Integration of mass spectral fingerprinting analysis with precursor ion (MS1) quantification for the characterisation of botanical extracts: application to extracts of *Centella asiatica* (L.) Urban

Armando Alcazar Magana¹,⁵ | Kirsten Wright² | Ashish Vaswani¹ |
Maya Caruso² | Ralph L. Reed⁴,⁵ | Conner F. Bailey¹ | Thuan Nguyen⁶ |
Nora E. Gray² | Amala Soumyanath² | Joseph Quinn²,³ | Jan F. Stevens⁴,⁵ |
Claudia S. Maier¹,⁵

¹Department of Chemistry, Oregon State University, Corvallis, OR, USA
²Department of Neurology, Oregon Health and Science University, Portland, OR, USA
³Department of Neurology and Parkinson’s Disease Research Education and Clinical Care Center (PADRECC), VA Portland Healthcare System, Portland, OR, USA
⁴Department of Pharmaceutical Sciences, Oregon State University, Corvallis, OR, USA
⁵Linus Pauling Institute, Oregon State University, Corvallis, OR, USA
⁶OHSU-PSU School of Public Health, Oregon Health & Science University, Portland, OR, USA

Correspondence
Claudia S. Maier, Linus Pauling Institute, Oregon State University, 2900 SW Campus Way, Corvallis, OR 97331, USA.
Email: claudia.maier@oregonstate.edu

Funding information
NIH, Grant/Award Numbers: R01AT008099, R61AT009629, S10RR027878, T32 AT002688

Abstract

**Introduction:** The phytochemical composition of plant material governs the bioactivity and potential health benefits as well as the outcomes and reproducibility of laboratory studies and clinical trials.

**Objective:** The objective of this work was to develop an efficient method for the in-depth characterisation of plant extracts and quantification of marker compounds that can be potentially used for subsequent product integrity studies. *Centella asiatica* (L.) Urb., an Ayurvedic herb with potential applications in enhancing mental health and cognitive function, was used as a case study.

**Methods:** A quadrupole time-of-flight analyser in conjunction with an optimised high-performance liquid chromatography (HPLC) separation was used for in-depth untargeted fingerprinting and post-acquisition precursor ion quantification to determine levels of distinct phytochemicals in various *C. asiatica* extracts.

**Results:** We demonstrate the utility of this workflow for the characterisation of extracts of *C. asiatica*. This integrated workflow allowed the identification or tentative identification of 117 compounds, chemically interconnected based on Tanimoto chemical similarity, and the accurate quantification of 24 phytochemicals commonly found in *C. asiatica* extracts.

**Conclusion:** We report a phytochemical analysis method combining liquid chromatography, high resolution mass spectral data acquisition, and post-acquisition interrogation that allows chemical fingerprints of botanicals to be obtained in conjunction with accurate quantification of distinct phytochemicals. The variability in the composition of specialised metabolites across different *C. asiatica* accessions was substantial, demonstrating that detailed characterisation of plant extracts is a...
Plants are extraordinary factories of specialised metabolites, producing more than 200,000 distinct compounds across the plant kingdom.\textsuperscript{1,2} Phytochemicals play a pivotal role in defence, communication, as well as signalling or regulation of primary metabolic pathways. Their large structural diversity is thought to be due to specific adaptations or functions inherited by evolutionary pressure.\textsuperscript{3}

Phytochemicals are the primary source of medicines in many countries.\textsuperscript{4} Their use can reach as much as 80% of people in indigenous populations,\textsuperscript{5} and they are becoming increasingly popular in Western countries.\textsuperscript{6-8} Upon ingestion, the dose-dependent effects of these compounds may have a variety of benefits for human health.\textsuperscript{9} Numerous new reports popularise the use of plant-derived supplements, including phytochemicals, for promoting human health. The industry is responding to this demand by expanding the diversity of plant-derived products and dietary supplements. However, concurrently, concerns regarding the quality, safety and purity of these products have been stimulated by an increasing number of toxicity reports involving plant-derived supplements.\textsuperscript{10}

Botanical supplements are typically complex mixtures of specialised metabolites of which some are biologically active compounds. The phytochemical profiles of botanical preparations can differ significantly from batch to batch since they can be affected by geography, genetics, ontogenetic stage, plant materials used and post-harvest processing methods among others.\textsuperscript{11,12} The stochastic nature of the phytochemical composition influences the biological and pharmacological activity of the product and thus severely impacts the reproducibility of preclinical studies and clinical trials. Many botanical supplements are plant extracts instead of raw materials. In addition to concerns of adulteration and misidentification of raw materials, the phytochemical profiles in extracts depend highly on the extraction methods used.\textsuperscript{13} Thermal and chemical treatments of plant extracts may cause the degradation of phytochemicals.\textsuperscript{14} During the preparation of botanical supplements, composition and levels of phytochemicals may change, which in turn may affect the bioactivity of the preparations.

Unless disease treatment is claimed, the US Food and Drug Administration (FDA) regulates plant extracts as food; therefore, the requirements to demonstrate safety and efficacy are less stringent compared to pharmaceuticals.\textsuperscript{6} Due to enhanced efforts in improving the characterisation of unregulated over-the-counter botanicals, advanced analysis strategies are needed to ascertain authentication and consistency of botanical supplements.\textsuperscript{6,10,15-17}

In the past, the analysis of phytochemical preparations was largely based on thin layer chromatography (TLC) or liquid chromatography (LC) in conjunction with ultraviolet (UV) spectrophotometric or fluorometric detection. Currently, LC methods in combination with mass spectrometry (MS) are considered state-of-the-art for chemical analysis of plant materials and derived products. Modern high resolving MS platforms offer accurate mass measurements in combination with collision-induced dissociation techniques for structural analysis and quantification of compounds in phytochemical preparations. Combining high resolution MS with tandem mass spectrometric techniques allows for establishing mass spectral fingerprints of the phytochemical diversity of plant materials and derived products.

In this proof-of-concept study, we applied a high resolution MS-based workflow in conjunction with post acquisition interrogation for the characterisation of aqueous extracts of the medicinal plant Centella asiatica (CA), a member of the Apiaceae family, which has been used to improve memory and mental health.\textsuperscript{18-21} Recent studies in humans and rodent models are supportive of C. asiatica preparations as complementary medicine to improve memory in ageing-related mild cognitive decline and potentially Alzheimer's disease.\textsuperscript{21-23} Centella asiatica has also been reported to possess other biological activities benefiting human health including anti-inflammatory and immunostimulant properties, promoting wound healing, and ameliorating leprosy, lupus, tuberculosis and gastric ulcers.\textsuperscript{19,21,24-26} The activities of C. asiatica have largely been attributed to its constituent triterpenes, saponins and sapogenins.\textsuperscript{27} Almost 60 compounds belonging to these and other phytochemical classes have been reported in C. asiatica.\textsuperscript{21} For many of these, their role in C. asiatica's biological activity or their mode of action is not known, and many others, as yet unidentified compounds are present which may also contribute to its activity. We report here a new phytochemical analysis workflow that allows both untargeted fingerprinting for determining distinct phytochemicals in various C. asiatica plant extracts, the tentative identification of 117 of these compounds, and simultaneously the accurate, targeted quantification of eight caffeoylquinic acids, seven flavonoids, five hydroxycinnamic acids and four pentacyclic triterpenoids. We report the analytical method's characteristics including limit of detection (LOD), limit of quantification (LOQ), dynamic range and reproducibility for the quantification method. The integrated workflow was applied to the characterisation of aqueous extracts of C. asiatica plant materials from multiple sources.
**EXPERIMENTAL**

### 2.1 Chemicals

LC–MS grade methanol and water were purchased from EMD Millipore (Burlington, MA, USA). Formic acid ACS reagent was from Fisher Chemicals (Hampton, NH, USA). The following certified standard compounds were used: 5-O-cafeoylquinic acid (1), epigallocatechin (2), catechin (3), dihydrocaffeic acid (4), 4-O-cafeoylquinic acid (5), 3-O-cafeoylquinic acid (6), caffeic acid (7), epicatechin (8), 1,5-dicafeoylquinic acid (9), 1,3-dicafeoylquinic acid (10), rutin (11), dihydroferulic acid (12), 3,4-dicafeoylquinic acid (13), 3,5-dicafeoylquinic acid (14), ferulic acid (15), 4,5-dicafeoylquinic acid (16), naringin (17), isoflavic acid (18), quercetin (19), madecassoside (20), asiaticoside (21), kaempferol (22), madecassic acid (23) and asiatic acid (24). Compounds 2, 3, 6, 12, 15 and 18 were from Sigma Aldrich (St Louis, MO, USA); 1, 11, 17 and 19 were from TCI America (Portland, OR, USA); 13 and 16 were from ChromaDex (Irvine, CA, USA) and the remaining compounds were from Toronto Research Chemicals (Canada).

Caffeoylquinic acids are prone to degradation or isomerisation under certain conditions including pH, light exposure, and temperature. To protect compounds from degradation, all standards and samples were prepared in methanolic solutions containing 0.1% v/v formic acid and kept in the dark at −20°C until analysis.

### 2.2 Plant materials and preparation of aqueous extracts of *Centella asiatica*

The identity of the plant materials was confirmed at the supplier (Oregon’s Wild Harvest) by organoleptic analysis and Fourier-transform infrared analysis (FTIR). Identity was further verified at Oregon Health and Science University by thin-layer chromatographic comparison of zone profiles with earlier batches of *C. asiatica*, and with reference standards of characteristic triterpenes (asiatic acid, madecassic acid, asiaticoside, madecassoside) as well as caffeoylquinic acids known to be found in *C. asiatica*. Voucher samples of all plant materials have been deposited in the Herbarium at Oregon State University. Voucher numbers for plant materials from Oregon’s Wild Harvest are CA1 (OSC-V-258635), CA2 (OSC-V-258632), CA3 (OSC-V-258631), CA4 (OSC-V-258629), CA5 (OSC-V-258627), CA6 (OSC-V-258630), CA7 (OSC-V-258634), and CA8 (OSC-V-258633). The preparation of the *C. asiatica* water extracts was reported previously. In brief, dried *C. asiatica* extracts were prepared by refluxing aerial parts of the plant (stems, leaves and flowers but not roots; 80 grams per 1 litre of deionised water) for 1.5 h, cooling for 30 min to allow for handling, and then filtering the suspension to remove plant debris. The aqueous extracts were freeze-dried and stored at −20°C and analysed within 2 months of preparation. This method was modified from earlier work by Veerendra et al. who showed that exhaustive water extraction of *C. asiatica* produced a residue with greater cognitive enhancing properties than extracts made with methanol or chloroform.

For quantification of the individual compounds in dried *C. asiatica* extracts, a stock solution was prepared as follows. Briefly, 10 mg of each freeze-dried extract powder was resuspended in 10 mL of aqueous methanol (70% v/v with 0.1% v/v of formic acid) by sonication (30 min, 25°C; see Supporting Information Figure S9), centrifuged (14000 × g for 10 min) and filtered with 0.22-µm polyvinylidene fluoride (PVDF) Whatman filters before analysis. This procedure was used to prepare extracts from eight different accessions of the plant materials labelled from CA1 to CA8. Aliquots of 1 mL from each extract were pooled to generate a quality control sample (QC) used for evaluating LC–MS/MS platform performance.

### 2.3 Fingerprinting of *Centella asiatica* extracts by untargeted data-dependent analysis

For the chemical profiling analyses, a pooled CA sample (QC sample) was used. Untargeted high-performance liquid chromatography (HPLC) combined with high resolution accurate LC–MS/MS was conducted using a Shimadzu Nexera UHPLC system connected to an AB SCIEX TripleTOF® 5600 mass spectrometer equipped with a Turbo V ionisation source operated in the electrospray ionisation (ESI) mode.

Chromatographic separation was achieved using an Inertsil Phenyl-3 column (4.6 mm × 150 mm, 100 Å; GL Sciences, Torrance, CA, USA). The injection volume was 10 µL. Three technical replicates were conducted. Gradient elution was performed using a mobile phase consisting of solvent A, water containing 0.1% v/v formic acid, and solvent B, methanol containing 0.1% v/v formic acid. Flow rate was 0.4 mL/min. The chromatographic method was 30 min, and the gradient design was as follows: an initial 1 min at 5% B, followed by 5 to 30% B from 1 to 10 min, then 30 to 100% B from 10 to 20 min, hold at 100% B from 20 to 25 min, and then return to 5% B from 25 to 30 min.

Data-dependent acquisitions (DDAs) were conducted for obtaining precursor and fragment ion information for aiding in annotating compounds in the CA extracts. DDA analyses were conducted using the negative (ESI–) and positive ionisation (ESI+) mode. For detecting negative ions, the following parameter settings were used to operate the mass spectrometer: spray voltage −4200 V; source temperature 550°C and a period cycle time of 950 ms was used. The following settings were used: full scan with ion accumulation of 150 ms, followed by a dynamic MS/MS selection of the eight most intense ions with 100 ms accumulation; after two MS/MS acquisitions the precursor (fragmented) ions were excluded for 30 s; collision energy 35 V with collision energy spread (CES) of 15 V ramped through each MS/MS scan using a range of m/z 100–1200. For ESI+ acquisitions, the instrument settings were the same as used in the negative ion mode except that the spray voltage was set to 4500 V. The mass spectrometer was equipped with a calibrant delivery system. Mass calibration was automatically performed after every fifth LC run.
2.4 | Data processing and annotation of plant metabolites

Annotation confidence was established according to reporting criteria for chemical analysis suggested by the Metabolomics Standards Initiative.\(^{35,36}\) For Level 1 (L1) annotations (Tables 1 and 2) we used accurate mass, fragment ion spectral similarity and retention time (RT) co-elution based on an authentic commercially available standards. Standard addition experiments were also conducted as part of developing the quantitative method (see later).

In Figure S1 the workflow is described to obtain in putative annotations [Level 2 (L2) annotations\(^{35,36}\)] based on exact mass, isotopic pattern and MS/MS spectral data. In addition, three manual data evaluations were included: (1) examination of the metabolite structures with respect to the suitability of the ionisation mode in which a compound was detected (i.e. basic sites in a molecule that can be protonated in ESI+ or labile protons for ESI–), (2) elution peaks for tentatively annotated features were interrogated to omit compounds originating from in-source fragmentation, and (3) only compounds previously reported in plants were kept and listed in Table 1. It is important to note, that these compounds are tentative annotations only in accordance to the guidelines described by Sumner et al.\(^{35}\) and will need to be further confirmed in future work.

Raw data processing was performed using Progenesis QI™ software with METLIN™ plugin V1.0.6499.51447 (NonLinear Dynamics, Newcastle Upon Tyne, UK) and entailed peak picking, alignment and searching of multiple databases to assist in compound annotations. For the current study, we searched the mass spectral data against METLIN,\(^{37}\) Human Metabolome Database (HMDB),\(^{38}\) Chemical Entities of Biological Interest (ChEBI)\(^{39}\) (online versions, April 2018) and an in-house compound library based on the Mass Spectrometry Metabolite Library of Standards (MSMLS) consisting of 619 standards (IROA Technologies, Bolton, MA, USA) and other commercially available standards including the 24 compounds used as phytochemical markers in this study (650 total).\(^{40}\) Progenesis QI™ uses as built-in search engine Metascope and provides a "score" for the quality of the compound annotation, using a range from 0 to 100, with 100 being a perfect match based on the mean of multiple similarity metrics.\(^{41}\) The current data were evaluated based on the accurate mass similarity, isotope similarity, and fragmentation score (ranging from 0 to 60 representing how well the observed data matches the spectral library entries or the theoretical fragment data based on the bond dissociation approach which is a computational method that calculates expected fragments based on theoretically derived bond dissociation energies\(^{42}\)). Progenesis QI’s fragmentation algorithm is described by Wolf et al.\(^{43}\) and Horai et al.\(^{44}\) A Progenesis QI score ≥ 50 is typically reached when isotopic pattern similarity is above 90%, MS/MS spectral data similarity is > 50% and the deviation of the accurate mass from the exact mass is lower than 5 ppm. Progenesis QI score ≥ 50 was considered as adequate for being considered as a candidate for putative annotation (L2 annotations according to Sumner et al.\(^{35}\)). This score is more rigorous than previous reports using Progenesis QI™ with a score > 31.6,\(^{44}\) putative annotations based only on accurate mass and isotope similarity or with a mass error of 20 ppm.\(^{46}\) Additional features were assigned by querying and comparison with KNApSAcK online library.\(^{47}\) Supporting Information Table S1 lists identified (L1) and putatively assigned (but unverified) metabolites (L2 annotations), and provides access to the following properties: RT, monoisotopic ion mass, ions observed and molecular formula. Figure S7 compiles positive matches (red lines) with the entries in the respective spectral libraries. In the case of structural isomers, the best match (highest score) against the MS/MS spectral data was selected. To illustrate chromatographic performance, we used extracted ion chromatograms (XICs) of 22 annotated ions detected in the positive ion mode (Figure S2) and 24 detected using the negative ion mode (Figure S3).

2.5 | Chemical similarity network and clustering

We built a chemical space network based on the compounds listed in Tables 1 and S1. In this network, terminal nodes represent compounds and edges (branches) identify similarity relationships based on two-dimensional (2D) chemical structures. We used the Tanimoto algorithm, \(T(A,B) = A \cap B / A \cup B\), often referred to as intersection over union, with \(A\) and \(B\) representing structures of molecule \(A\) and molecule \(B\), for calculating measures of similarity. For this purpose, we used PubChem Score Matrix Service V1.3 according to Sungwhan Kim 2016.\(^{48}\) The PubChem server uses the simplified molecular input line-entry system (SMILES) identifiers to compute Tanimoto coefficients and then creates edges between similar structures if the coefficients are greater than or equal to the set threshold value (0.68). The derived Tanimoto coefficient represents an associative coefficient with a value ranging from 0 to 1, numerically expressing the structural similarity between a 2D binary comparison (0 being no similarity and 1 being complete similarity).\(^{49,50}\) A Tanimoto coefficient greater or equal to 0.68 indicates that the compounds being compared are structurally similar and statistically significant at the 95% confidence interval.\(^{51}\) Tanimoto coefficients were exported to Cytoscape (V3.6.1) for graphic visualisation and a 2D structural similarity network was created for the compounds listed in Table 1.

2.6 | Method development for quantification of selected phytochemicals in extracts

A quantification method was developed for 24 compounds (Figure S4). The method uses the same chromatographic conditions as described in the untargeted analysis section and the mass spectrometer was operated in ESI+ mode using the following settings: spray voltage –4200 V; source temperature 550°C; period cycle time 950 ms; precursor ions accumulation time 100 ms; scan range \(m/z\) 100–1200 (as described earlier).

External calibration curves were acquired for the 24 authentic compounds based on using the area under the curve of the precursor ion (MS1-based). Solutions containing analytical blanks, 0.005, 0.01,
**Table 1** Summary of detected compounds in *Centella asiatica* extracts (pooled sample) by extensive querying and comparison with spectral data (METLIN, HMDB, ChEBI and our in-house library) and compound libraries (KNApSAcK and PantMAT) using Progenesis QI™ and applying the workflow shown in Figure S1. Compounds are labelled with their respective PubChem CID. Additional parameters are shown in Table S1.

Categories were assigned according to structural similarity using Tanimoto algorithm, and they may correspond to more than one compound class. Compounds confirmed using authentic standards are shown in bold [Level 1 (L1) identifications]; all other compound assignments are based on Level 2 annotations (MS/MS spectral matches are presented in Figure S7). Eighty-seven compounds that were detected for the first time in *C. asiatica* extracts are denoted with an asterisk **

| Compound | CID | InChIKey | Molecular formula |
|----------|-----|----------|-------------------|
| **Amino acid derivatives** | | | |
| 2-Pyrrolidone-5-carboxylic acid | 499 | ODHCTXKNWHXJC-UHFFFAOYSA-N | C₅H₇NO₃ |
| 1-beta-D-Glucopyranosyl-L-tryptophan* | 11772967 | ZHBH2ZMTVJASV-JOSVURMMS-N | C₁₇H₂₂N₂O₇ |
| 2,6-Piperidinedicarboxylic acid* | 557515 | SOOPBZRXJMNXTF-UHFFFAOYSA-N | C₇H₁₁NO₄ |
| 4-Guainidobutanoic acid* | 25200642 | TUHVEAXIMEOSA-UHFFFAOYSA-N | C₉H₁₆N₂O₃ |
| 5-Methoxy-L-tryptophan* | 557515 | SOOPBZRXJMNXTF-UHFFFAOYSA-N | C₁₂H₁₄N₂O₃ |
| 6-amino-9H-purine-9-propanoic acid* | 255450 | QXAYJKFBMWMARF-UHFFFAOYSA-N | C₉H₁₄N₂O₃ |
| 6-Oxo-2-piperidinecarboxylic acid* | 3014237 | FZXCPFJMYOQZCA-UHFFFAOYSA-N | C₆H₉NO₃ |
| L-Arginine* | 28782 | ODKSFYDXXFIFQN-BYPYZUCNSA-N | C₆H₁₄N₄O₂ |
| N-(1-Deoxy-1-fructosyl)phenylalanine* | 101039148 | FAVRCPNIVJIPN-VJDSNFAGSA-N | C₁₅H₂₁NO₇ |
| N-Acetyl-L-glutamic acid* | 938 | PVNIIIVLPJYAGWF-UHFFFAOYSA-N | C₆H₁₄N₄O₂ |
| Niacin (nicotinic acid)* | 993 | PAVNIIIVLPJYAGWF-UHFFFAOYSA-N | C₆H₁₄N₄O₂ |
| Pantothenic acid | 6613 | GPOKWTGJZEAQD-ZETCQYHSA-N | C₁₂H₁₈N₄O₆ |
| Succinyl-L-proline* | 194156 | NEBOPDYAEPDYH-LURUJMIEA-N | C₁₀H₁₄N₄O₄ |
| Vincosamide* | 10163855 | LBRPPLCNZUXLS-AZVRDJBXSA-N | C₂₆H₃₀N₂O₈ |
| **Amino sugar derivatives** | | | |
| Enicoflavine* | 5281564 | GBJQPSBGSKNYHV-YVMONPNESA-N | C₁₀H₁₃NO₄ |
| Muramic acid* | 433580 | MSFSPUZLOGKHJ-UHFFFAOYSA-N | C₁₀H₁₅N₄O₇ |
| N-Acetyl-D-glucosamine* | 899 | OVRNRDQMRJTHS-UHFFFAOYSA-N | C₁₀H₁₄N₄O₇ |
| Soyacerebroside I* | 131751281 | HOMYIIYLJRTDK-AUYIJDJPSA-N | C₁₇H₂₂N₄O₉ |
| o-1-[(3-Carboxypropyl)amino]-1-deoxyfructose* | 131752417 | HUEOABWGBTQXNF-SFJDOBXSA-N | C₁₆H₂₀N₂O₇ |
| **Choline derivatives** | | | |
| Betaine* | 247 | KWIUHFFTVRNATP-UHFFFAOYSA-N | C₁₀H₁₄N₂O |
| Choline* | 305 | OEUIOHPDSNJKLS-UHFFFAOYSA-N | C₁₀H₁₄N₂O |
| Choline O-sulphate* | 486 | WXCQAWGWVXRGP-UHFFFAOYSA-O | C₁₀H₁₄N₄O₅S |
| Phosphocholine | 1014 | YHHSOZNZFOIEMCP-UHFFFAOYSA-O | C₁₀H₁₄N₄O₅P |
| **Di-caffeoylquinic acids** | | | |
| 1,3-Di-caffeoylquinic acid | 6474640 | YDDUMTOHZQPO-PSETPKNSA-N | C₂₆H₃₀O₁₂ |
| 1,4-Di-caffeoylquinic acid* | 12358846 | IYQXRCQXWUFQV-RDJKMHVSNA-N | C₂₆H₃₀O₁₂ |
| 1,5-Di-caffeoylquinic acid | 5281769 | YDDUMTOHZQPO-RDJKMHVSNA-N | C₂₆H₃₀O₁₂ |
| 3,4-Di-caffeoylquinic acid | 5281780 | UFCLZKMFSLNL-PSETPKNSA-N | C₂₆H₃₀O₁₂ |
| 3,5-Di-caffeoylquinic acid | 6474310 | KRZBCHWVBQONTZ-RDJKMVHSNA-N | C₂₆H₃₀O₁₂ |
| 4,5-Di-caffeoylquinic acid | 6474309 | UFCLZKMFSLNL-RDJKMVHSNA-N | C₂₆H₃₀O₁₂ |
| **Fatty acid derivatives** | | | |
| Caprylic acid* | 379 | WWZKQHOICKIZLMA-UHFFFAOYSA-N | C₈H₁₆O₂ |
| Palmitic acid* | 985 | IPCSVZSZVZIGE-UHFFFAOYSA-N | C₁₆H₃₂O₂ |
| Tetradecanedioic acid* | 13185 | HQHCKZLJHKEB-UHFFFAOYSA-N | C₁₆H₃₂O₄ |
| 16-Hydroxypalmitic acid* | 10466 | UGAGPNCDRTDHP-UHFFFAOYSA-N | C₁₆H₃₂O₃ |
| Traumatic acid* | 5283028 | MAZWDMBCDUFUFWQVHLKOHSA-N | C₁₆H₃₂O₄ |
| 12-Oxodihydrophytodienoic acid* | 5716902 | BZXDFDKIZJBP-GTOOHNYSNA-N | C₁₈H₃₄O₃ |
| Nomilinic acid 17-glucoside* | 444212 | MUZNCCNBAPYIF-UWHTYGYSNA-N | C₄₄H₆₈O₁₆ |
| Compound | CID | InChIKey | Molecular formula |
|----------|-----|----------|-------------------|
| 3,5-Dihydroxyphenyl 1-O-(6-O-galloyl-β-D-glucopyranoside) | 131752603 | WVHDGXKUZIDTIN-UHFFFAOYSA-N | C_{19}H_{20}O_{12} |
| 6-C-α-L-Arabinosyl-8-C-β-L-arabinosylapigenin | 122391238 | LDVNKZYMPZDAI-DKEQBBAAAS-N | C_{29}H_{26}O_{13} |
| Apimaysin* | 101920411 | LCQVQAZLYBMGJ-YWOWJUJADAS-N | C_{29}H_{26}O_{13} |
| Astragalin | 5282120 | JPUKWEQWBDQOB-QSOFNLRSA-N | C_{23}H_{22}O_{11} |
| Kaempferol | 5280863 | IYRMMWYSZQPKJC-UHFFFAOYSA-N | C_{19}H_{16}O_{6} |
| Mangiferin* | 5281647 | AEEDIBAIWPIBD-ZIKJAXJQBANS-N | C_{23}H_{22}O_{11} |
| Naringin | 442428 | DFPSMGMNTNNDN-HZPHTFPEAS-N | C_{29}H_{26}O_{14} |
| Pelargonidin 3-O-glucoside* | 443648 | ABVCUBUKXJYSE-GQPQBVGS-N | C_{21}H_{21}O_{10}+ |
| Quercetin | 5280343 | REFJWTPEDVJJI-UHFFFAOYSA-N | C_{19}H_{16}O_{7} |
| Quercetin 3-(6′-acetylglucoside)* | 44259187 | IGLUNMMDNDNWZOA-CHUITGBKSA-N | C_{23}H_{22}O_{13} |
| Quercetin 3-O-glucoside* | 5280804 | OVSQVDMCBVZGQ-QSOFNLRSA-N | C_{23}H_{22}O_{12} |
| Rutin | 5280805 | IKGXBQEEMLURG-NVNPHEPSANA-N | C_{20}H_{16}O_{16} |
| Glabraoside A* | 102393599 | OKKMHZTVGQAGP-FVLETCQYSA-N | C_{30}H_{20}O_{13} |
| Hexoside derivatives | | | |
| Stachyose* | 439531 | UQZIYBHXAGNOE-XNSRJBNMNAS-N | C_{22}H_{22}O_{21} |
| Carlusic acid methyl ester* | 122391261 | MQWPZWMPWXSL-ZETCQYMNAS-N | C_{13}H_{12}O_{6} |
| Daucic acid* | 5316316 | KUKCROTFBRNU-UHFFFAOYSA-N | C_{7}H_{8}O_{2} |
| Digalacturonate* | 439694 | IGSYEZFPZOFNC-LIKWRGPLSA-N | C_{13}H_{14}O_{13} |
| Dihydroactinidiolide* | 27209 | IMKHDGBJNLREDU-EHFFFAOYSA-N | C_{11}H_{12}O_{2} |
| Furanol 4-(6-malonylglucoside)* | 131750900 | QJYOBEMAMLWZTF-UHFFFAOYSA-N | C_{12}H_{16}O_{11} |
| Isovalerylglucuronide* | 137383 | VOJALAAOYUSSCT-ZCLKEDUABS-N | C_{11}H_{14}O_{8} |
| Linustatin* | 119301 | FERSMFQBWVBKQT-CXTTVELSAS-N | C_{15}H_{27}NO_{11} |
| Purgic acid B* | 16091605 | YTQXXUYELKIK-GLYQDNRSA-N | C_{21}H_{26}O_{29} |
| Hydroxybenzonic acids | | | |
| Caffeic acid* | 689043 | QAIPRVGONGVQAS-DUXPYHPUSA-N | C_{7}H_{8}O_{4} |
| Isoferulic acid* | 736186 | QRCVMIEKCOAJU-HWKANZROSA-N | C_{10}H_{10}O_{4} |
| Dihydrocaffeic acid* | 348154 | DZAUWHJDUNRCTF-UHFFFAOYSA-N | C_{9}H_{10}O_{4} |
| Dihydroferulic acid* | 14340 | BOLJTPHS5ZHR-UHFFFAOYSA-N | C_{10}H_{12}O_{4} |
| Ferulic acid | 445858 | KSEBMYQBYZTDHS-HWKANZROSA-N | C_{10}H_{12}O_{4} |
| Monocaffeoylquinic acids | | | |
| 3-O-Caffeoylquinic acid | 1794427 | CWVRJTMFETXNAS-UXMJACGLSA-N | C_{13}H_{18}O_{9} |
| 4-O-Caffeoylquinic acid | 9798666 | GYFFKZTYAFCTR-AVXJPLUSA-N | C_{13}H_{18}O_{9} |
| 5-O-Caffeoylquinic acid | 5280633 | CWVRJTMFETXNAS-NXLLHMKUSA-N | C_{13}H_{18}O_{9} |
| Organic acids | | | |
| Citric acid* | 19782904 | KRKNYBCHKNGOX-UHFFFAOYSA-N | C_{6}H_{8}O_{7} |
| L-Ribulose* | 644111 | ZAQJHHRNXZUBTE-UCORVYFPSA-N | C_{5}H_{10}O_{5} |
| Malate* | 20130941 | BJEPKYJPYRNK-OHFFFAOYSA-N | C_{4}H_{8}O_{5} |
| Succinate* | 1110 | KDYPGRWQOYBRFD-UHFFFAOYSA-N | C_{4}H_{8}O_{4} |
| Phenolic compounds | | | |
| 3-Hydroxy-2-oxo-3-phenylpropanoic acid | 71581094 | ZHLWCBBWULYSF-ZETCQYMNAS-N | C_{10}H_{8}O_{4} |
| 1-Caffeoyl-5-feruloylquinic acid* | 121225501 | DXJURPUTSYEZSV-BCYRLKMSA-N | C_{18}H_{20}O_{12} |
| 3,4-Dihydroxybenzaldehyde* | 8768 | IBGBRVPALMCC-UHFFFAOYSA-N | C_{12}H_{10}O_{3} |
| 3,5-Dihydroxy-2-methylphenyl beta-D-glucopyranoside* | 46184089 | AXTQBXDFUAFPD-UIPOAAUSA-N | C_{13}H_{18}O_{8} |
| 3-Hydroxycoumarin | 13650 | MJKVTMPWOKAWS-UHFFFAOYSA-N | C_{8}H_{6}O_{3} |
| 4-Hydroxybenzaldehyde* | 126 | RGHHSNMVTDUWLL-UHFFFAOYSA-N | C_{7}H_{6}O_{2} |
| Compound                                            | CID     | InChIKey                                  | Molecular formula |
|-----------------------------------------------------|---------|-------------------------------------------|-------------------|
| 5-Methoxysalicylic acid*                            | 75787   | IZZIWIAYZOBULF-UHFFFAOYSA-N               | C_8H_8O_4         |
| 8-Acetoxyl-4'-methoxypinoresinol 4-glucoside*       | 73830447| ZKCRENDTQINGLGO-UHFFFAOYSA-N             | C_29H_36O_13      |
| Aesculin*                                           | 5281417 | XHCADHNYFIFUHF-TVKYYDDYS-A                | C_13H_20O_4       |
| Catechin                                            | 9064    | PTFAWBLQPZVEMU-DZGCQCFKS-A                | C_13H_14O_6       |
| Coumarin                                            | 323     | ZYGHJZDHTUPRJ-UHFFFAOYSA-N               | C_10H_8O_2        |
| Epicatechin                                         | 72276   | PTFAWBLQPZVEMU-UKRROHQQ-S                 | C_13H_20O_8       |
| Epigallocatechin*                                   | 72277   | XMOSCLCSDHDWHP-IUODEOFL-RA                | C_13H_14O_7       |
| Folinic acid*                                       | 6006    | VIAGPDDUFNURO-ABBLWVSNPS-A               | C_20H_23N_7O_7    |
| Ginkgoic acid*                                      | 5281858 | YXHVCCZLLWZIYSA-FPLPWNLS-A               | C_20H_20O_3       |
| Kuwanon Y*                                          | 1433407 | YYUHPJIKWHNMSV-JOAOHLMBS-A               | C_40H_30O_9       |
| Kynurenic acid*                                     | 3845    | HCHZHEIFKROPDI-UHFFAOYSA-N               | C_10H_10O_3       |
| N1,N5,N10,N14-tetra-trans-p-Coumaroyslpermine*     | 9810941 | KKJYIHSHTUGJL-PBRJPQHQSNA-N              | C_40H_30N_4O_8    |
| Phlorin*                                            | 476785  | WXTPOHTGNYFSD-RMPHPYRLSA-N               | C_13H_20O_8       |
| Tropic acid*                                        | 10726   | JACRWUPXAESPB-UHFFAOYSA-N                | C_10H_20O_3       |
| Xanthurenic acid*                                   | 5699    | FBZONXHGGPHHII-UHFFAOYSA-N               | C_10H_10N_2O_6    |
| Purine derivatives                                   |         |                                           |                   |
| 2'-O-Methyladenosine*                               | 102213  | FPUGCISOXNNP-C1OSLPPCCSA-N                | C_11H_15N_5O_4    |
| Adenine*                                            | 190     | GGGGBOXGBJISGV-UHFFAOYSA-N               | C_10H_15N_5O_4    |
| Adenosine*                                          | 60961   | OIRDTQYFTABQQP-KQYNNXS-U                 | C_10H_15N_5O_4    |
| cAMP*                                               | 6076    | IVOMOWUHPKRL-KQYNNXS-A                   | C_10H_15N_5O_4    |
| 5'-Deoxy-5'-[methylsulfanyl]adenosine*              | 165114  | WXQJULRQRWHRMGT-JLGSQGANS-A              | C_10H_15N_5O_4    |
| Succinoadenosine*                                   | 126969142| VKGZCEJTCPKMLR-PWBTYLA-S                 | C_10H_15N_5O_4    |
| Guanosine*                                          | 6802    | NYHBFQYNKIQIF-UOKFHMZSA-N                | C_10H_15N_5O_4    |
| Terpenoids                                           |         |                                           |                   |
| 26-(2-Glucosyl-6-acetylglucosyl)-1,3,11,22-tetrahydrxyergosta-5,24-dien-26-oate* | 131752817| CRRPFTZRFACDM-ZCUXNETKSA-N               | C_24H_42O_17      |
| Asiatic acid                                        | 119034  | JXSVIVRWWRRQRT-UYDOISQ-A                 | C_26H_30O_5       |
| Asiaticoside                                        | 24721205| WYQVAPQRGARQUB-HCJNVKSA-A                | C_20H_20O_5       |
| Dysolenticin B*                                     | 56601655| RPPAVMFOBDKIDO-HALKBXBS-A                | C_20H_20O_5       |
| Gentionicoside*                                     | 88708   | DUAGQYUORDTXOR-GPQXQLAS-A                | C_14H_20O_9       |
| Madecassic acid                                      | 73412   | PRAUHVZPJOEIC-AOLYGAPSA-N                | C_20H_20O_6       |
| Madecassoside                                       | 91885295| BNMGLJRUUDDLHWR-RSQPUDYA-A               | C_20H_20O_5       |
| Sambacin*                                           | 131752486| SIVWXQOSOMQCC-PANGBKQDHS-A               | C_16H_20O_12      |
| Swertiamarin*                                       | 442435  | HEYZWPRKKGDC-QBXMVCASA-N                 | C_10H_18O_10      |
| Tsangane L 3-glucoside*                             | 73981648| URMJRTXXZFGB-UHFFAOYSA-N                 | C_13H_20O_7       |
| b-Chlorogenin 3-[4'-[2'-glucosyl-3'-xylosyl(glucosyl)galactoside]   | 74193143| WRRAISMNVHHXH-UHFFAOYSA-N               | C_10H_18O_23      |
| Shanzhise*                                          | 11948668| YSIFYNVXOGADM-KDYWQABDSN-A               | C_14H_20O_11      |
| Others                                              |         |                                           |                   |
| Cytosine*                                           | 597     | OPTASPRLRRNAP-UHFFAOYSA-N                | C_8H_8N_3O         |
| Longicamphenylone*                                  | 91747202| VMMYUHUQMAMNC-UHFFAOYSA-N                | C_12H_20O_6       |
| Longifolenedaldehyde*                               | 5655584 | PBMTGOFWRRJFS-UHFFAOYSA-N                | C_12H_20O_6       |
| Uric acid*                                          | 1175    | LEHOTFKMEONL-UHFFAOYSA-N                 | C_6H_4N_3O         |
| 6-Docosanamide*                                     | 44584605| COUPYRANTUKY-MUUIHNZSA-N                 | C_20H_20N_2O       |
| Deoxyfructosazine*                                  | 73452   | FBDICDJCVXZLIP-VSSNQEPJSA-N              | C_20H_20N_2O       |
| Ginsenosyne K*                                      | 15736266| SYNNBWELEYGBFQT-NTCAYCPXSA-N             | C_17H_22O_3        |
0.05, 0.10, 0.50, 1.00, 5 and 10 mg/L of all compounds were prepared in 70% v/v methanol containing 0.1% v/v of formic acid. For quantification, SCIEX MultiQuant™ V3.0.2 analysis software was used, calculating the peak areas under the curve for precursor ions.

### 2.6.1 Accuracy and recovery experiments

To test the accuracy of the method using precursor ions, three standard mixtures of known concentrations (low, 0.05 mg/L; medium, 0.50 mg/L; high, 5.00 mg/L) were evaluated. Standard addition of authentic standards was performed, and recovery experiments were conducted for CA extracts using precursor ions. Quality control samples were spiked with the 24 available standards at two different concentration levels (0.25 ng and 5 ng on-column for each compound). Thus, 1 mL of standard mix containing 0.0, 0.05 or 1.0 mg/L of each authentic compound was added separately to 1.0 mL of the pooled sample (200 mg dried CA powder/L).

### 2.6.2 Application of precursor ion (MS1) quantification method for plant extracts

For precursor ion quantification (MS1 quantification) of phytochemicals in extracts, the same chromatographic runs for untargeted analysis were used for quantification of distinct phytochemicals.

### TABLE 2

Analytical parameters for authentic standards. Exact m/z used for extracted ion chromatogram (XIC), retention times (RTs), limit of detection (LOD) and limit of quantification (LOQ), percentage of accuracy for three concentrations, and percentage of relative standard deviation (%RSD) are given for 24 selected compounds. Compounds are sorted by RT.

| Compound                        | [M-H]$^{-}$ | RT$^{b}$ (min) | LOD$^{c}$ (μg/L) | LOQ$^{d}$ (μg/L) | Low QC$^{e}$ 0.05 mg/L | Medium QC$^{e}$ 0.5 mg/L | High QC$^{e}$ 5 mg/L | % RSD$^{f}$ |
|---------------------------------|-------------|----------------|------------------|------------------|------------------------|------------------------|---------------------|-------------|
| 5-O-Caffeoylquinic acid         | 353.0867    | 11.85          | 0.078            | 0.260            | 109                    | 108                    | 93                  | 13.11       |
| Epigallocatechin                | 305.0656    | 13.48          | 0.146            | 0.485            | 101                    | 104                    | 99                  | 21.70       |
| Catechin                        | 289.0707    | 14.48          | 0.018            | 0.060            | 123                    | 106                    | 98                  | 11.54       |
| Dihydrocaffeic acid            | 181.0495    | 14.85          | 0.005            | 0.015            | 109                    | 117                    | 93                  | 11.18       |
| 4-O-Caffeoylquinic acid         | 353.0867    | 15.08          | 0.040            | 0.134            | 125                    | 116                    | 93                  | 12.13       |
| 3-O-Caffeoylquinic acid         | 353.0867    | 15.30          | 0.040            | 0.134            | 122                    | 116                    | 97                  | 17.33       |
| Caffeic acid                   | 179.0339    | 15.83          | 0.319            | 1.064            | 111                    | 110                    | 99                  | 11.71       |
| Epicatechin                    | 289.0707    | 16.80          | 0.013            | 0.045            | 107                    | 100                    | 98                  | 12.52       |
| 1,5-Di-cafeoylquinic acid      | 515.1184    | 17.49          | 0.061            | 0.202            | 110                    | 104                    | 97                  | 15.32       |
| 1,3-Di-cafeoylquinic acid      | 515.1184    | 17.49          | 0.061            | 0.202            | 110                    | 104                    | 97                  | 15.32       |
| Rutin                          | 609.145     | 18.96          | 0.016            | 0.052            | 107                    | 102                    | 100                 | 11.65       |
| Dihydroferulic acid            | 195.0652    | 19.02          | 0.027            | 0.089            | 123                    | 104                    | 96                  | 8.98        |
| 3,4-Di-cafeoylquinic acid      | 515.1184    | 19.11          | 0.169            | 0.562            | 105                    | 99                     | 100                 | 15.30       |
| 3,5-Di-cafeoylquinic acid      | 515.1184    | 19.45          | 0.166            | 0.552            | 106                    | 100                    | 102                 | 15.25       |
| Ferulic acid                   | 193.0495    | 19.55          | 0.001            | 0.004            | 111                    | 105                    | 93                  | 6.79        |
| 4,5-Di-cafeoylquinic acid      | 515.1184    | 19.92          | 0.108            | 0.358            | 101                    | 99                     | 102                 | 12.09       |
| Naringin                       | 579.1708    | 20.07          | 0.025            | 0.083            | 97                     | 100                    | 88                  | 10.66       |
| Isoferulic acid                | 193.0495    | 20.42          | 0.001            | 0.003            | 94                     | 102                    | 103                 | 16.39       |
| Quercetin                      | 301.0342    | 21.10          | 0.068            | 0.227            | 111                    | 100                    | 97                  | 12.15       |
| Madecassoside                  | 973.5003    | 21.47          | 0.008            | 0.025            | 110                    | 97                     | 87                  | 16.94       |
| Asiatoside                     | 957.5054    | 21.97          | 0.001            | 0.004            | 102                    | 99                     | 98                  | 18.87       |
| Kaempferol                     | 285.0390    | 22.01          | 0.028            | 0.092            | 112                    | 103                    | 97                  | 10.46       |
| Madecassic acid                | 503.3367    | 23.74          | 0.009            | 0.031            | 97                     | 100                    | 97                  | 11.41       |
| Asiatic acid                   | 487.3418    | 24.41          | 0.005            | 0.015            | 96                     | 95                     | 92                  | 12.61       |

$^{a}$Exact mass; negative ionisation; mass error < 5 ppm.
$^{b}$Retention time.
$^{c}$Calibration detection limit evaluated as S/N ratio 3:1.
$^{d}$Calibration quantification limit evaluated as S/N ratio 10:1.
$^{e}$% of accuracy = (measured concentration/true concentration) × 100 in quality control sample.
$^{f}$%RSD measured for 1 mg/L based on nine measurements over a span of 6 months. Values were calculated using the following equation %RSD = (standard deviation/average) × 100 in Microsoft Excel 2016.
3 | RESULTS AND DISCUSSION

3.1 | Untargeted fingerprinting analysis of CA extracts

We developed a chromatographic method for chemical profiling of botanical samples that uses a phenyl-bonded phase, whereby the phenyl groups are directly bonded to the silica surface and separations are governed by $\pi-\pi$ interactions. This stationary phase was selected to take advantage of the presence of phenolic scaffolds in many specialised metabolites, in particular those originating from the phenylpropanoid biosynthesis pathway. The method requires 30 min per chromatographic run. This LC separation method provides suitable resolution and peak capacity for chemical fingerprinting based on DDA, while minimising peak suppression and matrix effects. Representative XICs are shown in Figures S2 and S3. The peak width at half height is sufficient to conduct quantification of phenolic compounds using precursor ion (MS1) quantification with sufficient reproducibility of peak area determinations.

Figure 1 shows a typical total ion chromatogram (TIC) for a *C. asiatica* water extract and the most intense molecular features fragmented in the DDA experiment acquired in negative ion mode. From over 20000 $m/z$-features detected, 117 compounds were annotated (Table 1) after applying the workflow outlined in Figure S1. To our knowledge, this analysis includes 87 compounds that have been reported in plants previously but have been now detected for the first time in *C. asiatica* extracts. MS/MS spectra and spectral matches of the newly detected compounds in CA are provided in Figure S7. Some of the most abundant compounds include six di-cafeoylquinic acid isomers, quinic acid, mono-cafeoylquinic acids, and several glycosides, such as asiaticoside, madecassoside and quercetin 3-O-glucoside. It is noteworthy that the current chromatographic separation conditions resolved di-cafeoylquinic acids isomers 3,4-, 3,5- and 4,5-di-cafeoylquinic acids (Table S1, Figure S5). Analytical parameters for the annotated compounds, namely $m/z$, RT, detected adducts and molecular formulas are shown in detail in Table S1. When compounds were detected in both ion modes, the one with the highest signal-to-noise ($S/N$) ratio was included. Annotated compounds include five hydroxycinnamic acids, nine mono- and di-cafeoylquinic acids, 12 terpenoids, 13 flavonoids, 11 hexosides among other phytochemicals (Table 1).

In order to capture the similarities and differences in metabolite composition observed for the eight CA accessions available to us, we used principal component analysis (PCA) on log-transformed and Pareto-scaled data using 5512 $m/z$ features containing MS/MS spectral information in negative ion mode (Figure 2A). PCA revealed...

**FIGURE 1** Examples of typical data obtained by untargeted analysis of a pooled *Centella asiatica* (CA) water extract using the data dependent acquisition mode. (A) Total ion chromatogram (ESI$^-$) (10 µL injection, 1 mg/L). (B) Distribution map of precursor ions submitted to collision induced dissociation along the elution period. The $y$-axis provides $m/z$ information for the precursor ion; the $x$-axis represents the elution times for each precursor ion. Each one of the 5512 dots contains a fragmentation spectrum. The intensity of the colour blue represents the ion abundance of the precursor ion. TIC, total ion chromatogram; DDA, data dependent acquisition [Colour figure can be viewed at wileyonlinelibrary.com]
significant differences across the CA accessions. The PCA loading plots show the constituents with more variability among the C. asiatica accessions (Figure 2B). Di-caffeoylquinic acids and triterpenes showed high variation between the different accessions. For additional contrasting of CA accessions, we also used as tools a correlation matrix and a heatmap based on area under the curve for extracted ion peaks. The correlation matrix aids in evaluating similarities and dissimilarities of extracts based on the correlation score (Figure 2C). Higher correlation scores (between 0.75 and 1) are indicated by red; scores between 0.74 and 0.51 are indicated by white; scores between 0.5 and 0.25 are represented by purple; and lower correlation scores (less than 0.25) are represented by blue. The Pearson correlation value calculated between samples CA6 and CA2 sample is 0.27, which indicates that there is a small linear relationship between CA2 and CA6. The Pearson correlation value calculated between CA2 and CA1 samples is 0.48, which also indicates a small linear relationship between CA2 and CA1 samples. The Pearson correlation between CA6 and CA4 samples is 0.87 which indicates that...
these two samples are linearly related, indicating similarities in metabolite contents for CA4 and CA6 extracts. The heatmap with hierarchical clustering (Figure 2D) visualises the precursor ion peak areas for 14 compounds evaluated in all CA extracts; these compounds were selected because they were present at relatively high concentration in all plant accessions and authentic standards were commercially available. Peak areas were averaged across three replicates. The dendrogram on the y-axis indicates the degree of similarity or difference between the CA compound levels in the CA accessions, e.g. CA3 and CA8 are closer in the clustering tree indicating higher similarity of the compound levels in extracts CA3 and CA8, whereas extract CA6 is separated in the dendrogram from CA3 and CA8 indicating little similarity of the compound levels of the C6 extract with the levels found in the extracts of CA3 and CA8.

3.2 Structural similarity network

Figure 3 shows a 2D structural similarity network of 117 assigned compounds (Table S1) found consistently in the aqueous extract of all eight CA accessions, which was created using the Tanimoto similarity score. Compounds are arranged in 14 interconnected clusters that are structurally similar at the boundary nodes at the 95% confidence level. We used the 2D structural similarity network to support our tentative annotations of metabolites and the associated categorisation into compound classes or clusters (Table 1). Compounds fall into the following clusters: 14 amino acid derivatives, five amino sugar derivatives, four choline derivatives, six di-caffeoylquinic acids, seven fatty acid derivatives, 13 flavonoids, nine hexoside derivatives, five hydroxycinnamic acids, three mono-caffeoylquinic acids, four organic acids,
21 phenolic compounds, seven purine derivatives, 12 terpenoids, and seven other compounds. Classification was established according to structural similarity (Tanimoto algorithm); consequently, some compounds may belong to more than one compound class. While C. asiatica is most known as a rich source of pentacyclic triterpenoids,\textsuperscript{19,24} relatively high percentages of caffeoylquinic acids and flavonoids have also been identified.\textsuperscript{19,55} These specialised metabolites, specifically phenylpropanoid derivatives, have been associated with C. asiatica’s anti-inflammatory, antioxidant, or other biological activities.\textsuperscript{60,61}

Noteworthy, but beyond the current study, the enzymatic machinery required to produce specialised metabolites in most plants, including C. asiatica, is largely uncharacterised.\textsuperscript{62} The lack of intermediate specialised metabolites along the metabolic pathways makes this characterisation more complex. The detection of specialised metabolites that have not yet been reported in plant extracts, in conjunction with a Cytoscape network clustered according to the structural similarity (Tanimoto algorithm), can potentially help with characterising new metabolic pathways by searching for potential enzymes responsible for the interconversion of metabolites clustering together.

### 3.3 Accurate quantification of phytochemicals in extracts using precursor ion (MS1) quantification

In this work, we demonstrate the suitability of using a LC–MS/MS DDA method for the screening of compounds and molecular ion (MS1) extraction for quantification of selected compounds in the same chromatographic run in conjunction with external calibrations for the selected marker compounds. From over 20000 recorded m/z features, the 5512 and 6906 most prominent m/z features acquired in negative (Figure 1) and positive ion mode, respectively, were fragmented in the DDA experiment. This untargeted approach provides a spectral library for thousands of potential compounds that can be mined in future applications. From the 117 tentatively assigned compounds (Table 1), 24 compounds were selected as a proof of concept.

| Table 3 | Recovery experiment. Recovery percentage and mean concentration of individual quantified compounds were measured in a pooled CA sample (100 mg/L) using precursor ions with respective standard deviations obtained without standard addition and after addition of a mixture of 24 standards in two different concentration levels (0.25 and 5 ng of each standard on-column). All measurements are given in nanograms |

| Compound                        | QC (ng on column) | QC + 0.25 ng standards | % Recovery\textsuperscript{a} | QC + 5 ng\textsuperscript{b} standards | % Recovery\textsuperscript{a} |
|---------------------------------|-------------------|------------------------|-----------------------------|----------------------------------------|-----------------------------|
| 5-O-Caffeoylquinic acid         | 1.02 ± 0.01       | 1.28 ± 0.01            | 102                         | 6.53 ± 0.46                            | 110                         |
| Epigallocatechin                | <LOQ\textsuperscript{a} | 0.18 ± 0.02            | 71                          | 4.55 ± 0.20                            | 91                          |
| Catechin                       | <LOQ              | 0.33 ± 0.09            | 134                         | 5.88 ± 0.32                            | 117                         |
| Dihydrocaffeic acid            | <LOQ              | 0.28 ± 0.01            | 113                         | 5.34 ± 0.27                            | 107                         |
| 4-O-Caffeoylquinic acid         | 0.88 ± 0.02       | 1.17 ± 0.03            | 115                         | 6.31 ± 0.09                            | 109                         |
| 3-O-Caffeoylquinic acid         | 2.45 ± 0.03       | 2.79 ± 0.09            | 133                         | 7.68 ± 0.10                            | 105                         |
| Caffeic acid                   | 0.67 ± 0.07       | 0.97 ± 0.03            | 123                         | 6.06 ± 0.43                            | 108                         |
| Epicatechin                    | <LOQ              | 0.30 ± 0.06            | 119                         | 5.57 ± 0.14                            | 111                         |
| 1,5-Di-caffeoylquinic acid     | 0.38 ± 0.01       | 0.68 ± 0.06            | 121                         | 5.91 ± 0.28                            | 110                         |
| 1,3-Di-caffeoylquinic acid     | 0.38 ± 0.01       | 0.68 ± 0.06            | 121                         | 5.91 ± 0.28                            | 110                         |
| Rutin                          | 0.04 ± 0.01       | 0.31 ± 0.02            | 106                         | 5.44 ± 0.35                            | 108                         |
| Dihydroferulic acid            | <LOQ              | 0.25 ± 0.01            | 101                         | 5.71 ± 0.17                            | 114                         |
| 3,4-Di-caffeoylquinic acid     | 3.23 ± 0.02       | 3.57 ± 0.02            | 136                         | 8.86 ± 0.17                            | 113                         |
| 3,5-Di-caffeoylquinic acid     | 3.78 ± 0.02       | 4.14 ± 0.03            | 144                         | 8.94 ± 0.50                            | 103                         |
| Ferulic acid                   | 0.11 ± 0.01       | 0.40 ± 0.04            | 118                         | 5.59 ± 0.11                            | 109                         |
| 4,5-Di-caffeoylquinic acid     | 3.93 ± 0.01       | 4.16 ± 0.05            | 92                          | 9.16 ± 0.10                            | 105                         |
| Naringin                       | <LOQ              | 0.29 ± 0.03            | 117                         | 5.86 ± 0.84                            | 117                         |
| Isoferulic acid                | 0.18 ± 0.01       | 0.52 ± 0.04            | 135                         | 6.35 ± 0.21                            | 123                         |
| Quercetin                      | 0.29 ± 0.22       | 0.56 ± 0.17            | 109                         | 5.94 ± 0.38                            | 113                         |
| Madecassoside                  | 25.5 ± 0.95       | 25.77 ± 0.07           | 108                         | 31.6 ± 2.20                            | 110                         |
| Asiaticoside                   | 10.74 ± 0.12      | 11.0 ± 0.16            | 114                         | 16.1 ± 0.55                            | 108                         |
| Kaempferol                     | 0.31 ± 0.02       | 0.62 ± 0.01            | 121                         | 5.79 ± 0.49                            | 109                         |
| Madecassic acid                | 1.35 ± 0.01       | 1.65 ± 0.03            | 120                         | 7.58 ± 0.24                            | 124                         |
| Asiatric acid                  | 0.62 ± 0.02       | 0.97 ± 0.28            | 138                         | 7.25 ± 0.53                            | 132                         |

\textsuperscript{a}\% Recovery = (Cf – Ci)/Cspiked × 100. Ci, nanograms measured (on-column) before standard addition; Cf, nanograms measured after standard addition; Cspiked, nanograms of spiked standard.

\textsuperscript{b}Calibration quantification limit evaluated as S/N ratio 10:1.
for the MS1-based quantification, including three monoa-
caffeoylquinic acids, five di-caffeoylquinic acids, seven flavonoids, five hydroxycinnamic acid derivatives and four triterpenes (structures shown in Figure S4).

3.4 | Method validation for selected compounds

For accuracy, precision, repeatability, linearity, LOD, LOQ and range, the proposed method followed the typical validation procedure in accordance with the ICH Harmonised Tripartite Guideline.63

Figure S5 shows LC–MS XICs obtained from authentic standards of the 24 compounds that were selected as phytochemical marker compounds. Analytical parameters, namely [M-H]− m/z, RT, accuracy for three different concentration levels, LOD, LOQ and inter-day coefficient of variation [relative standard deviation (RSD)] were established for the 24 phytochemical markers using precursor ion (MS1) extraction (Table 2). The analytical accuracy for three known concentration samples at the low (0.05 ppm), medium (0.50 ppm) and high (5.0 ppm) calibration curve intervals ranged from 87 to 125% (Table 2). The RSD was measured for a solution of 1 mg/L and ranged from 6.8 to 24% for nine repetitions measured in a span of 6 months (Table 2).

The availability of high-resolution accurate mass data allowed us to obtain comprehensive fingerprints for botanical extracts that can be further interrogated post acquisition to obtain accurate quantification of phytochemicals by extracting the precursor ions and using the area under the peak for quantification (MS1 quantification) in the same analytical run. Quantified compounds showed good linearity over three orders of magnitude (0.005–5.0 mg/L, r > 0.990, Table S2).

Matrix effects are frequently observed when analysing complex samples. In order to evaluate matrix effects in the CA extracts, pooled CA extract samples were spiked with the 24 available standards. TICs obtained for a CA extract and the same sample after standard addition is shown in Figure S6. For plant extracts, recoveries of individual compounds ranged from 71 to 144% and 91 to 132% for 0.25 and 5.0 ng on-column, respectively (Table 3), confirming the feasibility of the proposed procedure for quantitative analysis of marker compounds in CA extracts.

A range of three orders of magnitude is typical for time-of-flight (TOF) analysers, which is a disadvantage when compared with the dynamic range of triple quadrupole analysers, which usually feature a linear dynamic range that extends over six orders of magnitude. Nevertheless, the high resolution allows us to obtain chemical fingerprint and quantification of marker compounds in the same analytical run.64

![FIGURE 4](image_url) Extracted ion chromatograms (XICs) of 18 compounds that were used for precursor ion (MS1) quantification. Individual analytical parameters are shown in Table 2. XICs were obtained using the data-dependent acquisition (DDA) (ESI−) mode obtained for a pooled Centella asiatica (CA) sample [Colour figure can be viewed at wileyonlinelibrary.com]
saving instrument time and solvents, and avoiding sample degradation due to storage.

For the developed quantification method, the combination of an optimised separation method with a high resolution quadrupole time-of-flight (q-TOF) mass spectrometer allowed the detection and quantification of phytochemicals in plant extracts at sub-parts per billion levels (except caffeic acid; LOQ 1.06 μg/L) with minimum sample processing. Modern q-TOF mass spectrometers possess sensitivities

**FIGURE 5** Precursor ion (MS1) quantification of 15 phytochemicals in the eight different *Centella asiatica* accessions (water extracts). Results are presented as milligrams per gram of dry extract; standard error derived from triplicates analysis [Colour figure can be viewed at wileyonlinelibrary.com]
typically associated with MS/MS-based selected reaction monitoring (SRM) methods. Reported LOD values for 15 phenolic acids and 17 flavonoids acquired using SRM in a triple quadrupole mass spectrometer range from 3.4 to 228 µg/L, and are comparable with our LODs. Contemporary q-TOF instruments typically offer mass resolving power of \( \geq 25000 \) [full width at half maximum (FWHM)] at \( m/z \) 195. These q-TOF platforms obtain accurate mass measurements with high resolution for precursor ions and fragment ions thus allowing structural characterisation and quantification of phytochemicals in complex mixtures with high confidence.

### 3.5 | Quantification of phytochemical marker compounds in CA extracts from different sources

In previous studies, comparisons of specialised metabolite production of \( C. asiatica \) were limited to four triterpenoids (asatic acid, madecassic acid and their glycosides asiaticoside and madecassoside). Some other compounds (flavonoids and caffeoyl esters) were analysed by LC-MS and HPLC-DAD (diode array detector). A comparative study of nutrient content and yield performance of \( C. asiatica \) at different harvesting periods was reported. The study focused on yield measured by dry weight of leaves and on nutrient comparisons [phosphorus (P), potassium (K), sulphur (S), calcium (Ca), magnesium (Mg), zinc (Zn), copper (Cu), iron (Fe), manganese (Mn) and nitrogen (N)].

Our study shows that CA extracts were particularly rich in mono-caffeoylquinic acids, such as 3-caffeoylquinic acid, 4-caffeoylquinic acid and 5-caffeoylquinic acid, and di-caffeoylquinic acids, such as 1,3-di-caffeoylquinic acid, 1,5-di-caffeoylquinic acid, 3,4-di-caffeoylquinic acid, 3,5-di-caffeoylquinic acid and 4,5-di-caffeoylquinic acid, as well as some triterpenoids such as isatiking and its aglycones (Table 3, Figures 4 and 5). CA extracts also contained several flavonoids and hydroxycinnamic acid derivatives. Figure S8 compiles positive matches with certified standards. Figure 4 shows XICs for 18 selected compounds quantified in CA water extracts using area under the curve of the precursor ion acquired in DDA mode.

Under the HPLC conditions used, 1,3- and 1,5-di-caffeoylquinic acid co-elute since they are stereoisomers (structures shown in Figure S4). For all other compounds quantified in this study, the combination of suitable separation conditions with extraction of molecular ions (MS1) chromatograms at accurate \( m/z \) values (Figure S5, Table 2) enabled detection limits in the low nanomolar to picomolar range for 24 phytochemicals (Tables 2 and S2). The optimised analytical procedure minimised interferences by improving the chromatographic separation of isomers with similar fragmentation patterns and thereby also optimised detection limits.

In this study, eight accessions of \( C. asiatica \) were quantified using precursor ion (MS1) quantification (Figure 5). Comparing samples CA6 and CA1, the concentrations of asiaticoside varied by 19.9-fold, asatic acid by 9.1, and madecassoside by 1.5. In the case of di-caffeoylquinic acids, comparing CA5 and CA1, 4,5-dicaffeoylquinic acid varied 11-fold. Mono-caffeoylquinic acids presented less variation across the accessions (Figure 5). This emphasises the importance of establishing rigorous analytical procedures for botanical extracts and supplements to ensure product integrity and batch-to-batch reproducibility.

To conclude, we developed a method for untargeted and targeted characterisation of CA extracts using the same chromatographic run. The combination of suitable separation conditions with mass spectral data acquired with high resolving power using DDA enables the extraction of high resolution accurate mass precursor ions with exact \( m/z \) values that allows accurate quantification of phytochemicals with LOQs at 1.06 µg/L or lower. The described method was validated for the quantification of (a) seven flavonoids, (b) three structural isomers of caffeoylquinic acids, (c) five di-caffeoylquinic acids, (d) five caffeic acids derivatives and (e) four terpenoids. The concentration of targeted metabolites across different CA accessions was substantial, with the largest difference (20-fold) observed for asiaticoside between CA1 versus CA6, demonstrating that standardisation and detailed characterisation of plant extracts are prerequisites for reliable and reproducible studies aiming to determine the biological activity of CA and botanical extracts in general.

Overall, the current study underscores the need for methods to efficiently analyse highly complex plant extracts to support the standardisation of botanicals destined for preclinical studies, clinical trials and commercial products.

### 3.6 | Mass spectrometry data deposition

MS/MS data have been deposited to the GNPS repository (http://gnps.ucsd.edu) with the dataset identifiers MSV000084588 (ESI+) and MSV000084621 (ESI−).

### ACKNOWLEDGEMENTS

The authors acknowledge NIH grants R01AT008099, R61AT009629, T32 AT002688 and T32 AT002688 and S10RR027878.

### ORCID

Armando Alcazar Magana [https://orcid.org/0000-0002-4731-0247](https://orcid.org/0000-0002-4731-0247)
Claudia S. Maier [https://orcid.org/0000-0003-0743-8956](https://orcid.org/0000-0003-0743-8956)

### REFERENCES

1. Pyne ME, Narcross L, Martin VJ. Engineering plant secondary metabolism in microbial systems. Plant Physiol. 2019;179(3):844-861.
2. Wang S, Alseekh S, Fernie AR, Luo J. The structure and function of major plant metabolite modifications. Mol Plant. 2019;12(7):899-919.
3. Hartmann T. Diversity and variability of plant secondary metabolism: a mechanistic view. Entomol Exp Appl. 1996;80(1):177-188.
4. Commissio M, Strazzer P, Toffall K, Stocchero M, Guzzo F. Untargeted metabolomics: an emerging approach to determine the composition of herbal products. Comput Struct Biotechnol J. 2013;4:1-7.
5. Mahady GB. Global harmonization of herbal health claims. J Nutr. 2001;131(3):1120S-1123S.
6. van Breemen RB. Development of safe and effective botanical dietary supplements. J Med Chem. 2015;58(21):8360-8372.
7. Neiberg RH, Ackin M, Grzywacz JG, et al. Occurrence and co-occurrence of types of complementary and alternative medicine use by age, gender, ethnicity, and education among adults in the United States: the 2002 National Health Interview Survey (NHIS). *J Altern Complement Med*. 2011;17(4):363-370.

8. Saydah SH, Eberhardt MS. Use of complementary and alternative medicine among adults with chronic diseases: United States 2002. *J Altern Complement Med*. 2006;12(8):805-812.

9. Dietz BM, Chen SN, Alvarenga RFR, et al. DESIGNER extracts as tools to balance estrogenic and chemopreventive activities of botanicals for women’s health. *J Nat Prod*. 2017;80(8):2284-2294.

10. van Breemen RB, Fong HH, Farnsworth NR. Ensuring the safety of botanical dietary supplements. *Am J Clin Nutr*. 2008;87(2):509S-513S.

11. Bruni R, Sacchetti G. Factors affecting polyphenol biosynthesis in wild and field grown St. John’s Wort (*Hypericum perforatum* L. Hypericaceae/Guttiferae). *Molecules*. 2009;14(2):682-725.

12. Alqahtani A, Tongkao-on W, Li KM, Razmovski-Naumovski V, Chan K, Li QG. Seasonal variation of triterpenes and phenolic compounds in Sunset® and Centella asiatica (L.) Urb. *Phytochem Anal*. 2015;26(6):436-443.

13. Howard L. Processing techniques and their effect on fruit and vegetable phytochemicals. *Food Science, Technology and Nutrition*. 2008;2008:449-472.

14. Magana AA, Wrobel K, Elguera JCT, Escobosa ARC, Wrobel K. Determination of small phenolic compounds in tequila by liquid chromatography with ion trap mass spectrometry detection. *Food Anal Method*. 2015;8(4):864-872.

15. Fong H, Pauli G, Bolton J, et al. Evidence-Based Herbal Medicine: Challenges in Efficacy and Safety Assessments. 2 UIC/NIH Center for Botanical Dietary Supplements Research; 2006.

16. Glade MJ. The dietary supplement health and education act of 1994 – focus on labeling issues. *Nutrition*. 1997;13(11-12):999-1001.

17. Fong HH-S, Pauli GF, Bolton JL, et al. Evidence-Based Herbal Medicine: Challenges in Efficacy and Safety Assessments. In: Current Review of Chinese Medicine. Vol.2 Chicago, IL: UIC/NIH Center for Botanical Dietary Supplements Research; 2006:11-26.

18. Kapoor LD. *CRC Handbook of Ayurvedic Medicinal Plants*. Boca Raton, FL: CRC Press; 1990.

19. Brinckhaus B, Lindner M, Schuppan D, Hahn EG. Chemical, pharmacological and clinical profile of the east Asian medical plant *Centella asiatica*. *Phytotherapy*. 2000;7(5):427-448.

20. Shinomol GK, Muralidhara, Bharath MM. Exploring the role of “Brahmi” (*Centella asiatica* L.) on cognitive and oxidative stress in an intracerebroventricular streptozotocin model of Alzheimer’s disease in rats. *Clinical & Experimental Pharmacology & Physiology*. 2003;30(5-6):336-342.

21. Decaestecker TN, Vande Casteele SR, Lallemaqc PE, Van Peteghem CH, Defore DL, Van Boeckel JR. Information-dependent acquisition-mediated LC-MS/MS screening procedure with semi-quantitative potential. *Anal Chem*. 2004;76(21):6365-6373.

22. Sumner LW, Amberg A, Barrett D, et al. Proposed minimum reporting standards for chemical analysis chemical analysis working group (CAWG) metabolomics standards initiative (MSI). *Metabolomics*. 2007;3:211-221.

23. Vian MR, Kurland UJ, Jones MR, Dunn WB. How close are we to complete annotation of metabolomes? *Curr Opin Chem Biol*. 2017;36:64-69.

24. Guilas C, Montenegro-Burke JR, Domingo-Almenara X, et al. METLIN: a technology platform for identifying knowns and unknowns. *Anal Chem*. 2018;90(5):3156-3164.

25. Wishart DS, Feunang YD, Marcu A, et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res*. 2018;46(D1):D608-D617.

26. Housley L, Magana AA, Hsu A, et al. Untargeted metabolomic screen reveals changes in human plasma metabolite profiles following consumption of fresh broccoli sprouts. *Mol Nutr Food Res*. 2018;62(19):1-6.

27. Housley L, Magana AA, Hsu A, et al. Untargeted metabolomic screen reveals changes in human plasma metabolite profiles following consumption of fresh broccoli sprouts. *Mol Nutr Food Res*. 2018;62(19):1-6.

28. Wolfs S, Schmidt M, Muller–Hannemann M, Neumann S. In silico fragmentation for computer assisted identification of metabolite mass spectra. *BMC Bioinformatics*. 2010;11:1-12.

29. Horai H, Arita M, Nishioka T. Comparison of ESI-MS spectra in MassBank database. Paper presented at: 2008 International Conference on BioMedical Engineering and Informatics; 27-30 May 2008.

30. Neiberg RH, Aickin M, Grzywacz JG, et al. Use of complementary and alternative medicine among adults with chronic diseases: United States 2002. *J Altern Complement Med*. 2006;12(8):805-812.

31. Wrobel K, Elguera JCT, Escobosa ARC, Wrobel K. Determination of small phenolic compounds in tequila by liquid chromatography with ion trap mass spectrometry detection. *Food Anal Method*. 2015;8(4):864-872.

32. Van Peteghem CH, Defore DL, Van Boeckel JR. Information-dependent acquisition-mediated LC-MS/MS screening procedure with semi-quantitative potential. *Anal Chem*. 2004;76(21):6365-6373.

33. Sumner LW, Amberg A, Barrett D, et al. Proposed minimum reporting standards for chemical analysis chemical analysis working group (CAWG) metabolomics standards initiative (MSI). *Metabolomics*. 2007;3:211-221.

34. Vian MR, Kurland UJ, Jones MR, Dunn WB. How close are we to complete annotation of metabolomes? *Curr Opin Chem Biol*. 2017;36:64-69.

35. Guilas C, Montenegro-Burke JR, Domingo-Almenara X, et al. METLIN: a technology platform for identifying knowns and unknowns. *Anal Chem*. 2018;90(5):3156-3164.

36. Wishart DS, Feunang YD, Marcu A, et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res*. 2018;46(D1):D608-D617.

37. Housley L, Magana AA, Hsu A, et al. Untargeted metabolomic screen reveals changes in human plasma metabolite profiles following consumption of fresh broccoli sprouts. *Mol Nutr Food Res*. 2018;62(19):1-6.

38. Wolfs S, Schmidt M, Muller–Hannemann M, Neumann S. In silico fragmentation for computer assisted identification of metabolite mass spectra. *BMC Bioinformatics*. 2010;11:1-12.

39. Horai H, Arita M, Nishioka T. Comparison of ESI-MS spectra in MassBank database. Paper presented at: 2008 International Conference on BioMedical Engineering and Informatics; 27-30 May 2008.

40. Neiberg RH, Aickin M, Grzywacz JG, et al. Use of complementary and alternative medicine among adults with chronic diseases: United States 2002. *J Altern Complement Med*. 2006;12(8):805-812.
