Received: 12 May 2018 | Revised: 20 November 2018 | Accepted: 20 November 2018

DOI: 10.1002/dta.2550

RESEARCH ARTICLE

WILEY

Application of a molecular networking approach for clinical and forensic toxicology exemplified in three cases involving 3-MeO-PCP, doxylamine, and chlormequat

Sophie Allard[1] | Pierre-Marie Allard[2] | Isabelle Morel[1,3] | Thomas Gicquel[1,3]

[1] Laboratory of Forensic and Clinical Toxicology, University Hospital of Rennes, Rennes, France

[2] School of Pharmaceutical Sciences, University of Geneva and University of Lausanne, Geneva, Switzerland

[3] INSERM, INRA, CHU Rennes, Institut NuMeCan (Nutrition, Metabolism and Cancer) University Rennes, Rennes, France

Correspondence
Sophie Allard, Laboratoire de toxicologie biologique et médico-légale, CHU de Pontchaillou, 2 rue Henri Le Guilloux, 35033 Rennes cedex 9, France.
Email: sophie.allard.pro@gmail.com

Abstract

Untargeted toxicological screening is an analytical challenge, given the high number of molecules and metabolites to be detected and the constant appearance of new psychoactive substances (NPS). The combination of liquid chromatography with high-resolution tandem mass spectrometry (HRMS/MS) in a data-dependent acquisition mode generates a large volume of high quality spectral data. Commercial software for processing MS data acquired during untargeted screening experiments usually compare measured features (mass, retention time, and fragmentation spectra) against a predefined list of analytes. However, there is a lack of tools for visualizing and organizing MS data of unknown compounds. Here, we applied molecular networking to untargeted toxicological screening. This bioinformatic tool allows the exploration and organization of MS/MS data without prior knowledge of the sample's chemical composition. The organization of spectral data is based on spectral similarity. Hence, important information can be obtained even before the annotation step. The link established between molecules enables the propagation of structural information. We applied this approach to three clinical and forensic cases with various matrices: (a) blood and a syringe content in a forensic case of death by self-injection, (b) hair segments in a case of drug-facilitated assault, and (c) urine and blood samples in a case of 3-methoxyphencyclidine intoxication. Data preprocessing with MZmine allows sample-to-sample comparison and generation of multisample molecular networks. Our present study shows that molecular networking can be a useful complement to conventional approaches for untargeted screening interpretation, for example for xenobiotics identification or NPS metabolism elucidation.

KEYWORDS

high-resolution mass spectrometry, molecular networking, NPS metabolism, sample-to-sample comparison, untargeted toxicological screening

1 | INTRODUCTION

In forensic and clinical toxicology, the analyst is often confronted with complex, multidimensional problems. The results of a toxicological analysis can have important clinical and legal consequences. Given that the reason for intoxication and the nature of the substances confiscated by the authorities are often unknown, broad screening strategies are required. Liquid chromatography coupled with high-resolution tandem mass spectrometry (LC–HRMS/MS) is now considered to be the reference method for untargeted toxicological screening.[1,2]

In a case of intoxication, various samples (ie, blood, urine, bile, hair, seized products) can be analyzed, and sample-to-sample

comparison can facilitate interpretation of the results. Advances in analytical techniques and the increased availability of HR mass spectrometers have led to the generation of large volumes of high-quality data. Sample-to-sample spectral comparisons are difficult to carry out on a routine basis, due to the lack of suitable tools. Efforts to achieve high analytical reliability, sensitivity, and specificity are important prerequisites. However, the interpretation of raw spectral information is equally important. As it is essential to consider the context and clinical findings when interpreting the results of toxicological screening, "the molecular environment" (including metabolites of the analyte of interest) is an important factor. This is particularly true for the ever-increasing problem of detection of new psychoactive substances (NPS). There may be a considerable delay between the initial report on an NPS by the authorities and its inclusion in a spectral library. Human metabolites of recent NPS are even less likely to be reported. Most of available proprietary or open software only allows for the annotation of known compounds (previously reported in spectral libraries).[3] To keep pace with the rapidly changing market for illicit NPS, there is a need to develop innovative methods for the detection and identification of unreported compounds.[4-6]

Molecular networking is an approach that is capable of representing MS/MS data in graphical form, which has been developed and validated for the organization of spectral data acquired on small molecules.[7] It has already been applied to several research fields such as metabolomics, drug metabolism,[8] natural products research, drug discovery and precision medicine.[9] A molecular network is a visual display of the spectra acquired in an MS/MS experiment, enabling the efficient comparison of MS profiles. In a molecular network, each node represents an ion with its associated fragmentation spectrum, whereas the links between the nodes indicate similarities between spectra. A molecular network provides an overview of (a) the molecules detected and fragmented during an untargeted MS experiment, and (b) their structural relationships. By propagating structural annotation information within the network, this approach offers interesting perspectives for analyzing the metabolites of xenobiotics. The organization of similar but non-identical spectral data is also promising for the specific case of NPS identification, allowing scaffold analogs to be easily highlighted.

The objective of the present study is to apply molecular networking to human cases where its use may open up valuable perspectives. By retrospective exploitation of HR spectral data from three representative clinical and forensic cases, we explore the relevance of a workflow combining (a) preprocessing of raw spectral data with MZmine 2 software and (b) the generation of molecular networks. To the best of our knowledge, the present study is the first to apply a molecular networking approach to routine case studies in clinical and forensic toxicology.

2 | EXPERIMENTAL

2.1 | Sample preparation

The sample extraction and analytical methods have been validated and described in detail elsewhere.[10] Briefly, 200 µL of biological samples (blood and urine) were added to 300 µL of a 0.1 M zinc sulfate solution, and then supplemented with 500 µL of methanol containing the internal standard (0.2 mg/L flurazepam) (in Cases 1, 2, and 3). Solutions were vortexed for 2 minutes, left for 10 minutes at 4.0°C, and then centrifuged for 10 minutes at 3000 g. Supernatants (the upper phase) were evaporated to dryness at 50°C under a stream of nitrogen. Residues were dissolved in 200 µL of LC–MS grade water, vortexed, and again centrifuged for 5 minutes at 3000 g before transferring the supernatants into vials for LC–HRMS/MS analysis. Hair samples (Case 2) were first decontaminated with dichloromethane, cut into 1 cm segments, extracted, and analyzed as per the Society of Hair Testing guidelines.[11] For the analysis of the liquid solution (Case 1), the sample was directly diluted in mobile phase.

2.2 | Analytical platform

Analyses were carried out on an Orbitrap Q Exactive™ mass spectrometer coupled to an Accela 1250 pump (both from Thermo Scientific, San Jose, CA, USA) using a heated electrospray ionization source (HESI-II). Data acquisition, calibration and instrument control were managed using Xcalibur® 2.1 software (Thermo Scientific, San Jose, CA, USA).

2.3 | LC settings

LC separation was performed on a Hypersil GOLD PFP column (150 x 2.1 mm, 5 µm) (Thermo Scientific, San Jose, CA, USA). The mobile phases were composed of 10mM ammonium acetate and 0.1% formic acid in water (A), and 0.1% formic acid in acetonitrile (B). An LC gradient was performed from 95% to 5% A for 20 minutes, followed by a 7-minute plateau with 5% A; and then a 5-minute re-equilibration step. The flow rate was 200 µL/min, the column temperature was 25°C, the injection volume was 20 µL, and the samples were maintained at 15°C in the autosampler.

2.4 | MS settings

The instrument operated alternately in positive and negative ESI modes during the same run. The range for acquisition was respectively 80–800 *m/z* in positive mode and 150–800 *m/z* in negative mode. Ion precursor selection was performed in a data-dependent mode, where the most intense ions from the previous scan were selected for fragmentation. Prior to analysis, the device was calibrated in both ionization modes (using Pierce™ calibration solutions, Thermo Scientific, San Jose, CA, USA). Full scan (MS1) data were acquired for each ionization mode at a resolution of 70 000 full width at half maximum (FWHM), with an AGC target of 1e6, and a maximum injection time of 250 ms. The source parameters were as follows: source voltage: + 4.0 and - 4.5 kV; sheath gas flow: 35 units; auxiliary gas flow: 15 units; sweep gas flow: 2.5 units; capillary temperature: 300°C; and S-Lens RF level: 50 units. MS/MS (MS2) data were acquired in profile mode at a resolution of 17 500 FWHM with an automatic gain control (AGC) target of 1e6, a maximum injection time 250 ms, a TopN of 8 in positive mode and 3 in negative mode, an isolation window of 3 *m/z*, a normalized collision energy (NCE) of 70, and a dynamic exclusion time of 10 seconds. MS data were recorded in profile mode.

3 | DATA PROCESSING

The main steps of the workflow are summarized in the Supporting Information, Figure SI-A. All the software programs used in these steps are open source, and can be accessed freely online.

3.1 | MS data conversion

Raw spectral data files were converted from Thermo Fisher's proprietary .raw format into the .mzXML open mass format. This step was performed with the command line version of MSConvert (http://proteowizard.sourceforge.net) as the graphical user interface (GUI) does not offer the function of separation between positive and negative ionization. Spectra were converted from profile to centroid mode and positive and negative ionization traces were separated in two files. Our analysis focused on the positive ionization mode.

3.2 | Data preprocessing and MS1 annotation

Preprocessing of the data was required for the generation of multisample molecular networks, and the use of precursor ion abundance for relative quantification. This also allows the retention time to be taken into account, and thus to separate isomers into individual nodes. The different preprocessing steps for molecular network generation using MZmine 2 software (http://mzmine.github.io) have been described by Olivon et al.[12] Briefly, the main steps were deconvolution, alignment, and annotation of MS1 data using a custom database (a .csv file) containing the following features: the compound's name, the exact mass of the protonated molecule, the retention time, and the molecular formula. We eventually built a database of 2980 compounds of toxicological interest, of which 653 had associated retention time information (according to our LC system). Lastly, each preprocessed spectral file was saved as a single .mgf file using the integrated Global Natural Products Social Molecular Networking (GNPS) export module form MZmine (https://ccms-ucsd.github.io/GNPSDocumentation/featurebasedmolecularnetworking/). Detailed parameters (provided as an .xml file, ready for use in MZmine) can be found in the Supporting Information.

3.3 | Generation of molecular networks

The spectral file .mgf was uploaded to the GNPS platform (http://gnps.ucsd.edu). Due to the use of HR data, the basic parameters were modified to 0.02 *m/z* for the mass tolerance of precursor and fragment ions used for MS/MS spectral library searching, and 0.02 *m/z* for the mass tolerance of fragment ions used for molecular networking. As MZmine 2 was used to preprocess our data, the MSCluster feature was disabled and the minimum cluster size was set to 1. We kept the following default settings: (i) Links between nodes were created when the cosine score was greater than 0.7, and the minimum number of common fragment ions shared by two MS/MS spectra was 6 (ii) Links between two nodes were only kept in the network if each node was in the top 10 most similar nodes. Once the network had been created, the MS2 spectra were compared with the GNPS spectral libraries. The minimum cosine score required for a match was 0.7, and the MS2 spectra had to share at least 6 fragment ions. All the available GNPS libraries (https://gnps.ucsd.edu/ProteoSAFe/libraries.jsp) were interrogated as untargeted toxicological screening must be as broad as possible, and research on molecules derived from natural products or pesticides may be of interest. Furthermore, the GNPS libraries include MassBank[13] and Human Metabolome Database,[14] which are relevant for the analysis of human biological samples.

3.4 | Visualization and analysis of molecular networks

Molecular networks and MS1- and MS2-level spectral annotations were viewed offline with Cytoscape 3.5.1 (http://www.cytoscape.org). Nodes were labeled with the mass of the precursor ion, and edges were labeled with the mass shift between the two connected nodes. In a multisample network, nodes were colored using a pie chart in which a different color was assigned to each sample. This means that (a) the pie chart for the node contains as many colors as there are samples containing the compound, and (b) that the size of each section of the chart will be proportional to the corresponding peak area for the compound.

First, initial information was provided by an overall visual analysis of the network, since the complexity of a matrix is proportional to the size of the network. Secondly, one focuses on clusters (forming sub-networks) and the nodes. In particular, one should note the distribution of the various colors assigned to the samples. In a second phase, the nodes were annotated on the basis of their chromatographic retention time (if available), exact mass, and the match between the fragmentation spectrum and the GNPS and/or mzCloud (https://www.mzcloud.org, HighChem LLC, Slovakia) spectral databases. Bibliographical information was also used for annotation.

Fragmentation spectra in Case 3 were interpreted based on fragmentation trees generated using Sirius CSI:FingerID software (https://bio.informatik.uni-jena.de/software/sirius/).[15,16]

4 | CASES STUDIES

Three relevant clinical and forensic cases were chosen to exemplify the molecular networking approach. We performed retrospective analysis of the raw data acquired in LC–HRMS/MS (Cf. Section 2). We initially focused on sample-to-sample comparison in a case of lethal self-injection (Case 1) and in a case of drug-facilitated assault (Case 2). Next, we explored the interest of molecular networking for the annotation of human metabolites of a NPS (Case 3). Full data processed through the GNPS platform are accessible through these links: Case 1 (https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=e7c5118ee1cb4a42a4de34d9b8661d5a), Case 2 (https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=f089f47cca2d4f15ac1c190aaeede97b), and Case 3 (https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=9dd76d377c1d4340a605d4ea03ea9b79). MS/MS spectra of the discussed compounds are available in Figure SI-B.

4.1 | Sample-to-sample comparison

4.1.1 | Case 1

A 50-year-old farmer was found dead in his barn. A veterinary syringe containing a translucent yellow liquid was lying near his body. An initial external examination revealed a skin lesion suggestive of an injection in the left arm. The raw spectral files of the injected solution and a femoral blood sample were processed according to the workflow described in Section 2. In this case, the forensic challenge was to establish whether one or more of the compounds present in the syringe were also present in the victim's blood. With regard to a multimatrix molecular network, the question was whether or not the injected solution and the blood had one or more nodes in common.

Red and green colors were assigned to the nodes corresponding to the femoral blood and the injected solution, respectively. The size of the nodes was mapped to the peak area of the blood compounds (Figure 1A). The multimatrix molecular network displayed several clusters, constituted by 856 red nodes and 13 green nodes, reflecting the relative complexity of the two studied matrices. The visual inspection provided additional information: Only one node was both red and green indicating the presence of the same molecule in the victim's blood and in the liquid. Other clusters of red nodes reflected the endogenous molecules present in the blood matrix. This unique node was annotated as the quaternary ammonium pesticide chlormequat, with a match for the exact mass (*m/z* 122.0734 $[M]^+$) and retention time. The identification of the analyte was further supported by the typical isotopic pattern of a monochlorinated ion, and a comparison with the mzCloud experimental spectral database. Lastly, the circumstances under which the body was discovered were similar to those reported in another case of chlormequat poisoning.[17] This result was consistent with the initial screening results and the official cause of death declared to the authorities was chlormequat poisoning by suicidal injection.

4.1.2 | Case 2

A young woman reported that she had been sexually assaulted during a party but could not remember all the details of the incident (partial amnesia). Hair samples were taken six weeks later to be screened for the uptake of xenobiotics at the time of the alleged incident. The hair strand was cut into three 1 cm segments; segment 1 corresponded to the month after the incident, segment 2 covered the time of the incident, and segment 3 corresponded to the month preceding the incident. This segmentation of the hair strand helps differentiate between isolated and repeated exposures to a xenobiotic.[11] The raw spectral files of the three hair segments were processed according to the workflow described in Section 2. In this case, the forensic challenge was to determine whether one or more compounds reportedly used in drug-facilitated sexual assaults were present in hair segment 2 only and not in segments 1 or 3. With regard to a multisample molecular network, the question was whether one or more nodes were specific to hair segment 2. Orange, pink, and blue colors were assigned respectively to data from hair segments 1, 2, and 3. The multisample molecular network featured a total of 555 nodes including the three colors equitably (Figure 2A). The overall visual analysis confirmed that the three samples came from the same relatively complex matrix. The network did not contain any completely orange or completely blue nodes. However, a completely pink node stood out. It belonged to segment 2, which corresponded to the time of the alleged incident (Figure 2B). Annotation of this node matched the molecule doxylamine with regard to the exact mass (*m/z* 271.1832 $[M + H]^+$), the retention time, and spectral matching with a GNPS database (MassBank of North America). The match was confirmed by comparison with the mzCloud spectral database. Doxylamine is an H1 antihistamine that has been used in cases of drug-facilitated sexual assault, due to its sedative effects.[18] This result was consistent with the initial screening results that had not been declared to the authorities as a second (thinner) hair strand was negative.

4.2 | NPS metabolism exploration

Case 1. A 17-year-old man was admitted to the emergency department for altered consciousness and agitation. The patient had a history of substance abuse, and a small bag of white powder labelled "3-MeO-PCP" was found in his belongings. The young man presented alternating periods of hypotonia and hypertonia. His pupils were partly dilated and reactive to light. The patient was transferred to the intensive care unit. The main complication was rhabdomyolysis, with slightly elevated serum creatine kinase levels. The patient was discharged on day two. Urine and blood samples were sent to our laboratory for analysis. The spectral files of the two samples were processed according to the workflow described in Section 2. 3-methoxyphencyclidine (3-MeO-PCP) was identified in the patient's urine and blood with regard to the exact mass (*m/z* 274.2188 $[M + H]^+$), the retention time, and spectral matching with a commercial 3-MeO-PCP standard. The node annotated as 3-MeO-PCP was found within a cluster of molecules whose fragmentation spectra were similar enough (cosine score > 0.7) to be grouped together, indicating the potential similarity of their chemical structures (Figure 3). These nodes were annotated by rationalization of the observed mass shifts within the cluster, manual inspection of the fragmentation spectra, and comparison with the data from a recent study in which the in vitro metabolism (pooled human liver microsomes) and in vivo (rat urine) metabolism of 3-MeO-PCP was studied.[19] In order to make the graphic representation easier to interpret, only previously described metabolites and five unknown metabolites were depicted (see the GNPS link for the complete dataset).

In this case of drug intoxication, we found 12 previously described metabolites of 3-MeO-PCP. There were seven phase I metabolites (the piperidine-hydroxy, cyclohexyl-hydroxy, cyclohexyl-hydroxy piperidine-hydroxy, *O*-demethyl, *O*-demethyl piperidine-hydroxy, *O*-demethyl piperidine-dihydroxy and piperidine-dihydroxy metabolites) and five phase II metabolites (the piperidine-hydroxy glucuronide, *O*-demethyl glucuronide, *O*-demethyl piperidine-hydroxy glucuronide, *O*-demethyl piperidine-dihydroxy glucuronide and *O*-demethyl aryl-hydroxy glucuronide metabolites). As it is the case in the work of Michely et al, we did not find a phase I metabolite corresponding to

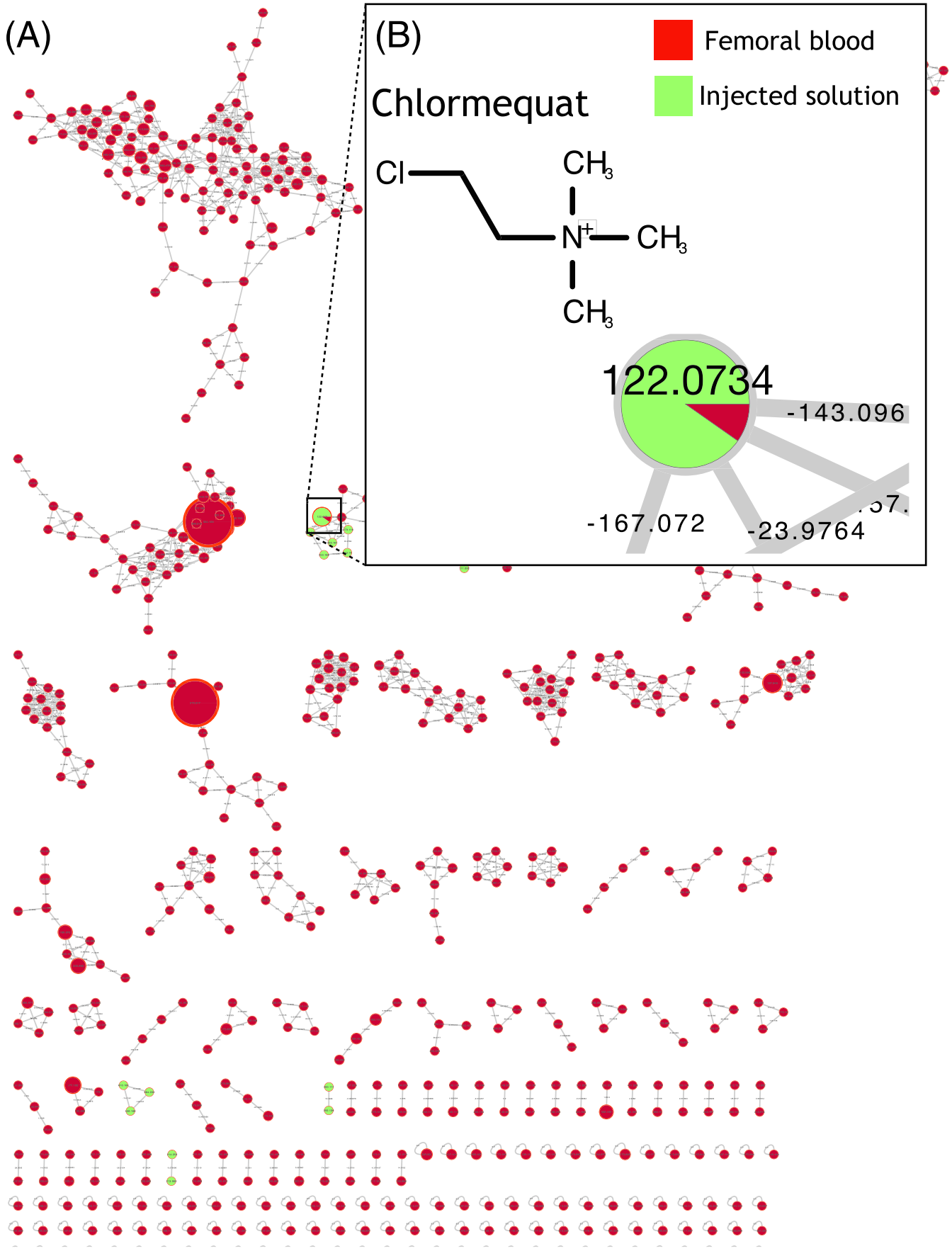

FIGURE 1 A, The overall molecular network, composed of nodes from the blood sample and the injected liquid. B, A magnified inset showing the chlormequat node. The node size is mapped to the peak area of the blood compounds [Colour figure can be viewed at wileyonlinelibrary.com]

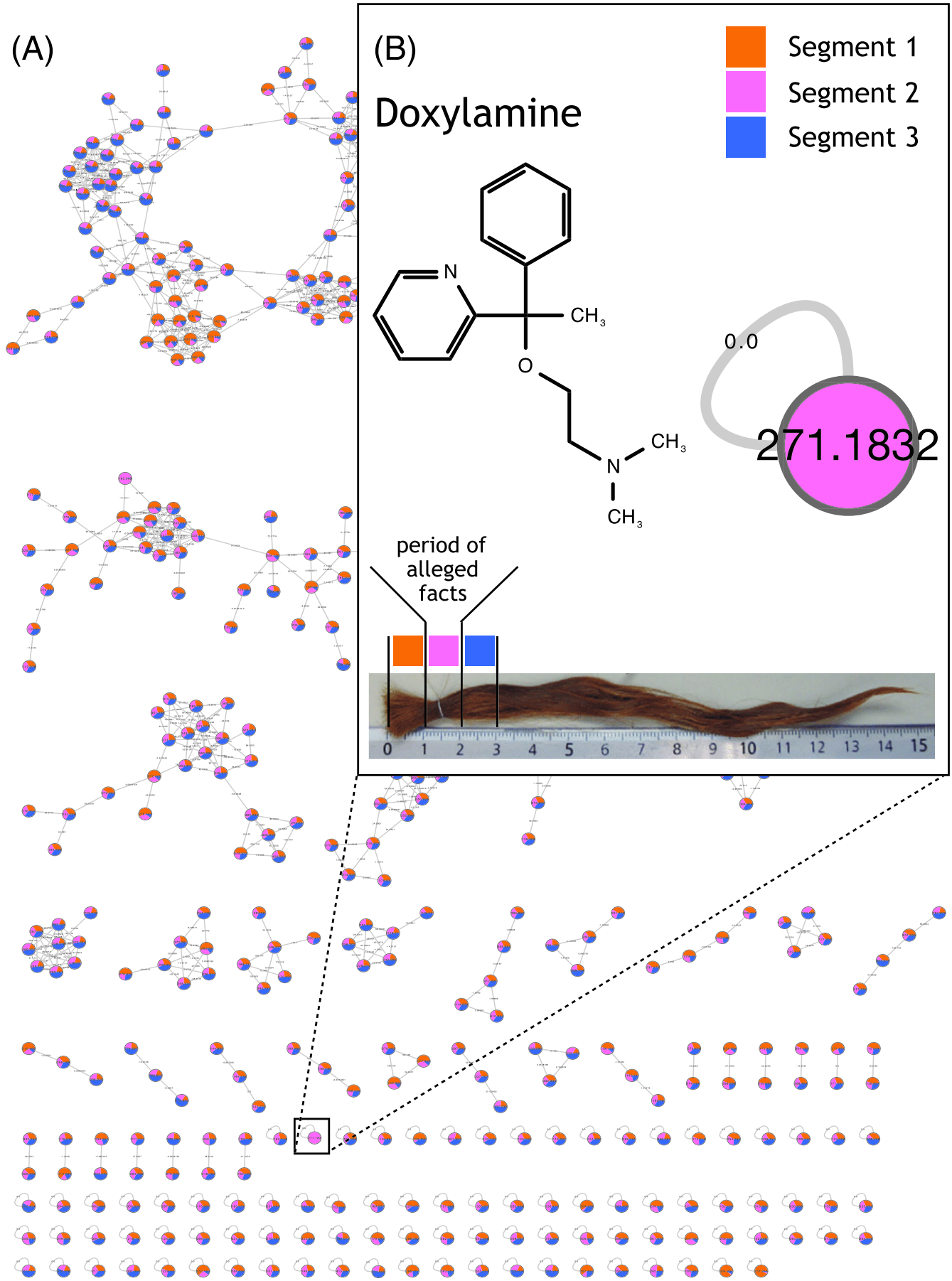

FIGURE 2 A, The overall molecular network, composed of clusters and nodes from the three hair segments. B, A magnified inset showing the doxylamine node [Colour figure can be viewed at wileyonlinelibrary.com]

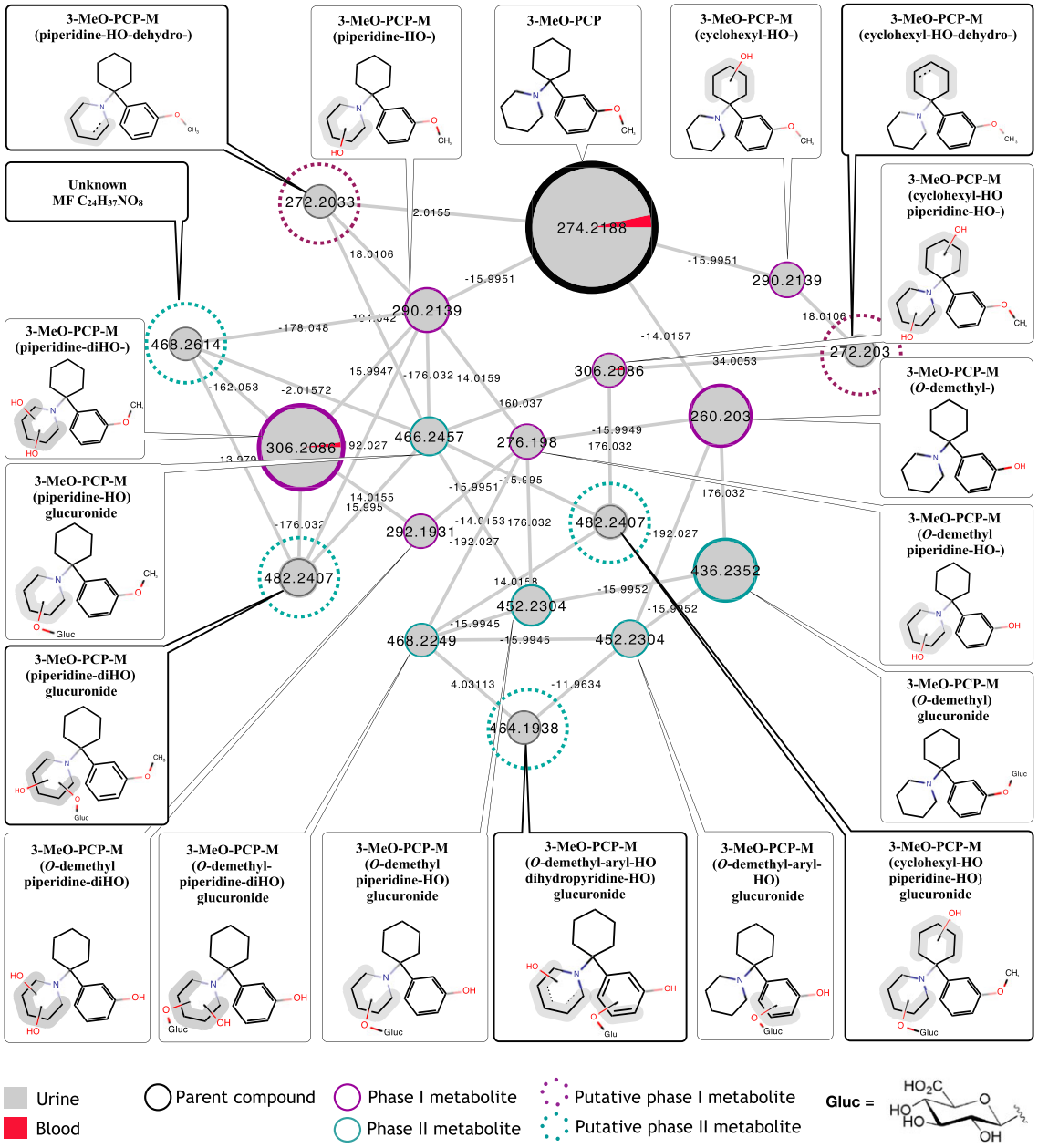

FIGURE 3 Selected nodes in the 3-MeO-PCP cluster in urine and blood, showing the parent compound linked with 12 known metabolites and 5 unreported metabolites (suggested structures) [Colour figure can be viewed at wileyonlinelibrary.com]

the putative *O*-demethyl aryl-hydroxy glucuronide.[19] Michely et al suggested that "the corresponding phase I metabolite could not be detected, either due to low concentrations or insufficient conjugate cleavage during work-up procedure for this particular metabolite." Our present findings support the first hypothesis of a low concentration as we did not use glucuronidase in our work-up procedure.

Structures of the unreported metabolites were established using the mass shifts between individual nodes. Analyses of the fragmentation patterns was performed as following to confirm the proposed structures. Mass spectra and fragmentation trees are accessible in Figure SI-C. Starting from *m/z* 274.21 (3-MeO-PCP), two mass shifts of 15.995 indicated hydroxylation of 3-MeO-PCP on the cyclohexyl moiety and on the piperidine moiety for features at *m/z* 290.21 at 6.7 minutes (cyclohexyl-hydroxy metabolite) and 7.9 minutes (piperidine-hydroxy metabolite), respectively. This was confirmed by the presence of fragments of MF $C_5H_{11}N$ (*m/z* 86.09) and MF $C_5H_{11}NO$ (*m/z* 102.09) in their fragmentation spectra, respectively (SI-C1 and SI-C2). From these two nodes, two mass shifts of 18.01 indicated a loss of H_2O leading to two nodes at *m/z* 272.20 at 9.2 minutes and *m/z* 272.20 at 10.05 minutes. Feature *m/z* 272.20 at 9.2 minutes was assigned to cyclohexyl-hydroxy dehydro metabolite thanks to its direct connectivity with cyclohexyl-hydroxy metabolite and the presence of a fragment of MF $C_5H_{11}N$ (*m/z* 84.08), indicative of a piperidine moiety and fragment of MF $C_{13}H_{14}O$ (*m/z* 187.11) indicative of a reduced 3-methoxy-phenyl-cyclohexyl fragment (SI-C3). Feature *m/z* 272.20 at 10.05 was assigned to piperidine-HO-dehydro-3-MeO-PCP thanks to its direct connectivity with piperidine-HO-3-MeO-PCP and the presence of a fragment of MF C_5H_9N (*m/z* 84.08), indicative of a reduced piperidine moiety (SI-C4). Both features at *m/z* 306.21 at 5.4 minutes and *m/z* 306.21 at 8.4 minutes were

assigned to cyclohexyl-hydroxy piperidine-hydroxy metabolite and piperidine-dihydroxy metabolite, respectively. This was confirmed by the presence of fragments of MF $C_5H_{11}NO$ (*m/z* 102.09) and MF C_5H_9N (*m/z* 84.08) indicating the presence of a piperidine-hydroxy moiety and by the presence of a fragment of MF $C_{13}H_{14}O$ (*m/z* 187.11) indicative of a reduced 3-methoxy-phenyl-cyclohexyl fragment in the fragmentation spectra of cyclohexyl-hydroxy piperidine-hydroxy metabolite (SI-C5). Likewise, the structure of piperidine-dihydroxy metabolite (*m/z* 306.21 at 8.4 minutes) was confirmed by fragments of MF C_5H_9NO (*m/z* 100.07) and MF C_5H_7N (*m/z* 82.07) in its fragmentation spectrum (SI-C6). From these two nodes, two mass shifts of 176.032 indicated the addition of a glucuronic acid leading to two features at *m/z* 482.24 at 5.54 minutes and *m/z* 482.24 at 6.94 minutes. Fragmentation patterns showed the same characteristic fragments as those observed in the corresponding phase I metabolites (SI-C7 and SI-C8). Feature *m/z* 468.23 at 4.7 minutes was assigned to O-demethyl piperidine-dihydroxy glucuronide. This was confirmed by the presence of fragments of MF $C_{12}H_{12}O$ (*m/z* 173.09) indicating the presence of 3-hydroxy-phenyl-cyclohexyl fragment and fragment of MF $C_5H_{11}NO$ (*m/z* 102.09) (SI-C9). Starting from this node, a mass shift of 4.03 was indicative of the loss of 4H for the feature *m/z* 464.19 at 5.13 minutes. This feature was annotated as O-demethyl aryl-hydroxy dihydropyridine-hydroxy glucuronide. This was confirmed by the presence of a fragment of MF C_5H_7NO (*m/z* 98.06) indicative of a dihydropyridinol moiety and a fragment at MF $C_{12}H_{12}O$ (*m/z* 173.09) indicative of a 3-hydroxy-phenyl-cyclohexyl moiety (SI-C10). The suggested Markush structures of these five newly described metabolites are shown in Figure 3.

Within a molecular network, the links between nodes reflect not only spectral similarity between ions but also the metabolic processes occurring in biological samples. Mapping of the node color to the sample type provides additional information. (i) All metabolites were detected in urine. (ii) The parent compound, the piperidine-dihydroxy metabolite, and the cyclohexyl-hydroxy piperidine-hydroxy metabolite were also detected in blood. Mapping of the node size to the peak area for metabolites in urine provides semi-quantitative information and in the present case, indicated that the metabolite with the most intense signal in urine is the piperidine-dihydroxy compound.

5 | DISCUSSION

With hundreds or thousands of MS2 fragmentation scans obtained during untargeted data-dependent MS/MS experiments and the simultaneous processing of several samples, the amount of analytical data can be overwhelming. Advanced spectral analysis tools are needed to fully exploit structural information present in the fragmentation data; the molecular networking approach provides a means of extracting important information from such large, complex data sets.

In all three cases described, the combination of several bioinformatics tools allowed us to efficiently compare samples and explore the spectra acquired during an untargeted MS/MS experiment. In Cases 1 and 2, the molecular networking approach gave the same results as the usual approach. Those cases exemplified the ability of multisamples molecular networks to efficiently compare datasets by highlighting common features and differences between samples. In Case 2, no doxylamine metabolites were detected which is consistent with the fact that substances incorporated in the hair usually correspond to the parent molecules and less to their metabolites. In a recent drug-facilitation sexual assault (DFSA) case report involving doxylamine, the authors only quantified doxylamine in hair matrix.[20] Our last case (Case 3) shows that clustering by spectral similarity (a) improves the annotation and classification of previously described metabolites, and (b) facilitates the identification of new metabolites by indirectly representing metabolic pathways. However, thorough manual interpretation and validation remain essential for the reliable evaluation of metabolite structures. The unified graphical representation of multiple metadata (exact mass, retention time, annotations, semi-quantitative information, mass shifts, and structural relationships/metabolic pathways) is a major strength for metabolite annotation.

The molecular networking approach has already been used for the chemical exploration of biological matrices. Here, we assessed the relevance of the approach by applying it to three representative forensic cases. Our results show that the use of molecular networking opens new perspectives in analytical toxicology in general and drug screening in biological matrices using LC–HRMS/MS in particular. However, our present work must be extended, in order to optimize the acquisition parameters and in particular evaluate the impact of various collision energies on spectral clustering. The MS settings (such as electrospray polarity, source voltage, collision energy, resolution, range for acquisition, and maximum injection time) must be optimized for each mass spectrometer, in order to avoid common pitfalls and provide high-quality data. In our study we observed that a resolution of 70 000 is not optimal since it is obtained at the expenses of lower MS/MS acquisition scan rates. For optimal molecular networks generation, a preferred acquisition method would use a 35 000 resolution in MS1 and 17 500 in MS2 (allowing a short duty cycle, <50 ms for an AGC target of 1E6) combined with a separate acquisition of positive mode and negative mode data. Since the goal of the project was to evaluate the data treatment methods rather than the MS acquisition parameters we nevertheless decided to keep them identical to the routine conditions used in the laboratory. The interpretation of molecular networking cannot surpass classical challenges in MS.[21] In addition to these analytical settings, MZmine data preprocessing requires careful adjustments of the parameters at each step. This data processing method is complementary to usual approaches in forensics and analytical toxicology, and enables the grouping of sets of structurally related compounds (drugs and their metabolites, for example) in the absence of prior knowledge of their fragmentation behavior. These free tools are becoming more user-friendly, and their use does not require advanced bioinformatics skills. Furthermore, the "spectral tolerance" implied by spectral similarity allows this approach to accommodate datasets produced by different mass spectrometers, ionization modes, and LC-based methods.[22]

Annotation is performed during the molecular networking generation process. The full GNPS spectral libraries are available at the following address: https://gnps.ucsd.edu/ProteoSAFe/libraries.jsp. They were built by the merging of various previously available spectral libraries (eg, part of NIST, HMDB, MassBank) and by the

contribution of GNPS users. Those latest contributions are compiled as the GNPS libraries. Thus, GNPS is also a living MS repository in addition to a data treatment platform. As this web platform was initially developed by the natural product research community there is a lack of NPS spectra in the libraries. In those cases, annotation has to be done manually.

6 | CONCLUSION

Bioinformatic methods are increasingly used to structure untargeted MS/MS datasets. A combination of in silico metabolite prediction with in silico fragmentation could help to annotate NPS metabolites. This type of combined strategies has already been proposed to extend the annotation potential of an in silico fragmentation database in natural products research.[23,24] These approaches may lead to additional metabolite annotations and complement in vitro conventional studies. It should nevertheless be borne in mind that a comparison with reference standards remains essential for safely addressing legal toxicological issues.

According to the three clinical and forensic cases presented herein, molecular networking appears to be of value in the analysis of complex data obtained by untargeted toxicological screening methods. Important information can be obtained prior to the annotation stage by observing the distribution of structurally linked molecules within various samples. Molecular networking and its potential applications in the field of toxicological analysis should be studied further as they may offer interesting perspectives.

ORCID

Sophie Allard https://orcid.org/0000-0001-5540-6372
Pierre-Marie Allard https://orcid.org/0000-0003-3389-2191
Thomas Gicquel https://orcid.org/0000-0001-7448-9732

REFERENCES

1. Rentsch KM. Knowing the unknown – state of the art of LCMS in toxicology. *TrAC Trends Anal Chem*. 2016;84:88-93.
2. Maurer HH, Meyer MR. High-resolution mass spectrometry in toxicology: current status and future perspectives. *Arch Toxicol*. 2016;90(9):2161-2172.
3. Oberacher H, Arnhard K. Compound identification in forensic toxicological analysis with untargeted LC-MS-based techniques. *Bioanalysis*. 2015;7(21):2825-2840.
4. Pasin D, Cawley A, Bidny S, Fu S. Current applications of high-resolution mass spectrometry for the analysis of new psychoactive substances: a critical review. *Anal Bioanal Chem*. 2017;409(25):5821-5836.
5. Meyer MR, Maurer HH. Review: LC coupled to low- and high-resolution mass spectrometry for new psychoactive substance screening in biological matrices – where do we stand today? *Anal Chim Acta*. 2016;927:13-20.
6. Mercolini L, Protti M. Biosampling strategies for emerging drugs of abuse: towards the future of toxicological and forensic analysis. *J Pharm Biomed Anal*. 2016;130:202-219.
7. Wang M, Carver JJ, Phelan VV, et al. Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nat Biotechnol*. 2016;34(8):828-837.
8. Van Der Hooft JJJ, Padmanabhan S, Burgess KEV, Barrett MP. Urinary antihypertensive drug metabolite screening using molecular networking coupled to high-resolution mass spectrometry fragmentation. *Metabolomics*. 2016;12:125.
9. Quinn RA, Nothias LF, Vining O, Meehan M, Esquenazi E, Dorrestein PC. Molecular networking as a drug discovery, drug metabolism, and precision medicine strategy. *Trends Pharmacol Sci*. 2017;38:143-154.
10. Gicquel T, Hugbart C, Le Devehat F, et al. Death related to consumption of Rauvolfia sp. powder mislabeled as Tabernanthe iboga. *Forensic Sci Int*. 2016;266:e38-e42.
11. Cooper GAA, Kronstrand R, Kintz P. Society of Hair Testing. Society of Hair Testing guidelines for drug testing in hair. *Forensic Sci Int*. 2012;218(1–3):20-24.
12. Olivon F, Grelier G, Roussi F, Litaudon M, Touboul D. MZmine 2 data-preprocessing to enhance molecular networking reliability. *Anal Chem*. 2017;89(15):7836-7840.
13. Horai H, Arita M, Kanaya S, et al. MassBank: a public repository for sharing mass spectral data for life sciences. *J Mass Spectrom*. 2010;45(7):703-714.
14. 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.
15. Böcker S, Dührkop K. Fragmentation trees reloaded. *J Chem*. 2016;8(1):5.
16. Dührkop K, Shen H, Meusel M, Rousu J, Böcker S. Searching molecular structure databases with tandem mass spectra using CSI:FingerID. *Proc Natl Acad Sci U S a*. 2015;112(41):12580-12585.
17. Boumrah Y, Gicquel T, Hugbart C, Baert A, Morel I, Bouvet R. Suicide by self-injection of chlormequat trademark C5SUN(®). *Forensic Sci Int*. 2016;263:e9-e13.
18. Anderson LJ, Flynn A, Pilgrim JL. A global epidemiological perspective on the toxicology of drug-facilitated sexual assault: a systematic review. *J Forensic Leg Med*. 2017;47:46-54.
19. A Michely JA, Manier SK, Caspar AT, Brandt SD, Wallach J, Maurer HH. New psychoactive substances 3-Methoxyphencyclidine (3-MeO-PCP) and 3-Methoxyrolicyclidine (3-MeO-PCPy): metabolic fate elucidated with rat urine and human liver preparations and their detectability in urine by GC-MS, "LC-(high resolution)-MSn" and "LC-(high resolution)-MS/MS". *Curr Neuropharmacol*. 2017;15(5):692-712.
20. Larabi IA, Martin M, Etting I, Penot P, Fabresse N, Alvarez JC. Drug-facilitated sexual assault (DFSA) involving 4-methylethcathinone (4-MEC), 3,4-Methylenedioxypyrovalerone (MDPV), and doxylamine highlighted by hair analysis. *Drug Test Anal*. 2018;10(8):1280-1284.
21. Quinn RA, Nothias L-F, Vining O, Meehan M, Esquenazi E, Dorrestein PC. Molecular networking as a drug discovery, drug metabolism, and precision medicine strategy. *Trends Pharmacol Sci*. 2017;38(2):143-154.
22. Yang JY, Sanchez LM, Rath CM, et al. Molecular networking as a dereplication strategy. *J Nat Prod*. 2013;76(9):1686-1699.
23. Allard P-M, Genta-Jouve G, Wolfender J-L. Deep metabolome annotation in natural products research: towards a virtuous cycle in metabolite identification. *Curr Opin Chem Biol*. 2017;36:40-49.
24. Beauxis Y, Genta-Jouve G. Metwork: a web server for natural products anticipation. *Bioinformatics*. 2018.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Allard S, Allard P-M, Morel I, Gicquel T. Application of a molecular networking approach for clinical and forensic toxicology exemplified in three cases involving 3-MeO-PCP, doxylamine, and chlormequat. *Drug Test Anal*. 2019;11:669–677. https://doi.org/10.1002/dta.2550