Analytica Chimica Acta 1365 (2025) 344241

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

Overcoming limitations in exhaled breath aerosol analysis – Untargeted breathomics for the identification of endogenous breath biomarkers using LC-HRMS/MS

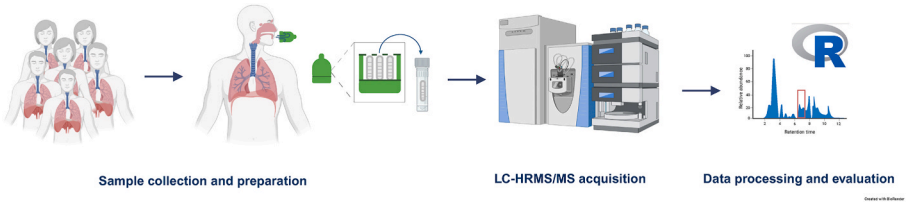
Juel Maalouli Schaar, Selina Hemmer, Sascha K. Manier, Lea Wagmann, Markus R. Meyer *

Experimental and Clinical Toxicology and Pharmacology, Center for Molecular Signaling (PZMS), PharmaScienceHub (PSH), Saarland University, Homburg, Germany

HIGHLIGHTS

- Application of an untargeted breathomics approach for exhaled breath (EB) aerosol analysis.
- Potential biomarker candidate found in healthy volunteer and patient EB samples.
- Results contribute toward standardization of EB sample collection.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Exhaled breath
Untargeted breathomics
Endogenous biomarker
LC-HRMS/MS

ABSTRACT

Background: Human exhaled breath (EB) aerosol is a suitable matrix for the non-invasive bioanalysis of both volatile but also non-volatile compounds. One of the key challenges in drug monitoring studies is the lack of standardization in EB sample collection. The aim of the present study was to use untargeted breathomics for identification of potential endogenous biomarkers in EB that could address these challenges.
Results: The Breath Explor impaction device was used to collect microparticles in EB aerosol from six healthy volunteers and six different sampling time points over a 1-h period (total number of samples = 36). Sample preparation consisted of an extraction with methanol and analysis was performed using reversed-phase and hydrophilic interaction-phase liquid chromatography coupled to high-resolution mass spectrometry. Results revealed the presence of a feature that could be reliably detected across all samples, thus representing a potential biomarker for standardizing EB aerosol sample collection. Although an annotation of this feature was not possible, it showed robustness and reproducibility in EB samples from healthy volunteers and could also be successfully detected in patient samples.
Significance and novelty: This study represents a step forward in overcoming the current limitation of EB analysis, the lack of standardization in sampling procedure. Nevertheless, further investigation including correlation studies and more diverse populations will be essential for the confirmation of the clinical relevance and application of this biomarker.

* Corresponding author. Experimental and Clinical Toxicology and Pharmacology, Center for Molecular Signaling (PZMS), PharmaScienceHub (PSH), Saarland University, Homburg, Germany
E-mail address: markus.meyer@uni-saarland.de (M.R. Meyer).

https://doi.org/10.1016/j.aca.2025.344241
Received 4 March 2025; Received in revised form 29 April 2025; Accepted 22 May 2025
Available online 23 May 2025
0003-2670/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Human exhaled breath (EB) has emerged as a promising alternative and innovative sample matrix that allows the non-invasive analysis of both volatile and non-volatile compounds [1–4]. Both follow a similar *in vivo* pathway prior to exhalation, which is generally characterized by diffusion from the bloodstream into the airway fluid [2,5]. However, the mechanism by which volatile and non-volatile compounds are exhaled differs due to their different physico-chemical properties. Volatile compounds are primarily exhaled through a liquid-air partitioning process occurring within the alveoli [3,6,7]. In contrast, non-volatile compounds, including certain drugs or drugs of abuse, are exhaled in the form of aerosolized microparticles that originate from the lining fluid of the airways. These microparticles are formed with each exhalation, encapsulating non-volatile compounds, and enabling their transport into the exhaled breath [8–12]. The selective detection of non-volatile compounds in these exhaled microparticles has been achieved by advanced collection devices, such as adsorption- or impaction-based systems. These devices capture microparticles from EB aerosol, allowing for the targeted analysis of specific non-volatile compounds and providing a powerful tool for non-invasive monitoring of drugs or drugs of abuse intake, therapeutic drug levels or exposure to various chemicals [3,9,11–16]. In this study, an impaction device is used. This device is equipped with three parallel impaction collectors with cylindrical polyhedral geometry and eight baffles arranged in a sequential order. The collectors are made of medically acceptable materials and designed to capture aerosol particles from the distal lung. Additionally, the device used is portable, facilitating practical use in a variety of settings, such as clinical and non-clinical environment. Its applicability in the detection of drugs and drugs of abuse in human EB aerosol has already been demonstrated in previous studies [11,12,15,17,18].

Despite the potential of EB analysis, the lack of standardization of the sample collection process remains a significant limitation, since there is no control over the total amount of EB passing through the device and actually retained on the collector. One promising solution might be the identification of an endogenous non-volatile biomarker, which can serve as a reference for the collection of EB. Such a marker might ideally correlate to the EB amount collected and allow a direct correlation of drug quantity with EB amount. For this purpose, simultaneous detection using one analytical workflow is of particular importance and represents an essential key element in the application of EB aerosol as a matrix in drug excretion and monitoring studies [12,19,20].

Untargeted metabolomics is a well-established approach for the identification of novel endogenous and exogenous biomarkers in various biological matrices [21–24]. The field of untargeted breathomics, a sub-discipline of metabolomics focusing on EB, has predominantly been applied for the identification of volatile endogenous compounds using Tedlar bags, sorbent tubes, or needle trap devices. Such analysis have been investigated primarily in the context of diseases, including respiratory infections. Besides volatile organic compounds, non-volatile compounds have been detected using single barrel electrostatic filter device and direct infusion Fourier transform ion cyclotron resonance mass spectrometry [25–29]. However, the available data on non-volatile endogenous biomarkers in human EB are limited and insufficient to address the aforementioned limitations. Therefore, the present study aims to use untargeted breathomics in EB aerosol of healthy volunteers using liquid chromatography coupled to high-resolution mass spectrometry (LC-HRMS) to identify a suitable non-volatile endogenous biomarker helping to overcome existing limitations.

2. Experimental section

2.1. Chemicals and materials

D-Glucose-d_7, glutamic acid-d_3, palmitic acid-d_{31}, stearic acid-^{13}C, and succinic acid-d_3 were purchased from Merck (Darmstadt, Germany). L-tryptophan-d_5 was obtained from Alsachim (Illkirch-Graffenstaden, France). Acetonitrile (ACN), ammonium formate, ammonium acetate, ethanol, formic acid, methanol, and all other chemicals (LC-MS or analytical grade) were purchased from VWR (Darmstadt, Germany). The water was purified in-house with a Millipore filtration unit (18.2 Ω x cm water resistance). Breath Explor sampling devices were from Munkplast AB (www.munkplast.se, Uppsala, Sweden).

2.2. EB aerosol sample collection

Sample collection was conducted using a device containing three collectors. Microparticles from EB were collected by impaction on the collectors [11]. For sampling procedure, the Breath Explor Operate Guide from Munkplast AB was used to assist the volunteers throughout the sampling procedure, including an illustrative video, a status message that described the current stage, and a timer, which indicated the recommended time for every exhalation. This study included EB aerosol samples from six volunteers (three females, three males, mean age 32, aged 24–43 years). Each volunteer provided their informed consent. A total of six samples were collected from each volunteer over a 1-h period. Therefore, every 10 min, twelve expiration cycles were collected into one sampling device. The cycles consisted of exhalation, holding the breath, inhalation, placing the device into the mouth, and then exhalation through the device. EB samples were collected from all six volunteers at the same time point and were stored at −80 °C prior analysis.

2.3. Sample preparation

EB carrying collectors were removed from the device and placed into 5 mL tubes. Subsequently, 2.2 mL of methanol containing 100 µM D-glucose-d_7, 100 µM glutamic acid-d_3, 100 µM palmitic acid-d_{31}, 100 µM stearic acid-^{13}C, 20 µM succinic acid-d_3, and 50 µM L-tryptophan-d_5 as internal standards were added. After 5 min soaking time, the mixture was shaken for 5 min at 1400 rpm and centrifugated three times for 2 min at 3000×*g*. Afterwards, 2 mL extraction solvent were evaporated to dryness using a vacuum centrifuge at 1400 rpm and 24 °C. The obtained residues were reconstituted in 100 µL of a mixture of acetonitrile (containing 0.1 % formic acid) and water (80:20, v/v). After centrifugation for 2 min at 18,407×*g*, supernatants were transferred to MS vials and injected onto the liquid chromatography-high resolution tandem mass-spectrometry (LC-HRMS/MS) system. Additionally, three blank collectors free of human EB were prepared as described above. Pooled QC sample was prepared by transferring equal proportion of each EB sample into one MS vial.

2.4. Proof-of-concept samples

For proof-of-concept testing, extracts, already prepared from EB samples of patients at Saarland University Hospital, Internal medicine III, Homburg, Germany, from a previous study were used [12]. These EB extracts derived from 28 patient samples (seven females, 21 males, mean age 70, aged 27–87 years) and were stored at −20 °C up to twelve months. Further information on patients, sample collection, and sample preparation can be found elsewhere [12]. After thawing, the extracts were evaporated to dryness using a vacuum centrifuge at 1400 rpm and 24 °C for 240 min to reconstitute them in the same volume and solvent as the volunteer samples. Residues were reconstituted in 100 µL of a mixture of acetonitrile (containing 0.1 % formic acid) and water (80:20, v/v). Pooled quality control (QC) sample was prepared by transferring equal portion of each sample into one MS vial.

2.5. LC-HRMS/MS analysis

Analyses were performed using a Thermo Fisher Scientific (TF,

Dreieich, Germany) Dionex UltiMate 3000 RS pump consisting of a degasser, a quaternary pump, and an UltiMate Autosampler, coupled to a TF Q Exactive Plus high-resolution mass spectrometer equipped with a heated electrospray ionization (HESI)-II source according to previously published studies [23,30,31]. Performance of the columns and the mass spectrometer was tested before using a test mixture as described by Maurer et al. [32,33]. In addition, instrument performance was verified by detection of a methanolic mixture of all internal standards included in the method at a concentration corresponding to that in the extraction solvent before each analysis. Mass spectrometry calibration was done according to the manufacturer's recommendations using TF calibration solutions (both positive and negative) prior to each analysis. Gradient reversed-phase (RP) elution was performed on a TF Accucore Phenyl-Hexyl column (2.6 μm, 100 mm × 2.1 mm) and hydrophilic interaction-phase chromatography (HILIC) elution using a Merck (Darmstadt, Germany) SeQuant ZIC cHILIC column (3 μm, 150 mm × 2.1 mm). The mobile phases for the RP chromatography consisted of 2 mM aqueous ammonium formate containing acetonitrile (1 %, v/v) and formic acid (0.1 %, v/v, pH 3, eluent A), as well as 2 mM ammonium formate solution with acetonitrile:methanol (1:1, v/v) containing water (1 %, v/v) and formic acid (0.1 %, v/v, eluent B). The flow rate was set from 0 to 10 min to 0.5 mL min^{-1} and from 10 to 13.5 min to 0.8 mL min^{-1} using the following gradient: 0.0–1.0 min hold 99 % A, 1.0–10 min to 1 % A, 10–11.5 min hold 1 % A, and 11.5–13.5 min hold 99 % A. The gradient elution for HILIC was performed using aqueous ammonium acetate (200 mM, eluent C) and acetonitrile containing formic acid (0.1 %, v/v, eluent D). The flow rate was set to 500 mL min^{-1} using the following gradient: 0.0 min–1.0 min hold 2 % C, 1.0–5.0 min to 20 % C, 5.0–8.5 min to 60 % C, 8.5–10 min hold 60 % C, and 10–12 min hold 2 % C. The injection volume was set to 5 μL and the column oven temperature was 40 °C. For preparation and cleaning of the injection system, isopropanol:water (90:10, v/v) was used. Mass spectrometric analysis for untargeted metabolomics was carried out in accordance with a previously optimized workflow [23,30,34]. The following settings were used: wash volume, 100 μL; wash speed, 4000 nL/s; loop wash factor, 2. HESI-II source conditions were as follows: ionization mode, positive or negative, sheath gas, 60 AU; auxiliary gas, 10 AU; sweep gas, 3 AU; spray voltage, 3.5 kV in positive and −4.0 kV in negative mode; heater temperature 320 °C; ion transfer capillary temperature, 320 °C; and S-lens RF level, 50.0. The settings for full-scan data acquisition were as follows: resolution 140,000 at *m/z* 200; microscan, 1; automatic gain control (AGC) target, 5e5; maximum injection time, 200 ms; scan range, *m/z* 50–750; spectrum data type; centroid. The analysis of all study samples was performed in randomized order to identify potential instability of analyte or instrument performance with five injections of methanol for RP chromatography or eluent D for HILIC chromatography at the beginning of the sequence, followed by five injections of the pooled QC sample as described by Manier et al. [23]. One pooled QC sample injection was performed every five samples to monitor potential batch effects as described by Wehrens et al. [35]. Feature identification was performed using parallel reaction monitoring (PRM) and data-dependent MS^2 (dd-MS^2) acquisition. PRM data acquisition was performed using the following settings: resolution, 35,000 at *m/z* 200; microscans, 1; AGC target, 5e5; maximum injection time, 200 ms; isolation window, 1.0 *m/z*; collisions energy (CE), 10, 20 or 40 eV; spectrum data type, centroid. dd-MS^2 acquisition was performed using the following settings: resolution, 35,000 at *m/z* 200; AGC target, 5e5; maximum injection time, 200 ms; loop count, isolation window, 1.0 *m/z*; CE, 10, 20 or 40 eV; minimum AGC target, 8e3; exclude isotopes, on. For both PRM and dd-MS2 acquisition, inclusion list containing the monoisotopic masses and a time window of their retention time ± 60 s was used for potential biomarkers. TF Xcalibur software version 4.1 was used for data handling.

2.6. Data processing and statistical analysis

For untargeted metabolomics data processing was performed in an R environment according to previously published workflows [23,30,31]. TF LC-HRMS/MS RAW files were converted into mzXML files using ProteoWizard [36]. Optimization of the XCMS parameters was performed in accordance with a previously optimized strategy [34]. Peak picking was performed using XCMS in an R environment and annotation of isotopes, adducts, and artifacts was performed using the R package CAMERA [37,38]. Summarized overview of all peak picking and alignment parameters is provided in Table S1 in the Supporting Information. In addition, all features with a retention time below 60 s or above 600 s were excluded, as they were detected during the column acquisition or the wash phase of the chromatography. According to Wehrens et al. feature abundances with a value of zero were replaced by the lowest measured abundance as a surrogate limit of detection and subsequently log10 transformed [35]. Principal component analysis (PCA) and hierarchical clustering were used to analyze patterns in the dataset. Box-plot of volunteer samples (I-VI) with six samples per volunteer range from the first quartile (25th percentile) to the third quartile (75th percentile) of distribution representing the interquartile range (IQR). Medians (50th percentiles) are indicated by the lines across the boxes. Tukey-style whiskers extend to a maximum of 1.5 x IQR beyond the boxes. Outliers are dots above the top whisker. Names for features were adopted from XCMS using "M" followed by rounded mass and "T" followed by retention time in seconds. The R script can be found on GitHub (https://github.com/JuelMS/Breathomics.git) and the mzXML files are available via Metabolights (study identifier MTBLS11937) [39].

2.7. Feature filtering and identification

Subsequently, all features resulting from the peak picking were filtered using the generated peak list according to the following two criteria. For a feature of interest to serve as a potential robust endogenous biomarker for standardization, the feature must be present in at least 35 of the total 36 volunteer EB samples. Additionally, features present in the analyses of the blank collectors were excluded to avoid false-positive results caused by system contamination, contaminations during the sample preparation, or materials used in the EB sampling device. After applying these criteria, remaining features were identified by measuring MS/MS spectra using PRM and full scan dd-MS^2 acquisition methods as described above. After the conversion of MS/MS RAW files into mzXML files using ProteoWizard, the PRM spectra were imported to NIST MS Search version 2.3 Library. The settings for library and MS/MS search were obtained from previously published protocols [40,41]. Following libraries were used: NIST 2014 (nist_msms and nist_msm2 sub library), Wiley METLIN Mass Spectral Database, Lipidblast, MHMW, Human Metabolome Database (HMDB), and GNPS library [32,42–46]. Full scan with dd-ms^2 spectra were imported to SIRIUS version 6.0.5 to generate candidate structures using CSI:FingerID. The following computation settings were applied: Instrument, orbitrap; MS^2 mass accuracy: 5 ppm; inclusion of fall back adducts: all; molecular formular generation: de novo, bottom up; allowed elements: H, C, N, O, P; CSI:FingerID Structure Database: PubChem, COCONUT, Blood Exposome, FooDB, GNPS, HMDB, HSDB, KEGG, LOTUS, LipidMaps, Maconda, MeSH, PubMed, Supernatural, TeroMOL, YMDB [47,48]. Additionally, the in-silico fragmentation tool MetFrag (https://msbi.ipb-halle.de/MetFrag/) was applied to MS/MS data to identify potential structures.

3. Results and discussion

3.1. Quality control and internal standards

In an untargeted approach, pooled QC samples are recommended for quality control purposes. These QCs are used for optimization of peak

picking parameters, for feature identification, and to detect and correct potential batch effects to ensure reliable and comparable data sets [34, 49]. As the analytes of interest are unknown in advance in untargeted approach, various internal standards are primarily. The selection of isotope labeled internal standards was based on the consideration of endogenous compounds various chemical structures and their detectability on both columns. Furthermore, internal standards were implemented for the purpose of verifying the sample preparation. According to Hemmer et al. inspection of the total ion chromatograms of the QCs and the extracted ion chromatograms of the internal standards was done visually to confirm the quality of the measurements [50].

3.2. Untargeted breathomics

Following peak picking of the untargeted breathomics analysis, a comprehensive peak list of all detected features for both columns and ionization modes was generated and is given in the Supporting Information Table S2 (sheet 1–4). After exclusion of isotopes and adducts annotated by CAMERA, the total number of features across all 36 EB samples were 365 (RP, positive ionization mode), 40 (RP, negative ionization mode), 143 (HILIC, positive ionization mode), and 35 (HILIC, negative ionization mode) as shown in Table S3. Additional 106 features were found in 35 out of 36 EB samples. They were further considered to avoid elimination of potential biomarkers that were not labeled as features in one individual sample. Thus, a total of 471 features are further processed (see Table S3). Inclusion and processing of the blank collectors is a critical step for excluding contamination from the sampling device materials or the analytical workflow [51]. As demonstrated by the data in Table S3, this process is essential, as manual filtering for features exclusively present in the EB samples resulted in a highly reduced feature count of six (RP, positive ionization mode), zero (RP, negative ionization mode), zero (HILIC, positive ionization mode), and two (HILIC, negative ionization mode). This filtering process ensured that only features associated with the biological sample, and not with potential external contaminants, were considered for further analysis. The reduction in the number of features following the incorporation of blanks can be explained by the systematic elimination of signals from diverse sources during the experimental procedure. These compromise the material components of the collection device, residues of material utilized, solvents and reagents used during the extraction process, and eluents or carry-over effects introduced during the mass spectrometric analysis. Although a substantial reduction of features was achieved, it was in accordance with expectations. The purpose of blank collector analysis is to differentiate between authentic biological signals and artifacts caused by the experimental setup. Therefore, the elimination of such features is consistent with the intended quality control objectives and ensures the robustness of subsequent analysis. The remaining features for RP and HILIC separation and positive ionization are given in Table S4. To validate the dataset, two approaches were investigated. One approach was the use of principal component analysis (PCA) to assess the differences between authentic EB samples, QCs, and blank collector samples. The PCA demonstrates a clear separation between EB samples and blank collector samples for RP positive ionization mode (Fig. 1), thereby confirming that the detected features were primarily derived from the biological matrix. This validation step was important for the subsequent investigation of the remaining features. In contrast no distinct separation was observed after RP and HILIC separation and negative ionization (Figures S1, S2 and S3). The absence of differentiation may be attributed primarily to the significantly reduced number of features detected in these analyses compared to those observed in after RP and positive ionization, which directly influenced the ability of PCA to distinguish between the mentioned samples. An additional approach to assess the quality of the data was to consider differences between male and female volunteers. For RP separation and positive ionization, where most features was detected, a trend in the separation of genders could be observed. However, this potential gender-specific grouping is not clear due to the presence of outliers as shown in Fig. S4. It is important to note that hormonal fluctuation in females can influence metabolomic profiles, potentially contributing to variability in EB samples [52]. In contrast, insufficient data were available for the RP separation and negative ionization and HILIC separation to compare genders [53].

3.3. Feature identification

After adjustment of the measured PRM spectra for the remaining eight features (Table S4) using NIST MS Search and a comprehensive selection of spectral libraries no match could be identified for any of the selected features. To complement the identification of unknown features and enhance the analytical capacity using dd-MS^2 spectra, SIRIUS 6.0.5

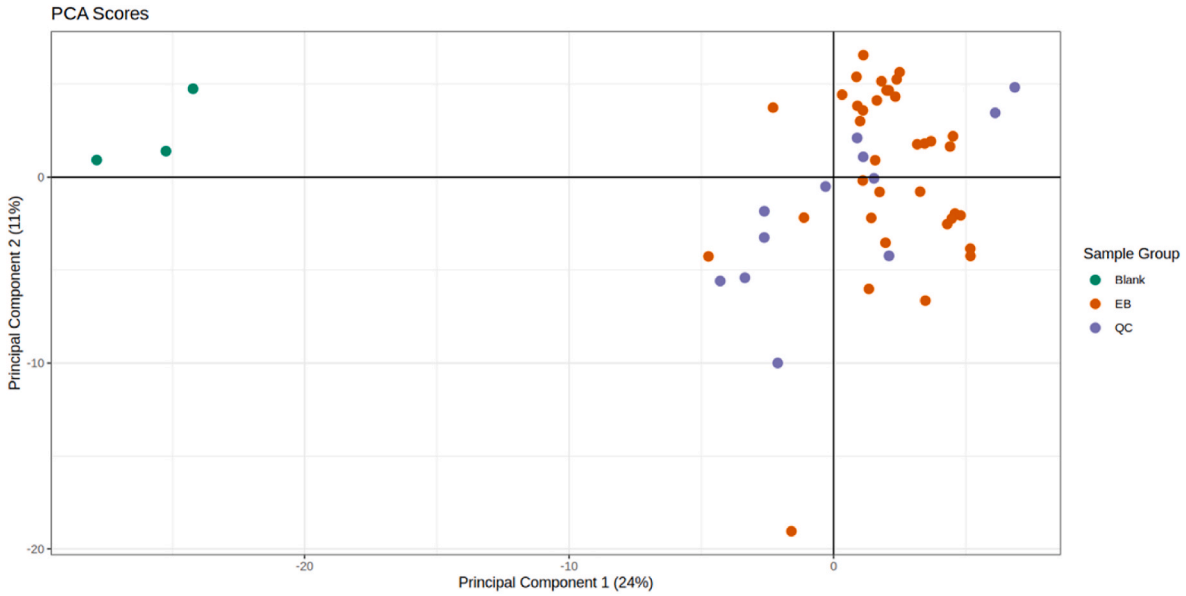


Fig. 1. Principal component analysis (PCA) after analysis using reversed-phase (RP) and positive ionization mode, exhaled breath samples of six volunteers (EB), blank collector samples (Blank), and pooled quality control (QC) injections.

software was used [47,48]. This computational analysis determines molecular formulas based on fragmentation trees and CSI:FingerID to propose potential structures by matching them against molecular databases. Nevertheless, no matches were obtained for the remaining features. This outcome highlights the challenges associated with identifying unknown biomarkers in untargeted workflows, particularly in rarely studied matrices such as human EB. Nevertheless, feature M331T465 (*m/z* 331.2972 and retention time of 465 s), as detected after RP and positive ionization mode, permitted the prediction of a plausible compound class using MS Search and SIRIUS 6.0.5. Manual inspection of the spectra revealed that a part of the spectrum of feature M331T465 as given in Fig. 2 showed similarities to those of fatty acids. Moreover, SIRIUS software suggested classification as a long-chain fatty acid compound. This identification is a level 3 identification according to the Metabolomics Standards Initiative (MSI), as this corresponds to an annotation of putatively characterized compound classes based on characteristic physicochemical properties of a chemical class of compounds, or by spectral similarity to known compounds of a chemical class [54]. This categorization is appropriate when considering the physiological composition of the respiratory system. Fatty acids, as essential components of pulmonary surfactants and cell membrane, are potential non-volatile compounds that can be exhaled in breath aerosol [19,55].

3.4. Endogenous biomarker candidate

The biological relevance of the so far unidentified feature M331T465 is one criterion for its potential use as an endogenous biomarker for standardizing EB sample collection. To further investigate this feature, its presence in each EB sample was studied. To facilitate visualization of the variability across all six EB samples per volunteers and the variability of the feature across the volunteers (I to VI) a boxplot of peak areas without normalization (Fig. S5) and a boxplot of total ion count (TIC)-normalized peak areas (Fig. S6) were generated. TIC normalization was applied to provide a standardized comparison base for each sample. Assessing the variability of the feature M331T465 across the volunteers is important, as a potential biomarker must demonstrate consistency across all volunteers. As shown in the TIC-normalized boxplot (Fig. S6), the median values were found to be consistent across most volunteers, ranging between 2×10^{-5} and 6×10^{-5} normalized areas units. However, for volunteer VI notable variability was observed, indicated by an increased interquartile range and the presence of outliers which could be a potentially biological variability. One outlier could be explained by a reported issue during sample collection. Volunteer VI indicated sneezing while exhaling during sample collection of one sample, which may have influenced the peak area in this sample. In contrast, the results of peak areas without normalization (Fig. S5) showed broader variability across the volunteers, particularly for volunteers IV and V with large interquartile ranges and outliers. Additionally, the median values of the non-normalized areas were less consistent, highlighting the significance of data normalization for reducing technical variability. It is critical to consider not only the variability across all volunteers but also its consistency within an individual. For a biomarker to serve as a reliable standard, its intra-individual variability is crucial. Nevertheless, median values for each volunteer were consistent. Reproducibility and consistency suggest that this feature has the potential to serve as an endogenous biomarker candidate and could enable the standardization of EB analysis.

3.5. Proof-of-concept testing

The potential of the filtered features as biomarkers was evaluated using extracts from 28 EB aerosol samples, which were prepared and analyzed as described above. The results of the targeted analysis for the eight features (Table S4) demonstrated that only the feature with an *m/z* 331.2972 could be detected in all patient samples. One reason for the non-detection of the remaining features in the patient samples could be their instability since the extracted samples were stored over a 12-month period before thawing and analysis compared to freshly prepared EB volunteer samples. Other reason could be differences in the sample preparations due to the use of a previous sample preparation protocol or insufficient concentration of these features in the patient extracts. Nevertheless, this finding is an additional aspect, which supports the consideration of the feature M331T465 as a potential endogenous biomarker for standardizing EB sample collection. It is crucial to acknowledge that the patient EB samples exhibit notable discrepancies when compared to the healthy volunteers EB samples. The patient cohort was predominantly composed of individuals with multiple chronic conditions, with an age range between 27 and 87 years and a mean age of 70 years. Despite these differences, based on the results of the patient samples the feature M331T465 showed sufficient robustness to be qualitatively detected in all EB samples across different populations.

3.6. Limitations and outlook

A structured overview of methodological strengths and limitations is provided in form of a SWOT-analysis in Table S5. This study was limited to a small cohort of six healthy volunteers and a total of six different sampling time points for tentative identification of an endogenous EB biomarker. All samples from both the volunteers and patients were collected within the same region of Germany, which may introduce a regional bias into the findings. It is necessary to expand the cohort to confirm the robustness and reproducibility of this endogenous biomarker candidate in future studies and to evaluate its consistency across different sampling devices. Nevertheless, differences between matrix and non-matrix components were observed using untargeted breathomics. The identification of the feature with M331T465

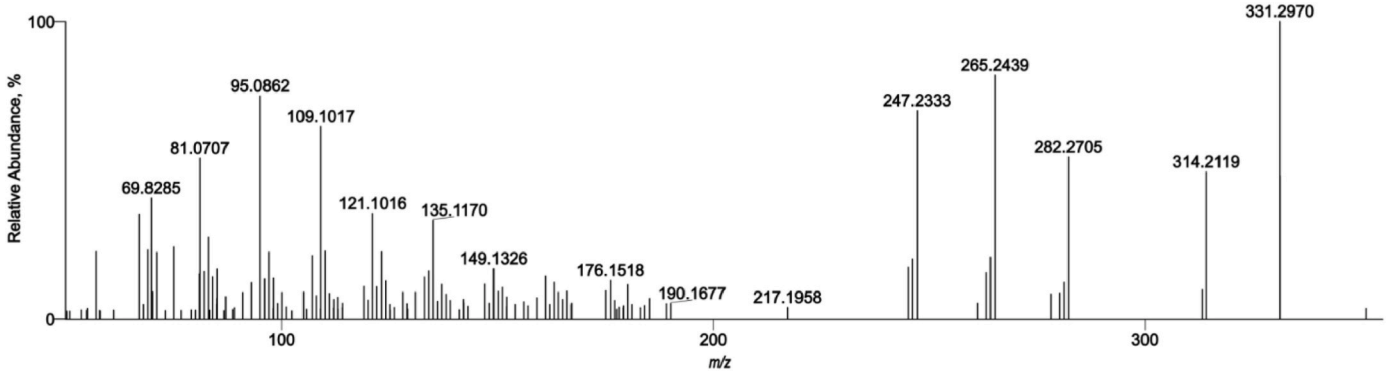


Fig. 2. Parallel reaction monitoring spectra of M331T465 after reversed-phase analysis in positive ionization mode in pooled quality control sample.

represents a key challenge for future studies, while it is also important to note that this biomarker can be applied as an incompletely identified (level 3 identification) standard, especially in matrices such as EB aerosol, where the composition is not sufficiently understood and the sample volume is difficult to control and thus the normalization and quantification in EB studies [54]. Additionally, further investigations are required to evaluate whether this feature correlates with variations in EB aerosol amount, a factor with high relevance that influences quantification and clinical interpretation. Furthermore, simultaneous detection of the feature M331T465 and drugs within an analytical method should be investigated, to prove clinical relevance of this endogenous biomarker and to address the limitations of using EB analysis in drug monitoring studies using this analytical workflow [12,20]. However, such EB studies can be easily performed due to their non-invasive sample collection compared to other biological matrices.

4. Conclusion

The present untargeted breathomics study demonstrates the potential of an untargeted approach for the investigation, detection, and preliminary validation of a potential new endogenous biomarker in EB. Despite the limited cohort size, an endogenous biomarker candidate with m/z 331.2972 potentially belonging to the group of fatty acids was detected in both healthy volunteers and a patient cohort underlaying its robustness across different samples and sample preparation procedures. However, the distinct identification of this feature remained unresolved. Conduction of studies on a diverse range of populations will be essential for the confirmation of the clinical relevance of this feature. Furthermore, drug monitoring studies should also be assessed for the confirmation of its clinical applicability and overcome the limitations currently present in EB aerosol analysis.

CRediT authorship contribution statement

Juel Maalouli Schaar: Writing – original draft, Visualization, Methodology, Investigation, Conceptualization. **Selina Hemmer:** Writing – review & editing, Methodology, Conceptualization. **Sascha K. Manier:** Writing – review & editing, Conceptualization. **Lea Wagmann:** Writing – review & editing. **Markus R. Meyer:** Writing – review & editing, Supervision, Resources, Methodology, Conceptualization.

Notes

The authors declare no competing financial interest.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank Amin A. Weber, Tilman Arnst, K. Simon Wellenberg, Tanja M. Gampfer, Fabian Frankenfeld, Matthias Kroesen, Philipp Schippers, and Aline C. Vollmer for their support and/or helpful discussion.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2025.344241.

Data availability

Data will be made available on request.

References

[1] O. Beck, A.C. Olin, E. Mirgorodskaya, Potential of mass spectrometry in developing clinical laboratory biomarkers of nonvolatiles in exhaled breath, Clin. Chem. 62 (2016) 84–91.
[2] H. Huang, J. Yang, C. Tao, L. Hu, T. Huan, W. Zhang, K. Zhang, X. Li, Exhaled breath analysis of non-volatile drugs: towards clinical applications, TrAC, Trends Anal. Chem. 171 (2024) 117541.
[3] P. Kuban, F. Foret, Exhaled breath condensate: determination of non-volatile compounds and their potential for clinical diagnosis and monitoring. A review, Anal. Chim. Acta 805 (2013) 1–18.
[4] A. Vasilescu, B. Hrinczenko, G.M. Swain, S.F. Peteu, Exhaled breath biomarker sensing, Biosens. Bioelectron. 182 (2021) 113193.
[5] C. Berchtold, M. Bosilkovska, Y. Daali, B. Walder, R. Zenobi, Real-time monitoring of exhaled drugs by mass spectrometry, Mass Spectrom. Rev. 33 (2014) 394–413.
[6] T. Bruderer, T. Gaisl, M.T. Gaugg, N. Nowak, B. Streckenbach, S. Muller, A. Moeller, M. Kohler, R. Zenobi, On-Line analysis of exhaled breath focus review, Chem. Rev. 119 (2019) 10803–10828.
[7] M.P. Hlastala, F.L. Powell, J.C. Anderson, Airway exchange of highly soluble gases, J. Appl. Physiol. 114 (2013) 675–680, 1985.
[8] O. Beck, N. Stephanson, S. Sandqvist, J. Franck, Detection of drugs of abuse in exhaled breath from users following recovery from intoxication, J. Anal. Toxicol. 36 (2012) 638–646.
[9] O. Beck, N. Stephanson, S. Sandqvist, J. Franck, Detection of drugs of abuse in exhaled breath using a device for rapid collection: comparison with plasma, urine and self-reporting in 47 drug users, J. Breath Res. 7 (2013) 026006.
[10] M.R. Meyer, S. Rosenborg, M. Stenberg, O. Beck, First report on the pharmacokinetics of tramadol and O-desmethyltramadol in exhaled breath compared to plasma and oral fluid after a single oral dose, Biochem. Pharmacol. 98 (2015) 502–510.
[11] S. Seferaj, S. Ullah, A. Tinglev, S. Carlsson, J. Winberg, P. Stambeck, O. Beck, Evaluation of a new simple collection device for sampling of microparticles in exhaled breath, J. Breath Res. 12 (2018) 036005.
[12] J. Maalouli Schaar, M. Kunz, L. Wagmann, O. Beck, F. Mahfoud, M.R. Meyer, Studying drug excretion into exhaled breath aerosol - a workflow based on an impaction sampling device and LC-HRMS/MS analysis, Anal. Chim. Acta 1323 (2024) 342991.
[13] N. Stephanson, S. Sandqvist, M.S. Lambert, O. Beck, Method validation and application of a liquid chromatography-tandem mass spectrometry method for drugs of abuse testing in exhaled breath, J. Chromatogr., B: Anal. Technol. Biomed. Life Sci. 985 (2015) 189–196.
[14] S. Ullah, S. Sandqvist, O. Beck, A liquid chromatography and tandem mass spectrometry method to determine 28 non-volatile drugs of abuse in exhaled breath, J. Pharm. Biomed. Anal. 148 (2018) 251–258.
[15] V. Malave, K. Jeerage, E. Garboczi, T. Lovestead, 3D computational fluid and particle dynamics simulations: metrics of aerosol capture by impaction filters, J. Breath Res. 18 (2023) 016002.
[16] M.A.G. Wallace, J.D. Pleil, Evolution of clinical and environmental health applications of exhaled breath research: review of methods and instrumentation for gas-phase, condensate, and aerosols, Anal. Chim. Acta 1024 (2018) 18–38.
[17] A.C. Almstrand, E. Ljungstrom, J. Lausmaa, B. Bake, P. Sjovall, A.C. Olin, Airway monitoring by collection and mass spectrometric analysis of exhaled particles, Anal. Chem. 81 (2009) 662–668.
[18] K. Feltmann, T.H. Elgan, M. Bottcher, S. Lierheimer, S. Hermansson, O. Beck, J. Gripenberg, Feasibility of using breath sampling of non-volatiles to estimate the prevalence of illicit drug use among nightlife attendees, Sci. Rep. 12 (2022) 20283.
[19] S. Ullah, S. Sandqvist, O. Beck, Measurement of lung phosphatidylcholines in exhaled breath particles by a convenient collection procedure, Anal. Chem. 87 (2015) 11553–11560.
[20] J.D. Pleil, M.A.G. Wallace, M.C. Madden, Exhaled breath aerosol (EBA): the simplest non-invasive medium for public health and occupational exposure biomonitoring, J. Breath Res. 12 (2018) 027110.
[21] A. Agin, D. Heintz, E. Ruhland, J.M. Chao de la Barca, J. Zumsteg, V. Moal, A. S. Gauchez, I.J. Namer, Metabolomics – an overview. From basic principles to potential biomarkers (part 1), Méd. Nucléaire 40 (2016) 4–10.
[22] X. Liu, J.W. Locasale, Metabolomics: a primer, Trends Biochem. Sci. 42 (2017) 274–284.
[23] S.K. Manier, A. Keller, J. Schaper, M.R. Meyer, Untargeted metabolomics by high resolution mass spectrometry coupled to normal and reversed phase liquid chromatography as a tool to study the in vitro biotransformation of new psychoactive substances, Sci. Rep. 9 (2019) 2741.
[24] S. Hemmer, S.K. Manier, S. Fischmann, F. Westphal, L. Wagmann, M.R. Meyer, Comparison of three untargeted data processing workflows for evaluating LC-HRMS metabolomics data, Metabolites 10 (2020).
[25] M. Malik, T. Demetrowitsch, K. Schwarz, T. Kunze, New perspectives on 'breathomics': metabolomic profiling of non-volatile organic compounds in exhaled breath using DI-FT-ICR-MS, Commun. Biol. 7 (2024) 258.
[26] T.C. Kuo, C.E. Tan, S.Y. Wang, O.A. Lin, B.H. Su, M.T. Hsu, J. Lin, Y.Y. Cheng, C. S. Chen, Y.C. Yang, K.H. Chen, S.W. Lin, C.C. Ho, C.H. Kuo, Y.J. Tseng, Human breathomics database, Database 2020 (2020).
[27] H. Zhao, Y. Yang, Y. Hao, W. Zhang, L. Cui, J. Wang, Y. Chen, T. Zuo, H. Yu, Y. Zhang, X. Song, Untargeted metabolomic analysis of exhaled breath condensate identifies disease-specific signatures in adults with asthma, Clin. Exp. Allergy (2025).
[28] E. Barberis, E. Amede, S. Khoso, L. Castello, P.P. Sainaghi, M. Bellan, P.E. Balbo, G. Patti, D. Brustia, M. Giordano, R. Rolla, A. Chiocchetti, G. Romani, M. Manfredi,

R. Vaschetto, Metabolomics diagnosis of COVID-19 from exhaled breath condensate, Metabolites 11 (2021).
[29] M. Bajo-Fernandez, E.A. Souza-Silva, C. Barbas, M.F. Rey-Stolle, A. Garcia, GC-MS-based metabolomics of volatile organic compounds in exhaled breath: applications in health and disease. A review, Front. Mol. Biosci. 10 (2023) 1295955.
[30] S. Hemmer, L. Wagmann, B. Pulver, F. Westphal, M.R. Meyer, In vitro and in vivo toxicometabolomics of the synthetic cathinone PCYP studied by means of LC-HRMS/MS, Metabolites 12 (2022).
[31] S.K. Manier, M.R. Meyer, Impact of the used solvent on the reconstitution efficiency of evaporated biosamples for untargeted metabolomics studies, Metabolomics 16 (2020) 34.
[32] H.H. Maurer, D.K. Wissenbach, A.A. Weber, Maurer/Wissenbach/Weber MWW LC-MSn Library of Drugs, Poisons, and their Metabolites, second ed., second ed., Wiley-VCH, Weinheim, Germany, 2019.
[33] H.H. Maurer, K. Pfleger, A.A. Weber, Mass Spectral Data of Drugs, Poisons, Pesticides, Pollutants and their Metabolites, Wiley-VCH, Weinheim, 2016.
[34] S.K. Manier, A. Keller, M.R. Meyer, Automated optimization of XCMS parameters for improved peak picking of liquid chromatography-mass spectrometry data using the coefficient of variation and parameter sweeping for untargeted metabolomics, Drug Test. Anal. 11 (2019) 752–761.
[35] R. Wehrens, J.A. Hageman, F. van Eeuwijk, R. Kooke, P.J. Flood, E. Wijnker, J. J. Keurentjes, A. Lommen, H.D. van Eekelen, R.D. Hall, R. Mumm, R.C. de Vos, Improved batch correction in untargeted MS-based metabolomics, Metabolomics 12 (2016) 88.
[36] R. Adusumilli, P. Mallick, Data conversion with ProteoWizard msConvert, Methods Mol. Biol. 1550 (2017) 339–368.
[37] C.A. Smith, E.J. Want, G. O'Maille, R. Abagyan, G. Siuzdak, XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification, Anal. Chem. 78 (2006) 779–787.
[38] C. Kuhl, R. Tautenhahn, C. Bottcher, T.R. Larson, S. Neumann, CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets, Anal. Chem. 84 (2012) 283–289.
[39] O. Yurekten, T. Payne, N. Tejera, F.X. Amaladoss, C. Martin, M. Williams, C. O'Donovan, MetaboLights: open data repository for metabolomics, Nucleic Acids Res. 52 (2024) D640–D646.
[40] S. Hemmer, L. Wagmann, M.R. Meyer, Altered metabolic pathways elucidated via untargeted in vivo toxicometabolomics in rat urine and plasma samples collected after controlled application of a human equivalent amphetamine dose, Arch. Toxicol. 95 (2021) 3223–3234.
[41] S.K. Manier, L. Wagmann, V. Flockerzi, M.R. Meyer, Toxicometabolomics of the new psychoactive substances alpha-PBP and alpha-PEP studied in HepaRG cell incubates by means of untargeted metabolomics revealed unexpected amino acid adducts, Arch. Toxicol. 94 (2020) 2047–2059.
[42] P.J. Linstrom, W.G. Mallard, The NIST chemistry WebBook: a chemical data resource on the internet, J. Chem. Eng. Data 46 (2001) 1059–1063.
[43] C. Guijas, J.R. Montenegro-Burke, X. Domingo-Almenara, A. Palermo, B. Warth, G. Hermann, G. Koellensperger, T. Huan, W. Uritboonthai, A.E. Aisporna, D. W. Wolan, M.E. Spilker, H.P. Benton, G. Siuzdak, METLIN: a technology platform for identifying knowns and unknowns, Anal. Chem. 90 (2018) 3156–3164.
[44] T. Kind, K.H. Liu, D.Y. Lee, B. DeFelice, J.K. Meissen, O. Fiehn, LipidBlast in silico tandem mass spectrometry database for lipid identification, Nat. Methods 10 (2013) 755–758.
[45] D.S. Wishart, A. Guo, E. Oler, F. Wang, A. Anjum, H. Peters, R. Dizon, Z. Sayeeda, S. Tian, B.L. Lee, M. Berjanskii, R. Mah, M. Yamamoto, J. Jovel, C. Torres-Calzada, M. Hiebert-Giesbrecht, V.W. Lui, D. Varshavi, D. Varshavi, D. Allen, D. Arndt, N. Khetarpal, A. Sivakumaran, K. Harford, S. Sanford, K. Yee, X. Cao, Z. Budinski, J. Liigand, L. Zhang, J. Zheng, R. Mandal, N. Karu, M. Dambrova, H.B. Schioth, R. Greiner, V. Gautam, Hmdb 5.0: the human metabolome database for 2022, Nucleic Acids Res. 50 (2022) D622–D631.
[46] M. Wang, J.J. Carver, V.V. Phelan, L.M. Sanchez, N. Garg, Y. Peng, D.D. Nguyen, J. Watrous, C.A. Kapono, T. Luzzatto-Knaan, C. Porto, A. Bouslimani, A.V. Melnik, M.J. Meehan, W.T. Liu, M. Crusemann, P.D. Boudreau, E. Esquenazi, M. Sandoval-Calderon, R.D. Kersten, L.A. Pace, R.A. Quinn, K.R. Duncan, C.C. Hsu, D.J. Floros, R.G. Gavilan, K. Kleigrewe, T. Northen, R.J. Dutton, D. Parrot, E.E. Carlson, B. Aigle, C.F. Michelsen, L. Jelsbak, C. Sohlenkamp, P. Pevzner, A. Edlund, J. McLean, J. Piel, B.T. Murphy, L. Gerwick, C.C. Liaw, Y.L. Yang, H.U. Humpf, M. Maansson, R.A. Keyzers, A.C. Sims, A.R. Johnson, A.M. Sidebottom, B.E. Sedio, A. Klitgaard, C.B. Larson, C.A.B. P, D. Torres-Mendoza, D.J. Gonzalez, D.B. Silva, L. M. Marques, D.P. Demarque, E. Pociute, E.C. O'Neill, E. Briand, E.J.N. Helfrich, E. A. Granatosky, E. Glukhov, F. Ryffel, H. Houson, H. Mohimani, J.J. Kharbush, Y. Zeng, J.A. Vorholt, K.L. Kurita, P. Charusanti, K.L. McPhail, K.F. Nielsen, L. Vuong, M. Elfeki, M.F. Traxler, N. Engene, N. Koyama, O.B. Vining, R. Baric, R. R. Silva, S.J. Mascuch, S. Tomasi, S. Jenkins, V. Macherla, T. Hoffman, V. Agarwal, P.G. Williams, J. Dai, R. Neupane, J. Gurr, A.M.C. Rodriguez, A. Lamsa, C. Zhang, K. Dorrestein, B.M. Duggan, J. Almaliti, P.M. Allard, P. Phapale, L.F. Nothias, T. Alexandrov, M. Litaudon, J.L. Wolfender, J.E. Kyle, T.O. Metz, T. Peryea, D. T. Nguyen, D. VanLeer, P. Shinn, A. Jadhav, R. Muller, K.M. Waters, W. Shi, X. Liu, L. Zhang, R. Knight, P.R. Jensen, B.O. Palsson, K. Pogliano, R.G. Linington, M. Gutierrez, N.P. Lopes, W.H. Gerwick, B.S. Moore, P.C. Dorrestein, N. Bandeira, Sharing and community curation of mass spectrometry data with global natural products social molecular networking, Nat. Biotechnol. 34 (2016) 828–837.
[47] K. Duhrkop, M. Fleischauer, M. Ludwig, A.A. Aksenov, A.V. Melnik, M. Meusel, P. C. Dorrestein, J. Rousu, S. Bocker, Sirius 4: a rapid tool for turning tandem mass spectra into metabolite structure information, Nat. Methods 16 (2019) 299–302.
[48] K. Duhrkop, H. Shen, M. Meusel, J. Rousu, S. Bocker, Searching molecular structure databases with tandem mass spectra using CSI:FingerID, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 12580–12585.
[49] E.M. Forsberg, T. Huan, D. Rinehart, H.P. Benton, B. Warth, B. Hilmers, G. Siuzdak, Data processing, multi-omic pathway mapping, and metabolite activity analysis using XCMS online, Nat. Protoc. 13 (2018) 633–651.
[50] S. Hemmer, S.K. Manier, L. Wagmann, M.R. Meyer, Impact of four different extraction methods and three different reconstitution solvents on the untargeted metabolomics analysis of human and rat urine samples, J. Chromatogr. A 1725 (2024).
[51] A.M. Evans, C. O'Donovan, M. Playdon, C. Beecher, R.D. Beger, J.A. Bowden, D. Broadhurst, C.B. Clish, S. Dasari, W.B. Dunn, J.L. Griffin, T. Hartung, P.C. Hsu, T. Huan, J. Jans, C.M. Jones, M. Kachman, A. Kleensang, M.R. Lewis, M.E. Monge, J.D. Mosley, E. Taylor, F. Tayyari, G. Theodoridis, F. Torta, B.K. Ubhi, D. Vuckovic, Q.C.C. Metabolomics Quality Assurance, Dissemination and analysis of the quality assurance (QA) and quality control (QC) practices of LC-MS based untargeted metabolomics practitioners, Metabolomics 16 (2020) 113.
[52] M. Costanzo, M. Caterino, G. Sotgiu, M. Ruoppolo, F. Franconi, I. Campesi, Sex differences in the human metabolome, Biol. Sex Differ. 13 (2022) 30.
[53] S. Barnes, H.P. Benton, K. Casazza, S.J. Cooper, X. Cui, X. Du, J. Engler, J. H. Kabarowski, S. Li, W. Pathmasiri, J.K. Prasain, M.B. Renfrow, H.K. Tiwari, Training in metabolomics research. II. Processing and statistical analysis of metabolomics data, metabolite identification, pathway analysis, applications of metabolomics and its future, J. Mass Spectrom. 51 (2016) 535–548.
[54] L.W. Sumner, A. Amberg, D. Barrett, M.H. Beale, R. Beger, C.A. Daykin, T.W. Fan, O. Fiehn, R. Goodacre, J.L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A.N. Lane, J.C. Lindon, P. Marriott, A.W. Nicholls, M.D. Reily, J. J. Thaden, M.R. Viant, Proposed minimum reporting standards for chemical analysis chemical analysis working group (CAWG) metabolomics standards initiative (MSI), Metabolomics 3 (2007) 211–221.
[55] W. Bernhard, Lung surfactant: function and composition in the context of development and respiratory physiology, Ann. Anat. 208 (2016) 146–150.