaric compounds (same molecular formula) with identical chromatographic behavior (same retention time). For example, the flavonoids 7,3′-dihydroxyflavone (**14**) and 2′6-dihydroxyflavone (**15**) could not be distinguished by the AMT tag approach but MS/MS patterns obtained at different collision energies were clearly discernible (Supplementary Fig. 1), thus demonstrating the potential of combining AMT tag and MS/MS data sets for aiding PNP identification.

2.4 Validating the AMT tag approach by screening chemically diverse plant extracts for known active principles

To evaluate the utility of the AMT tag library for the tentative identification of PNPs, extracts from medicinal plants with known active principles and diverse chemistry were analyzed. While presenting the results of these analyses, the pharmaceutical relevance of the PNPs of interest is also briefly described for each.

Alkaloid-containing extracts were obtained from various tissues of 16 different species (29 samples total) (Table 1). Using the Spektraris tool, the presence of 11 signature alkaloids was evaluated in these extracts. Most alkaloids, with the exception of **8** (detectable in all modes and polarities), were detected preferentially in ESI(pos) and APCI(pos), whereas only very weak signals, if any, were detected in ESI(neg) and APCI(neg). The quinoline alkaloid camptothecin (**3**) was detected, as expected, in leaf, root and stem tissue of the Happy Tree (*Camptotheca acuminata* Decne.; Nyssaceae). Semisynthetic analogues of **3**, including **4**, are approved chemotherapeutic agents for the treatment of ovarian and lung cancer (Venditto and Simanek, 2010). Both harringtonine (**16**) and homoharringtonine (**17**) were found to be present in leaves of *Cephalotaxus drupaceae* and *C. harringtonia* ‘Duke Gardens’ (Cephalotaxaceae), whereas only small amounts were detectable in leaves of *C. harringtonia* ‘Watnong’. Homoharringtonine (**17**) has been investigated in phase 2 and phase 3 clinical trials for the treatment of chronic myeloid leukemia (Quintás-Cardama et al., 2009). The quinoline alkaloid quinine (**8**), the first effective treatment for malaria (Achan et al, 2011), was detected in bark, leaf, root and xylem extracts from *Cinchona pubescens* Vahl (Rubiaceae). Based on our AMT tag data, colchicine (**18**) was present in corm, fruit and leaf tissue extracts from autumn crocus (*Colchicum autumnale* L.; Colchicaceae) and in extracts of fruit, leaf and tuber tissue from flame lilly (*Gloriosa superba*; Colchicaceae). *Colchicum* has been used for the treatment of gout since ancient times and isolated colchicine (**18**) has additional commercial uses in cellular genetics as a mitotic poison (Bhat et al., 2009). The tropane alkaloids atropine (**19**) and scopolamine (**20**) were detected in leaf extracts of 5 tested species of the genus *Datura* (Solanaceae) and Black henbane (*Hyoscyamus niger* L.). Atropine (**19**) has a long history as an ophthalmic agent and is also useful in treating bradycardia (slow resting heart rate) or heart block, while scopolamine (**20**) is used to treat motion sickness (Grynkiewicz and Gadzikowska, 2008). The alkaloid galanthamine (**21**), which is approved in several countries for the treatment of Alzheimer’s disease (Birks, 2006), was found to be present in bulb, flower and leaf tissues of wild daffodil (*Narcissus pseudonarcissus* L. (Amaryllidaceae). The sporophyte of *Huperzia squarrosa*, a clubmoss, was demonstrated to contain the alkaloid huperzine A (**22**), which is currently being investigated in advanced clinical trials for the treatment of vascular dementia in China (Hao

et al., 2009). The benzylisoquinoline alkaloids morphine (**7**) and codeine (**23**) were readily detectable in capsule, leaf and root extracts of the opium poppy (*Papaver somniferum* L.; Papaveraceae). The opioids have a long history of use as analgesics (**7** and **23**) and as a cough suppressant (**23**) (Trescot et al., 2008).

Methanolic extracts of 9 tissue samples obtained from 5 different terpenoid-producing and 7 samples from 3 different phenylpropanoid-accumulating species were used to further assess the utility of the AMT tag library for PNP screening purposes (Table 2). Diterpenoids were generally detectable in all modes and polarities (with the exception of triptolide (**24**), which did not ionize well in APCI(neg) and the triterpenoid glycoside digoxin (**25**), which was only detectable in ESI(neg) and APCI(neg)). Phenylpropanoids were detectable in all modes and polarities tested, as described below. The diterpenoid lactone andrographolide (**26**), which has anti-inflammatory, immunosuppressive and neuroprotective activities (Lim et al., 2012), was detected in leaf extracts of Kalmegh (*Andrographis paniculata* (Burm. F.) Nees; Acanthaceae). The labdane diterpenoid forskolin (**13**), which has been investigated in preliminary clinical trials as a weight loss aid (Godard et al., 2005), was found to be present in extracts of roots (fairly high amounts) and leaves (very low amounts) from *Plectranthus barbatus* (syn. *Coleus forskohlii*; Lamiaceae) but was undetectable in leaf and root extracts of *P. caninus* Roth (syn. *Coleus caninus*), which is consistent with the published occurrence of **13** in the genus *Plectranthus* (Waldia et al., 2011). The anti-cancer agent paclitaxel (taxol®) (**1**) (Kingston, 2007) was detected in extracts from *Taxus* x *media* cell suspension cultures. Root extracts of Regel's threewingnut (*Tripterygium regelii* Sprag. & Takeda; Celastraceae) were found to contain the diterpene triepoxide triptolide (**24**), which is currently being evaluated in clinical trials as a treatment for rheumatoid arthritis auto-immune diseases and cancers (Zhou et al., 2012). As expected, the cardiac glycoside digoxin (**25**), which is used for the treatment of certain heart conditions (Ehle et al., 2011), was readily detected in leaf and root extracts of foxglove (*Digitalis lanata* Ehrh.; Plantaginaceae). Leaf and rhizome extracts from *Podophyllum hexandrum* Royle and *P. peltatum* L. (Berberidaceae) contained the lignan podophyllotoxin (**5**), which is the pharmacophore for the anticancer drug etoposide (**6**) (Baldwin and Osheroff, 2005). Flower, leaf and root extracts from Golden flax (*Linum flavum* L.; Linaceae) also contained the structurally related lignan, 5-methoxypodophyllotoxin (**27**). Finally, the lignan nordihydroguaiaretic acid (**28**), an antioxidant with various promising biological activities and potential clinical applications (Lü et al., 2010), was detected in leaf samples of the creosote bush (*Larrea tridentata* (DC.) Coville).

Our Spektraris searches often returned hits for various metabolites, in addition to the known active principles mentioned above, but, when stringent criteria (small tolerances for m/z and relative retention time) were used, cases of highly unlikely hits were very rare. For example, in root extracts of *Cinchona pubescens* Vahl, in addition to quinine (**8**), several other known constituents such as cinchonamine (**29**), cinchonine (**30**), hydrocinchonine (**31**), and hydroquinidine (**32**) were tentatively identified (Kaufmann and Rúveda, 2005) (Supplementary Table 3), but other unrelated alkaloids were not found to be present. In *Taxus* cell cultures, baccatin III (**33**), 10-deacetylbaccatin (**34**), and cephalomannine (**35**) were detected with high confidence, in addition to the anticancer agent taxol (**1**) (Ketchum and Croteau, 2006) (Supplementary Table 3). A larger number of other putative taxanes was also detected but a unequivocal identification would require the inclusion of data for additional authentic taxane standards into the database. The expansion of our database with additional PNPs is thus our highest priority.

2.5 Testing the utility of AMT tag data acquisition methodology for PNP quantitation

The data acquisition parameters used to test medicinal plant extracts were chosen with the intention to enable qualitative analyses of PNPs. When screening various species within a genus, such data can be very valuable to support the decision process with regard to which plants or extracts to select for follow-up analyses. For PNP screening and follow-up biosynthetic studies it is also important to gather quantitative data on the tissue-specific accumulation of PNPs. Thus, methods were validated for the quantification of **26** and **13** in different tissues of *Andrographis paniculata* and *Plectranthus barbatus*, respectively (Supplementary Data File 1). These validation experiments, which assessed (1) the recovery of the analyte, (2) the precision, sensitivity, and linearity of analyte detection, and (3) ion suppression effects within the QTOF ion source, employed the same acquisition parameters as those used for the AMT tag library.

The andrographolide (**26**) concentration in greenhouse-grown *A. paniculata* plants was highest in leaves (up to 2.4 ± 0.3 % of dry weight biomass), but the metabolite was also detectable in stems (up to 0.4 ± 0.03 % of dry weight biomass), whole seedlings (up to 0.3 ± 0.04 % of dry weight biomass) and flowers (up to 0.07 ± 0.005 % of dry weight biomass) (Fig. 6A). Andrographolide (**26**) was undetectable in root and seed extracts. These patterns of accumulation are in general agreement with literature reports, although the determination of **26** in flowers and seeds has not yet been reported by others. Pholphana et al. (2004) used an HPLC-based method to determine the contents of **26** in leaves and stems (1.7 and 0.8 % of dry weight biomass) of *A. paniculata* harvested in Thailand. Akowuah et al. (2006) reported contents of **26** (determined by HPLC) between 2.1 to 3.0 % of dry weight biomass in leaves of *A. paniculata* harvested at different locations within Malaysia.

In greenhouse-grown *P. barbatus* plants, the concentrations of forskolin (**13**) were highest in roots (0.09 ± 0.02 % of dry weight biomass), with small amounts in stems and seedlings (< 0.002 % of dry weight biomass) (Fig. 6B). Forskolin (**13**) was undetectable in leaf and seed extracts. Literature reports for concentrations of **13** in roots vary between 0.01 and 0.2 % of dry weight biomass, whereas the metabolite was not detected in leaves (Schaneberg and Khan, 2003; Saleem et al., 2006; Ahmad et al., 2008; Zhang et al., 2009), which is consistent with the data presented herein. These examples demonstrate that the methods used for acquiring data for our AMT tag database can be readily adapted for quantitative follow-up studies.

3. Concluding remarks

Described in this study is an AMT tag library approach to improve the rapid preliminary identification of PNPs in LC/MS runs of medicinal plant extracts. By employing orthogonal data sets, including accurate mass/empirical formula, relative retention time and MS/MS fragmentation patterns, peak annotation can be carried out with high confidence. It would be desirable to include a range of more common metabolites to the AMT tag library in the future, so that users can benchmark their extractions. The LC method used for the existing library is very versatile and allows the separation of metabolites in the medium to low polarity range, and the inclusion of data for additional phenylpropanoids/flavonoids, alkaloids, terpenoids and lipids in the database would thus further increase its utility.

However, the current LC method has limited resolving power for polar metabolites (e.g., sugars and sugar phosphates, short-chain amino acids and organic acids, nucleotides and sugar nucleotides), which generally elute in the void volume. Hydrophilic interaction chromatography has been used successfully for the untargeted separation of polar metabolites for several years (reviewed in Cubbon et al., 2010). An emerging alternative is the use of silica hydride-based stationary phases (Pesek et al., 2008; 2011), which has been

employed successfully for polar metabolome analyses to differentiate microbial strains (Weisenberg et al., 2009), study microbial metabolism of specific carbon substrates (de Carvalho et al., 2010a), profile enzyme function (de Carvalho et al., 2010b), discover biomarkers to indicate human cancer progression (Putluri et al., 2011), and enable the broad-spectrum profiling of biological extracts (Callahan et al., 2009). The expansion of the AMT tag database to include polar metabolites would enable the tentative identification of a much broader range of metabolites, which would be another significant contribution to improving peak annotation in LC/MS-based metabolomics.

An alternative to acquiring MS/MS data at different fragmentation energies would have been to perform MS^n (higher order MS/MS) analyses with a high resolution orbitrap instrument. MSn analyses can also be used to calculate elemental compositions of analytes (Rojas-Chertó et al., 2011). The development of online resources for depositing and searching these types of data should be considered as a priority for the metabolomics community.

4. Experimental

4.1. Plant growth conditions

The sourcing of seeds and plant materials is summarized in Supplementary Table 4. This table also provides details about the growth conditions in greenhouses at different locations. Voucher specimens were deposited, after careful morphological and phytochemical evaluation, with the herbarium of the Field Museum of Natural History in Chicago, IL (Supplementary Data File 2).

4.2. Tissue harvest and metabolite extractions

A detailed description of the methods used for extracting various metabolite classes from different medicinal plants is provided as Supplementary Protocol 1.

4.3. Liquid chromatography mass spectrometry

HPLC-MS and HPLC-MS/MS runs were performed using a 1200 Rapid Resolution HPLC system coupled to a 6520 QTOF mass spectrometer (Agilent Technologies). Instrument control and data acquisition used the MassHunter Acquisition Software (Revision B.02.01; Agilent Technologies). MS and MS/MS data were processed and analyzed with the MassHunter Qualitative Analysis Software (Revision B.03.01; Agilent Technologies). Chromatographic separations were achieved on a ZORBAX SB-Aq column (2.1 × 50 mm, 1.8 μm particle size) connected to a ZORBAX SB-C8 guard column (2.1 × 30 mm, 3.5 μm particle size) (Agilent Technologies). The temperature of the column compartment was kept at 60°C and the solvent flow rate was set to 0.6 mL/min. Analytes were separated using a linear gradient from 98 % solvent A (0.2% AcOH) and 2 % solvent B (0.2% AcOH in MeOH) to 98 % B and 2 % A in 13 min. These final conditions were maintained for 6 min, before the column was returned to the initial solvent composition for another 5 min of equilibration. Injection volumes varied between 1 and 5 μL. A reference mass solution (containing hexakis(1H,1H,3H-tetrafluoropropoxy)phosphazine (250 nM) and purine (300 nM)) was infused into the ion source using an isocratic pump at a flow rate of 0.05–0.1 mL/min. Alternatively, the reference mass solution was introduced from a reservoir using a second sprayer needle. Varying flow rates were used in different modes and polarities to achieve a reference ion abundance of roughly 40,000 counts. The QTOF mass spectrometer was set to 2 GHz high gain in positive ion mode for MS experiments and 4 GHz high resolution mode for MS/MS experiments, with the following settings: the drying gas flow was 5 L min^{-1}, the nebulizer pressure was 60 psi, the drying gas temperature was 325°C, and the scan rates were set to 1.4 scans s^{-1} for MS and to 4 scans s^{-1} for MS/MS. Data

acquisition was performed in four different modes and polarities: electrospray (ESI)/positive polarity, ESI/negative polarity, atmospheric pressure chemical ionization (APCI)/positive polarity and APCI/negative polarity. The source vaporizer temperature was set to 450°C in APCI mode. MS/MS spectra were acquired at three different collision energies: 10, 20 and 40 eV.

4.4. Generation of the AMT tag library

Authentic standards were combined in bins (8 – 16 compounds per bin) at a final concentration of 1–10 ng/μL in MeOH. Each bin also contained the internal standard (IS), anthracene-9-carboxylic acid (**10**), thus allowing the calculation of relative retention times ($Rt_{(analyte)}$ / $Rt_{(IS)}$). The composition of the bins is listed in Supplementary Table 1. The raw data file for each bin was opened and extracted ion chromatograms were generated for all target m/z values. The target ion and its isotope peaks were selected and the Molecular Formula Generation (MFG) algorithm of the MassHunter software (Agilent Technologies) was used to calculate the elemental composition of each analyte (Bristow et al., 2008). Chromatograms were also screened for the occurrence of common adducts ($[M+Na]^+$ in ESI(pos), $[M+CH_3COO]^-$ in ESI(neg) and $[M+Cl]^-$ in APCI(neg)) and the presence of dimers and/or trimers (the target ions for all compounds are listed in Supplementary Table 1). A unique identifier was assigned to each PNP, and the monoisotopic mass, molecular formula, theoretic masses of the molecular ion and all adducts were calculated and tabulated in EXCEL format (Supplementary Table 1). The chromatographic properties (retention time and relative retention time) of each PNP and the relative intensities of all relevant ions in all modes and polarities were included in the AMT tag library file as well (Supplementary Table 1). A custom script, termed MS^1 Extractor, was then used to integrate entries from the AMT tag library file with additional information about each PNP from the PubChem online database (list of associated names, SMILES string, IUPAC name, formula, monoisotopic mass and the chemical structure in MDL .mol file format) and general information about data acquisition (instruments, HPLC protocol, MS parameters, and data acquisition and analysis software packages). The MS^1 Extractor tool generates an output file in MassBank format (Horai et al., 2010) that includes a listing of all detected peaks, their relative intensity (rounded to 999 for the peak of highest intensity according to MassBank record format requirements), and evaluates if the monoisotopic mass and formula for each PNP match between those calculated based on our mass spectrometric data and those listed in the PubChem database.

Based on the MS data sets for each PNP, a target ion for MS/MS fragmentation was selected, and MS/MS data were acquired at three different collision energies (10, 20 and 40 eV) in three separate chromatographic runs. Nitrogen was used as a collision gas. MS/MS raw data files were opened and the MS/MS data extracted using the MassHunter software (Agilent Technologies). Data for at least three MS/MS spectra were averaged and the results exported to a file in .csv format. A custom script, termed MS^2 Extractor, was then employed to merge the spectral data with meta information about the PNP and data acquisition conditions as listed above.

4.5. Searching against the AMT tag library

Raw data files were opened using the MassHunter qualitative analysis software (Agilent Technologies) and the Molecular Feature Extraction (MFE) algorithm was employed to locate individual sample components (Sana et al., 2008). The Molecular Formula Generation (MFG) algorithm was then used to calculate molecular formulas based on mss spectrometric data. These data sets were exported to a file in .csv format and saved. The exported file was opened in Microsoft EXCEL and the contents of columns with the molecular feature ID, retention time, calculated mass, m/z value, and the molecular formula (last value is optional)

copied into the query field of the Spektraris search tool (http://langelabtools.wsu.edu/amt/). The retention time of the internal standard was added and tolerances for selected search parameters were entered (± 0.05 for relative retention time (unitless) and ± 10 ppm for m/z value). MS/MS data files were also integrated into the online MassBank community resource (www.massbank.jp). All data sets can be downloaded, in MassBank format, from the Spektraris website (http://langelabtools.wsu.edu/amt/).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health (award number 1RC2GM092561-01). The authors would like to thank Mr. Blake K. Herron for technical assistance with running mass spectrometric analyses and Ms. Amy Hetrick for technical assistance with obtaining and maintaining plants. The conceptual input from Drs. T. Sana and S. Fisher (Agilent Technologies, Inc.) is also greatly acknowledged.

References

Achan J, Talisuna AO, Erhart A, Yeka A, Tibenderana JK, Baliraine FN, Rosenthal PJ, D'Alessandro U. Quinine, an old anti-malarial drug in a modern world: role in the treatment of malaria. Malaria J. 2011; 10:144.

Ahmad S, Rizwan M, Parveen R, Mujeeb M, Aquil M. A validated stability-indicating TLC method for determination of forskolin in crude drug and pharmaceutical dosage form. Chromatographia. 2008; 67:441–447.

Akowuah GA, Zhari I, Norhayati I, Mariam A. HPLC and HPTLC densitometric determination of andrographolides and antioxidant potential of *Andrographis paniculata*. J Food Comp Anal. 2006; 19:118–126.

Baldwin EL, Osheroff N. Etoposide, topoisomerase II and cancer. Curr Med Chem. 2005; 5:363–372.

Bhat A, Naguwa SM, Cheema GS, Gershwin ME. Colchicine revisited. Ann New York Acad Sci. 2009; 1173:766–773. [PubMed: 19758227]

Birks J. Cholinesterase inhibitors for Alzheimer's disease. Cochrane Database Syst Rev. 2006; (1):Art. No.: CD005593.

Bristow T, Constantine J, Harrison M, Cavoit F. Performance optimization of a new-generation orthogonal-acceleration quadrupole-time-of-flight mass spectrometer. Rapid Commun Mass Spectrom. 2008; 22:1213–1222. [PubMed: 18348222]

Bruce JE, Anderson GA, Wen J, Harkewicz R, Smith RD. High mass measurement accuracy for 100% sequence coverage of enzymatically digested bovine serum albumin from an ESI-FTICR mass spectrum. Anal Chem. 1999; 71:2595–2599. [PubMed: 10424157]

Buckingham, J. Dictionary of natural products on DVD. Chapman and Hall/CRC; London, UK: 2011. Version 19:1

Callahan DL, De Souza D, Bacic A, Roessner U. Profiling of polar metabolites in biological extracts using diamond hydride-based aqueous normal phase chromatography. J Sep Sci. 2009; 32:2273–2280. [PubMed: 19569107]

Cubbon S, Antonio C, Wilson J, Thomas-Oates J. Metabolomic applications of HILIC-LC-MS. Mass Spectrom Rev. 2010; 29:671–684. [PubMed: 19557839]

de Carvalho LP, Fischer SM, Marrero J, Nathan C, Ehrt S, Rhee KY. Metabolomics of *Mycobacterium tuberculosis* reveals compartmentalized co-catabolism of carbon substrates. Chem Biol. 2010a; 17:1122–1131. [PubMed: 21035735]

de Carvalho LP, Zhao H, Dickinson CE, Arango NM, Lima CD, Fischer SM, Ouerfelli O, Nathan C, Rhee KY. Activity-based metabolomic profiling of enzymatic function: identification of Rv1248c as a mycobacterial 2-hydroxy-3-oxoadipate synthase. Chem Biol. 2010b; 17:323–332. [PubMed: 20416504]

Ding J, Sorensen CM, Jaitly N, Jiang H, Orton DJ, Monroe ME, Moore RJ, Smith RD, Metz TO. Application of the accurate mass and time tag approach in studies of the human blood lipidome. J Chromatogr B. 2008; 871:243–252.

Ehle M, Chandni P, Giuliano RP. Digoxin: clinical highlights: a review of digoxin and its use in contemporary medicine. Crit Pathw Cardiol. 2011; 10:93–98. [PubMed: 21988950]

Godard MP, Johnson BA, Richmond SR. Body composition and hormonal adaptations associated with forskolin consumption in overweight and obese men. Obesity Res. 2005; 13:1335–1343.

Grynkiewicz G, Gadzikowska M. Tropane alkaloids as medicinally useful natural products and their synthetic derivatives as new drugs. Pharmacol Rev. 2008; 60:439–463.

Hao, Z.; Liu, M.; Lu, DH. Cochrane Database Syst Rev. 2009. Huperzine A for vascular dementia; p. Art. No.: CD007365

Horai H, Arita M, Kanaya S, Nihei Y, Ikeda T, Suwa K, Ojima Y, Tanaka K, Tanaka S, Aoshima K, Oda Y, Kakazu Y, Kusano M, Tohge T, Matsuda F, Sawada Y, Hirai MY, Nakanishi H, Ikeda K, Akimoto N, Maoka T, Takahashi H, Ara T, Sakurai N, Suzuki H, Shibata D, Neumann S, Iida T, Tanaka K, Funatsu K, Matsuura F, Soga T, Taguchi R, Saito K, Nishioka T. MassBank: a public repository for sharing mass spectral data for life sciences. J Mass Spectrom. 2010; 45:703–714. [PubMed: 20623627]

Kaufman TS, Rúveda EA. The quest for quinine: those who won the battles and those who won the war. Angew Chem Int Ed. 2005; 44:854–885.

Ketchum, REB.; Croteau, RB. The Taxus metabolome and the elucidation of the taxol® biosynthetic pathway in cell suspension cultures. In: Saito, K.; Dixon, RA.; Willmitzer, L., editors. Biotechnology in Agriculture and Forestry. Vol. 57. Springer Verlag; Berlin: 2006. p. 291-309.

Kind T, Fiehn O. Seven golden rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. BMC Bioinformatics. 2007; 8:105. [PubMed: 17389044]

Kind T, Fiehn O. Advances in structure elucidation of small molecules using mass spectrometry. Bioanal Rev. 2010; 2:23–60. [PubMed: 21289855]

Kingston DGI. The shape of things to come: structural and synthetic studies of taxol and related compounds. Phytochemistry. 2007; 68:1844–1854. [PubMed: 17184797]

Koehn FE, Carter GT. The evolving role of natural products in drug discovery. Nature Rev Drug Discov. 2005; 4:206–220. [PubMed: 15729362]

Koehn FE. High impact technologies for natural products screening. Progr Drug Res. 2008; 65:177–210.

Korfmacher WA. Principles and applications of LC-MS in new drug discovery. Drug Discov Today. 2005; 10:1357–1367. [PubMed: 16253874]

Kueger S, Steinhauser D, Willmitzer L, Giavalisco P. High-resolution plant metabolomics: from mass spectral features to metabolites and from whole-cell analysis to subcellular metabolite distributions. Plant J. 2012; 70:39–50. [PubMed: 22449042]

Li JWH, Vederas JC. Drug discovery and natural products: end of an era or an endless frontier? Science. 2009; 325:161–165. [PubMed: 19589993]

Lim JCW, Chan TK, Sagineedu SR, Stanslas J, Wong WSF. Andrographolide and its analogues: versatile bioactive molecules for combating inflammation and cancer. Clin Exp Pharmacol Physiol. 2012; 39:300–310. [PubMed: 22017767]

Lü JM, Nurko J, Wakley SM, Jiang J, Kougias P, Lin PH, Yao Q, Chen C. Molecular mechanisms and clinical applications of nordihydroguaiaretic acid (NDGA) and its derivatives: An update. Med Sci Monit. 2010; 16:RA93–R100. [PubMed: 20424564]

Moco S, Bino RJ, De Vos RCH, Vervoort J. Metabolomics technologies and metabolite identification. Trends Anal Chem. 2007; 26:855–866.

Newman DJ, Cragg GM. Natural products as sources of new drugs over the 30 years from 1981 to 2010. J Nat Prod. 2012; 75:311–335. [PubMed: 22316239]

Pesek JJ, Matyska MT, Fischer SM, Sana TR. Analysis of hydrophilic metabolites by high-performance liquid chromatography-mass spectrometry using a silica hydride-based stationary phase. J Chromatogr A. 2008; 1204:48–55. [PubMed: 18701108]

Pesek JJ, Matyska MT, Fischer SM. Improvement of peak shape in aqueous normal phase analysis of anionic metabolites. J Sep Sci. 2011; 34:3509–3516. [PubMed: 22009714]

Pholphana N, Rangkadilok N, Thongnest S, Ruchirawat S, Ruchirawat M, Satayavivad J. Determination and variation of three active diterpenoids in *Andrographis paniculata* (Burm F.). Nees Phytochem Anal. 2004; 15:365–371.

Putluri N, Shojaie A, Vasu VT, Vareed SK, Nalluri S, Putluri V, Thangjam GS, Panzitt K, Tallman CT, Butler C, Sana TR, Fischer SM, Sica G, Brat DJ, Shi H, Palapattu GS, Lotan Y, Weizer AZ, Terris MK, Shariat SF, Michailidis G, Sreekumar A. Metabolomics profiling reveals potential markers and bioprocesses altered in bladder cancer progression. Cancer Res. 2011; 71:7376–7386. [PubMed: 21990318]

Rojas-Chertó M, Kasper PT, Willighagen EL, Vreeken RJ, Hankemeier T, Reijmers TH. Elemental composition determination based on MS(n). Bioinformatics. 2011; 27:2376–2383. [PubMed: 21757467]

Quintás-Cardama A, Kantarjian H, Cortes J. Homoharringtonine, omacetaxine mepesuccinate, and chronic myeloid leukemia circa 2009. Cancer. 2009; 115:5382–5393. [PubMed: 19739234]

Saleem AM, Dhasan PB, Rafiullah MRM. Simple and rapid method for the isolation of forskolin from *Coleus forskohlii* by charcoal column chromatography. J Chromatogr A. 2006; 1101:313–314. [PubMed: 16325832]

Sana TR, Roark JC, Li X, Waddell K, Fischer SM. Molecular formula and METLIN personal metabolite database matching applied to the identification of compounds generated by LC/TOF-MS. J Biomol Tech. 2008a; 19:259–266.

Sana TR, Waddell K, Fischer SM. A sample extraction and chromatographic strategy for increasing LC/MS detection coverage of the erythrocyte metabolome. J Chromatogr B. 2008b; 871:314–321.

Sawada Y, Nakabayashi R, Yamada Y, Suzuki M, Sato M, Sakata A, Akiyama K, Sakurai T, Matsuda F, Aoki T, Hirai MY, Saito K. RIKEN tandem mass spectral database (ReSpect) for phytochemicals: a plant-specific MS/MS-based data resource and database. Phytochemistry. 2012; 82:38–45. [PubMed: 22867903]

Schaneberg BT, Khan IA. Quantitative analysis of forskolin in *Coleus forskohlii* (Lamiaceae) by reversed-phase liquid chromatography. J AOAC. 2003; 86:467–470.

Shen Y, Zhao R, Belov ME, Conrads TP, Anderson GA, Tang K, Paša-Tolić L, Veenstra TD, Lipton MS, Udseth HR, Smith RD. Packed capillary reversed-phase liquid chromatography with high-performance electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry for proteomics. Anal Chem. 2001; 73:1766–1775. [PubMed: 11338590]

Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, Custodio DE, Abagyan R, Siuzdak G. METLIN: a metabolite mass spectral database. Ther Drug Monit. 2005; 27:747–51. [PubMed: 16404815]

Sorensen CM, Ding J, Zhang Q, Alquier T, Zhao R, Mueller PW, Smith RD, Metz TO. Perturbations in the lipid profile of individuals with newly diagnosed type 1 diabetes mellitus: lipidomics analysis of a Diabetes Antibody Standardization Program sample subset. Clin Biochem. 2010; 43:948–956. [PubMed: 20519132]

Tarasova IA, Guryca V, Pridatchenko ML, Gorshkov AV, Kieffer-Jaquinod S, Evreinov VV, Masselon CD, Gorshkov MV. Standardization of retention time data for AMT tag proteomics database generation. J Chromatogr B. 2009; 877:433–440.

Trescot AM, Datta S, Lee M, Hansen H. Opioid pharmacology. Pain Physician. 2008; 11:S133–S153. [PubMed: 18443637]

Venditto VJ, Simanek EE. Cancer therapies utilizing the camptothecins: a review of the in vivo literature. Mol Pharm. 2010; 7:307–349. [PubMed: 20108971]

Waldia S, Joshi B, Pathak U, Joshi MC. The genus *Plectranthus* in India and its chemistry. Chem Biodiv. 2011; 8:244–252.

Weisenberg SA, Butterfield TR, Fischer SM, Rhee KY. Suitability of silica hydride stationary phase, aqueous normal phase chromatography for untargeted metabolomic profiling of *Enterococcus faecium* and *Staphylococcus aureus*. J Sep Sci. 2009; 32:2262–2265. [PubMed: 19569105]

Wolfender JL. HPLC in natural product analysis: the detection issue. Planta Med. 2009; 75:719–734. [PubMed: 19145552]

Zhang WW, Luo JG, Wang JS, Lu YY, Kong LY. LC-DAD-ESI-MS-MS for characterization of quantitative analysis of diterpenoids from *Coleus forskohlii*. Chromatographia. 2009; 70:1635–1643.

Zhou ZL, Yang YX, Ding J, Li YC, Miao ZH. Triptolide: structural modifications, structure-activity relationships, bioactivities, clinical development and mechanisms. Nat Prod Rev. 2012; 29:457–475.

- The development of an accurate mass–time tag approach for the identification of plant natural products is reported.
- LC/MS and MS/MS data sets were integrated into online spectral search tools and repositories.
- The utility was demonstrated by the detection of active principles in 27 medicinal plant species.

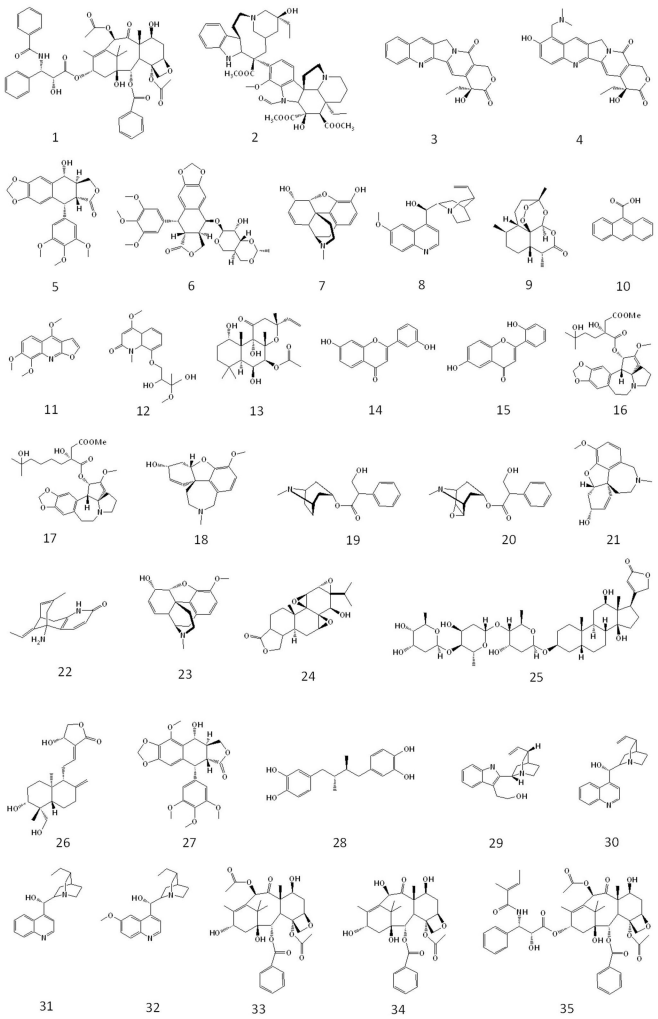

Fig. 1.
Structures of compounds.

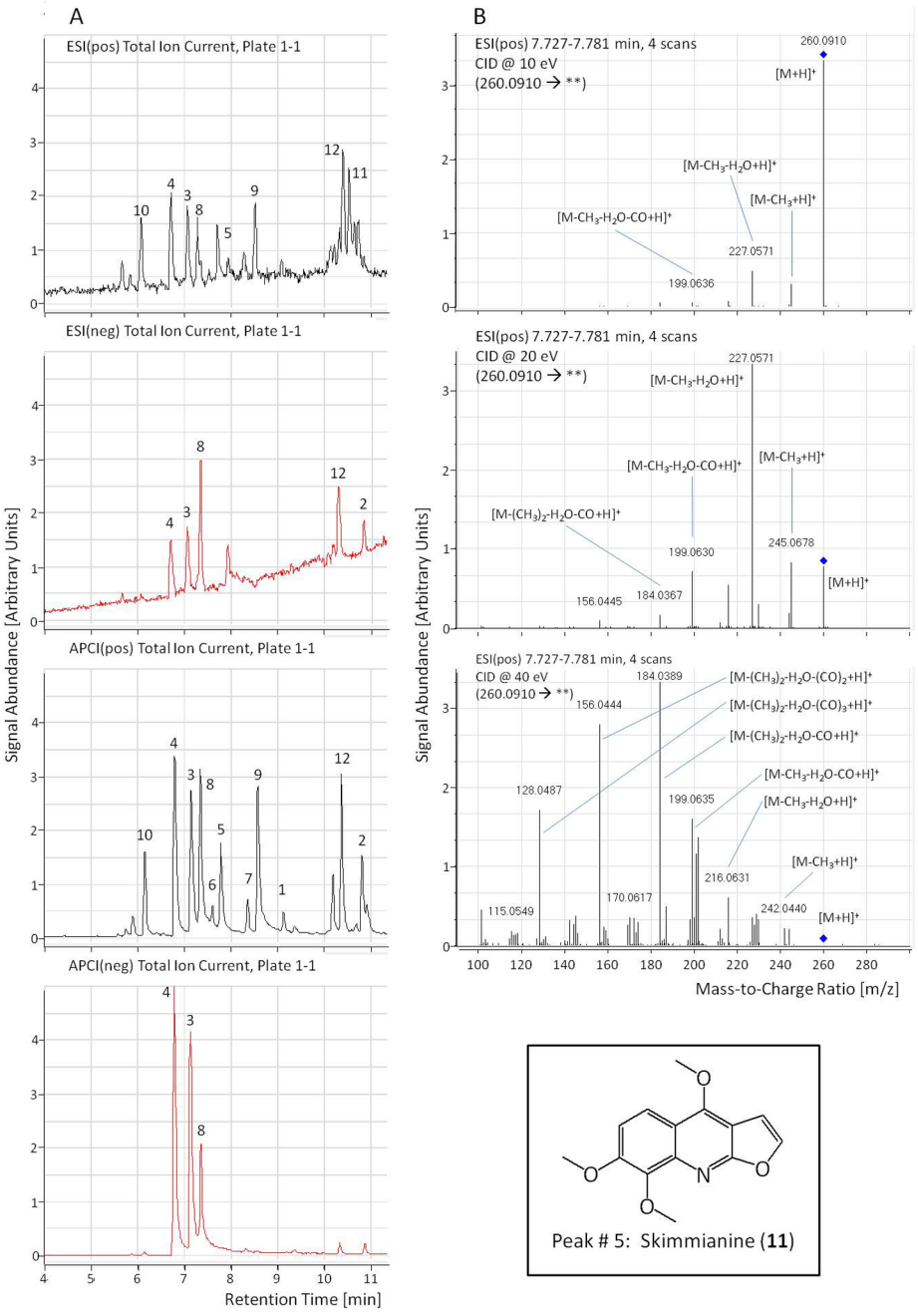

Fig. 2.

LC/MS and MS/MS analysis of plant natural products and detection using different ionization modes and polarities. **A**, Total ion current obtained with a mixture of authentic standards in ESI(pos) (upper panel), ESI(neg) (second panel from top), APCI(pos) (third panel from top) and APCI(neg) (lower panel). The elution order of the compounds in the standard mixture were (note that these are not the compound numbers used in the text): (1) isoalantolactone, (2) (6b-acetyl-13a-hydroxy-4a,6a-dimethyl-13-oxo-1,2,3,4,4b,5,6,7,8,9,10,10a,11,11a,11b,12-hexadecahydroindeno[2,1-a]phenanthren-2-yl) acetate, (3) 2-[1-(1H-benzimidazol-2-yl)-3-(4-methylphenyl)propan-2-yl]-1H-benzimidazole, (4) 2-[1-(1H-benzimidazol-2-yl)-3-(2-methoxyphenyl)propan-2-yl]-1H-benzimidazole, (5)

skimmianine, (6) evoxine, (7) bucharaine, (8) foliosidine, (9) 9-methoxy-2,2-dimethyl-6H-pyrano[3,2-c]quinolin-5-one, (10) methyllycaconitine, (11) 2-piperidin-1-ylethyl 10-acetyloxy-2,4a,6a,6b,9,9,12a-heptamethyl-13-oxo-3,4,5,6,6a,7,8,8a,10,11,12,14b-dodecahydro-1H-picene-2-carboxylate, and (12) [1,3,4a-trimethyl-4,4a,9,9a-tetrahydro-1H-9,2-(epoxymethano)xanthen-2(3H)-yl]methyl phenylcarbamate. **B**, MS/MS spectra of peak # 5 (skimmianine) at different collision energies; 10 eV (upper panel), 20 eV (middle panel); 40 eV (lower panel). An annotation of detected fragments is provided in the MS/MS spectral panels.

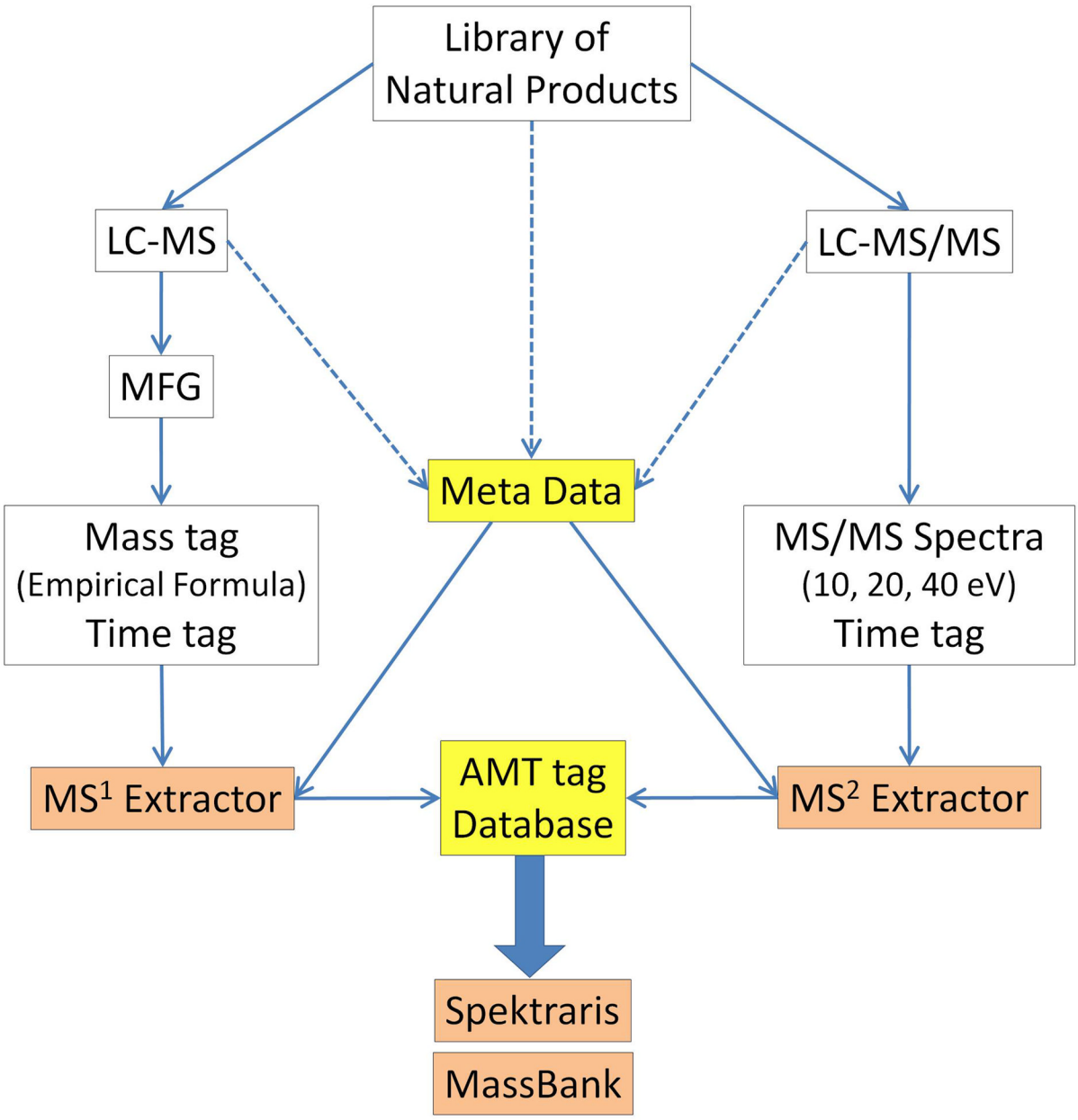

Fig. 3.

Flow diagram for the development of an accurate mass – time tag (AMT tag) database. For details see text. Acronyms: LC, liquid chromatography; MFG, molecular formula generation; MS, mass spectrometry.

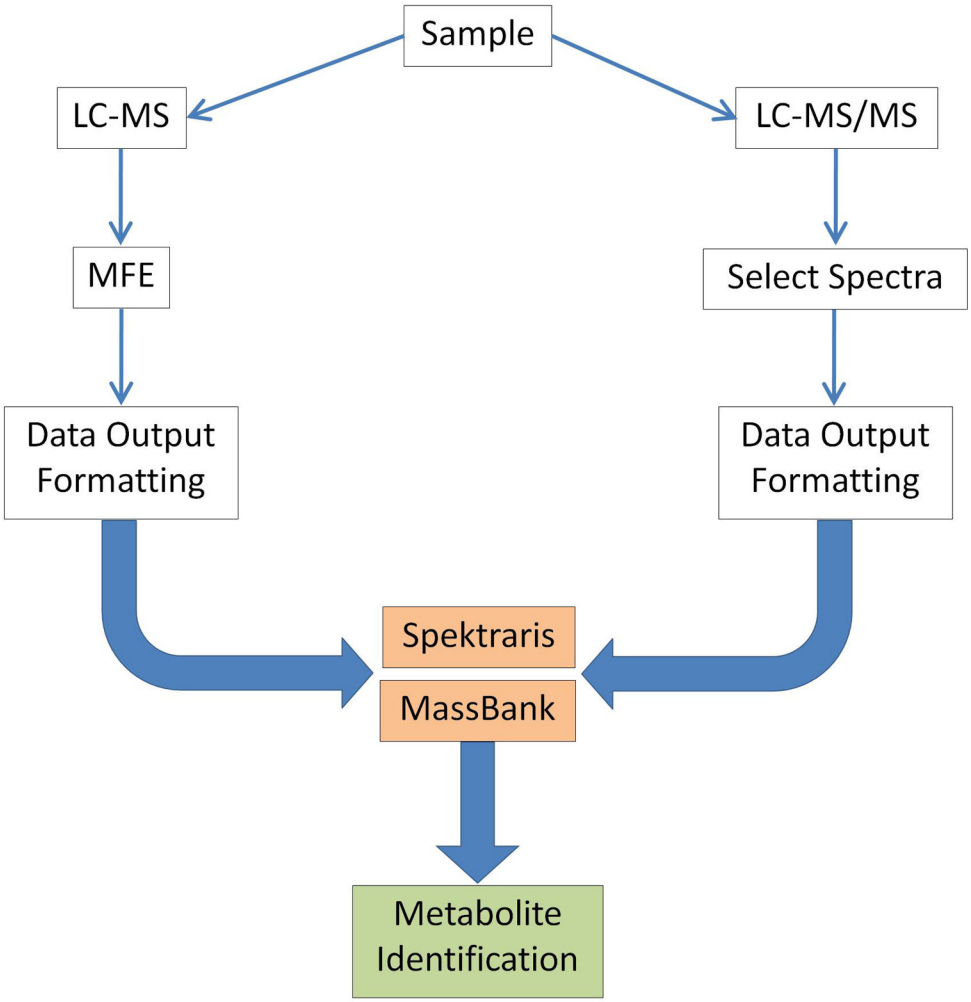

Fig. 4.
Flow diagram for the searching of an accurate mass – time tag (AMT tag) database using Spektraris or MassBank. For details see text. Acronyms: LC, liquid chromatography; MFE, molecular feature extractor; MS, mass spectrometry.

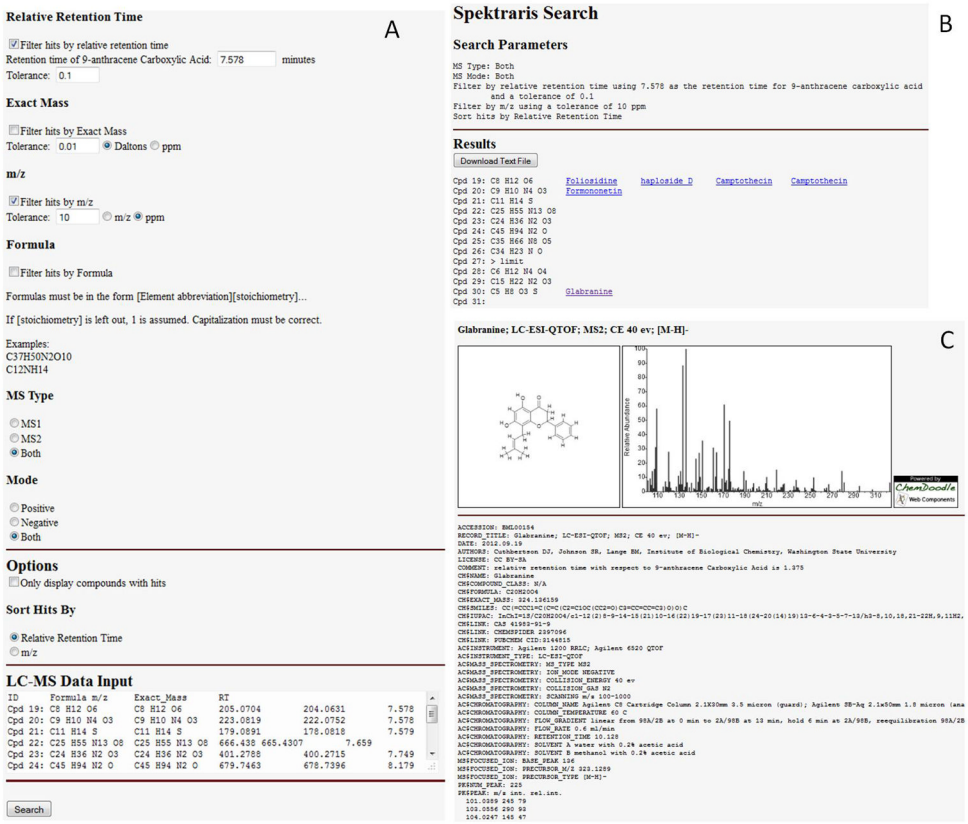

Fig. 5.

Workflow of accurate mass – time tag (AMT tag) database search using the Spektraris tool (http://langelabtools.wsu.edu/amt/). (**A**) Screenshot showing the search options and a sample data set (bottom). Executing the query brings up a new window with a summary of search results (**B**). Clicking on an individual compound hit opens a window with relevant compound information in MassBank format (**C**).

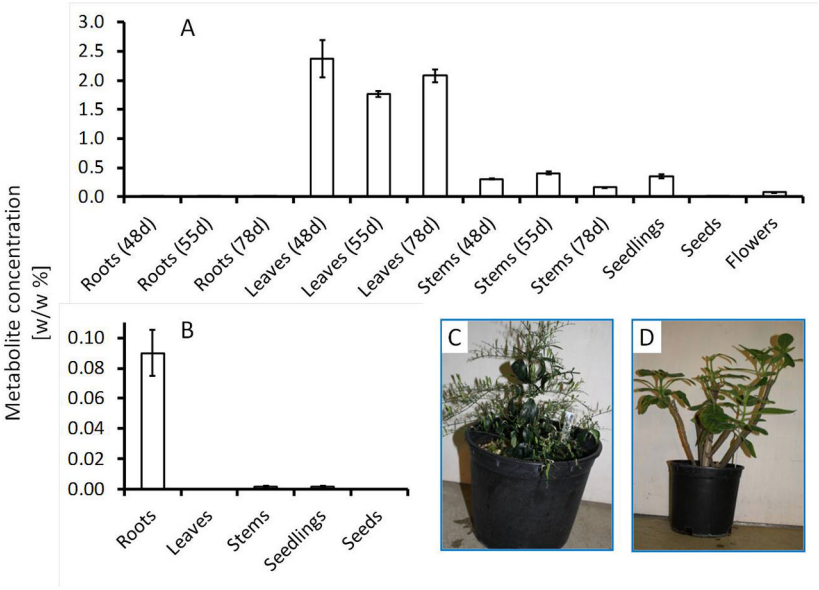

Fig. 6.

Quantitation of (**A**) andrographolide (**26**) in tissue extracts (harvest date shown as days after germination) obtained from *Andrographis paniculata* and (**B**) forskolin (**13**) in tissue extracts obtained from mature *Plectranthus barbatus* plants. Photos of representative plants are provided (**C**, *Andrographis paniculata*; **D**, *Plectranthus barbatus*).

Table 1

Identification of alkaloids in various plant extracts using an accurate mass – time tag approach.

Metabolite	Relative Retention Time	Empirical Formula (Hill Notation)	Mono- Isotopic Mass	m/z (Expected)				Species	Tissue Samples	m/z (Detected in Extracts)*			
				$[M+H]^+$	$[M+Na]^+$	$[M-H]^-$	$[M-Ac]^-$			ESI(+)	ESI(−)	APCI(+)	APCI(−)
Camptothecin (**3**)	1.030	C20H16N2O4	348.1110	349.1183	371.1002	347.1037	407.1243	*Camptotheca acuminata* Decne.	Leaf	+	−	+	−
									Root	+	−	+	−
									Stem	+	−	+	−
Morphine (**7**)	0.055	C17H19NO3	285.1370	286.1443	308.1262	284.1297	344.1508	*Papaver somniferum* L.	Capsule	+	−	+	−
Codeine (**23**)	0.301	C18H21NO3	299.1521	300.1594	322.1413	298.1449	358.1660		Leaf	+	−	+	−
Quinine (**8**)	0.723	C20H24N2O2	324.1838	325.1911	347.1730	323.1765	383.1971	*Cinchona pubescens* Vahl	Bark	+	+	+	−
									Leaf	+	+	+	−
									Root	+	+	+	−
									Xylem	+	+	+	−
Harringtonine (**16**)	0.576	C28H37NO9	531.2468	532.2541	554.2361	530.2396	590.2607	*Cephalotaxus drupacea* Siebold&Zucc.	Leaf	+	−	+	−
Homoharringtonine (**17**)	0.648	C29H39NO9	545.2625	546.2698	568.2517	544.2552	604.2758	*C. harringtonia* ‘Duke Gardens’	Leaf	+	−	+	−
								C. harringtonia ‘Watnong’	Leaf	(+)	−	(+)	−
Colchicine (**18**)	0.981	C22H25NO6	399.1682	400.1755	422.1574	398.1609	458.1815	*Colchicum autumnale* L.	Corm	+	−	+	−
									Fruit	+	−	+	−
									Leaf	+	−	+	−
								Gloriosa superba L.	Fruit	+	−	+	−
									Leaf	+	−	+	−
									Tuber	+	−	+	−
Atropine (**19**)	0.419	C17H23NO3	289.1678	290.1751	312.1570	288.1605	348.1811	*Datura candida* (Pers.) Saff	Leaf	+	−	+	−
Scopolamine (**20**)	0.280	C17H21NO4	303.1471	304.1543	326.1363	302.1398	362.1609	*D. ceratocaula* Ortega	Leaf	+	−	+	−
								D. innoxia Miller	Leaf	+	−	+	−
								D. metel L.	Leaf	+	−	+	−
								D. wrightii Regel	Leaf	+	−	+	−
								Hyoscyamus niger L.	Leaf	+	−	+	−

Metabolite	Relative Retention Time	Empirical Formula (Hill Notation)	Mono- Isotopic Mass	m/z (Expected)				Species	Tissue Samples	m/z (Detected in Extracts)*			
				$[M+H]^+$	$[M+Na]^+$	$[M-H]^-$	$[M-Ac]^-$			ESI(+)	ESI(−)	APCI(+)	APCI(−)
Galanthamine (**21**)	0.260	C17H21NO3	287.1521	288.1594	310.1413	286.1448	346.1654	*Narcissus pseudonarcissus* L.	Bulb	+	−	+	−
									Flower	+	−	+	−
									Leaf	+	−	+	−
Huperzine A (**22**)	0.355	C15H18N2O	242.1419	243.1492	265.1311	241.1346	301.1558	*Huperzia squarrosa* (G.Forst.) Trevis	Sporophyte	+	−	+	−

*A "+" indicates that a (de)protonated molecular ion or adduct of the target was detected. Details are given in Supplementary Table 1.

Table 2

Identification of terpenoids and phenylpropanoids in various plant extracts using an accurate mass – time tag approach.

Metabolite	Relative Retention Time	Empirical Formula (Hill Notation)	Mono- Isotopic Mass	m/z (Expected)				Species	Tissue Samples	m/z (Detected in Extracts)*			
				$[M+H]^+$	$[M+Na]^+$	$[M-H]^-$	$[M-Ac]^-$			ESI(+)	ESI(−)	APCI(+)	APCI(−)
Paclitaxel (taxol®) (**1**)	1.309	C47H51NO14	853.3310	854.3383	876.3202	852.3237	912.3443	*Taxus x media* L.	Cell cult.	+	+	+	+
Forskolin (**13**)	1.190	C22H34O7	410.2305	411.2377	433.2197	409.2232	469.2443	*Plectranthus barbatus* (syn. *Coleus forskohlii*)	Leaf	(+)	(+)	−	−
									Root	+	+	+	+
								Plectranthus caninus Roth (syn. *Coleus caninus*)	Leaf	−	−	−	−
									Root	−	−	−	−
Triptolide (**24**)	0.840	C20H24O6	360.1573	361.1646	383.1465	359.1500	419.1711	*Tripterygium regelii* Sprag. & Takeda	Root	+	+	+	−
Digoxin (**25**)	1.278	C41H64O14	780.4296	781.4369	803.4188	779.4223	839.4429	*Digitalis lanata* Ehrh.	Leaf	−	+	−	+
									Root	−	+	−	+
Andrographolide (**26**)	0.938	C20H30O5	350.2093	351.2166	373.1985	349.2020	409.2226	*Andrographis paniclata* (Burm.f.) Nees	Leaf	+	+	+	+
Podophyllotoxin (**5**)	0.992	C22H22O8	414.1315	415.1388	437.1207	413.1242	473.1448	*Podophyllum hexandrum* Royle	Leaf	+	+	+	+
									Rhizome	+	+	+	+
								Podophyllum peltatum L.	Leaf	+	+	+	+
									Rhizome	+	+	+	+
5-Methoxy- podophyllotoxin (**27**)	1.098	C23H24O9	444.1420	445.1493	467.1313	443.1348	503.1559	*Linum flavum* L.	Flower	+	+	+	+
									Leaf	+	+	+	+
									Root	+	+	+	+
Nordihydroguaiaretic	acid (**28**) 1.074	C18H22O4	302.1518	303.1591	325.1410	301.1445	361.1651	*Larrea tridentata* (DC.) Coville	Leaf	+	+	−	−

*A "+" indicates that a (de)protonated molecular ion or adduct of the target was detected. Details are given in Supplementary Table 1.