Journal of Pharmaceutical and Biomedical Analysis 262 (2025) 116874

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.journals.elsevier.com/journal-of-pharmaceutical-and-biomedical-analysis

Pseudotargeted metabolomics profiles potential damage-associated molecular patterns as machine learning predictors for acute pancreatitis

Wen-Lu Cai [a], Can Fang [a], Hong-Xu Leng [a], Jia-Yi Zheng [a], Li-Fang Liu [a], Guan-Wen Gong [b,*], Gui-Zhong Xin [a,**]

[a] Pukou Hospital of Chinese Medicine Affiliated to China Pharmaceutical University, Department of Chinese Medicines Analysis, China Pharmaceutical University, Nanjing, China

[b] Department of General Surgery, Affiliated Hospital of Nanjing, University of Chinese Medicine, Nanjing, China.

ARTICLE INFO

Keywords:
Metabolomics
Clinical decision support
Pancreatic disease
Machine learning
Damage-associated molecular patterns

ABSTRACT

Acute pancreatitis (AP) is a common gastrointestinal disease characterized by pancreatic cell damage and inflammation. Given the early clinical diagnosis and management challenges, exploring novel analytical frameworks from new orientations for interrogating AP is urgent. The release of damage-associated molecular patterns (DAMPs) and their receptor recognition initiate sterile inflammation, serving as key drivers in the development and progression of AP. Thus, this study aimed to delineate the underlying correlations between alterations in the DAMP profile and the AP state. We have developed a new framework combining potential DAMPs profiles obtained from pseudotargeted metabolomics method with machine learning (ML) models for AP prediction. 2-(1-Piperazinyl) pyrimidine chemical labeling was utilized to provide characteristic fragment ions and improve the quantitative sensitivity of targeted metabolites. A total of 49 potential DAMPs were identified and semi-quantified from collected serum samples (n = 84), positive or negative for APs. For modeling obtained datasets with five different ML algorithms, the support vector machine model was chosen as the optimal model to differentiate with high accuracy, achieving an area under the receiver-operating characteristic curve (AUROC) of 0.944. It also showed a strong performance in an external independent validation set (AUROC: 0.907). Moreover, the model was interpreted using the Shapley Additive exPlanations analysis to specify the important features and identify specific free fatty acids as key contributors. Overall, the novel framework enables high accuracy in predicting the presence of AP status. Meanwhile, it underlines the utility of DAMPs in inflammatory diseases and provides reference values for diagnosing in first-line clinics.

1. Introduction

Acute pancreatitis (AP) represents the most common gastrointestinal disease with an increasing incidence [1]. Most AP patients present with self-limiting acute inflammatory responses that resolve within a week, while about 20 % of patients develop severe AP (SAP) [2]. SAP is characterized by systemic inflammatory response syndrome (SIRS) associated with extra-pancreatic multiorgan failure (MOF) and mortality rates of up to 20–40 % [3,4]. Despite advances in research, such as clinical scoring systems for severity grading, predictors for diagnosis, and etiology definition, significant challenges remain for early and accurate prediction and management of AP in the clinical setting [5–7]. Therefore, exploring novel analytical frameworks from new orientations for interrogating AP is urgent.

The pathogenesis of AP involves acinar cell death, innate immune cell infiltration, and the release of inflammatory mediators. In the early stages, damage-associated molecular patterns (DAMPs) are released in abundance from damaged cells, triggering multiple pattern recognition receptors and leading to local tissue damage to the systemic inflammatory response (SIR) [8,9]. DAMPs are danger signals acted by endogenous molecules that undergo changes in their distribution, concentration, or physicochemical properties in a non-homeostatic state [10]. Of note, they are closely linked to causal or driving factors of the AP, which may reflect fundamental mechanisms and predictive factors.

* Corresponding author.
** Correspondence to: Pukou Hospital of Chinese Medicine Affiliated to China Pharmaceutical University, Department of Chinese Medicines Analysis, China Pharmaceutical University, No. 24 Tongjia Lane, Nanjing, China.
E-mail addresses: love_ggw2@163.com (G.-W. Gong), xingz@cpu.edu.cn (G.-Z. Xin).

https://doi.org/10.1016/j.jpba.2025.116874
Received 26 December 2024; Received in revised form 17 March 2025; Accepted 4 April 2025
Available online 7 April 2025

Currently, only a fraction of DAMPs, such as nucleic acids, proteins, ions, and glycans, have been extensively studied, while metabolite-derived DAMPs remain relatively rare; however, the metabolic changes observed in AP imply that they are significant and should not be neglected [11–13]. For instance, free fatty acids (FFAs) and bile acids (BAs) are considered potential target pools for DAMPs in AP. During AP episodes, the uncontrollable release of lipase from alveolar cells leads to excessive triglyceride hydrolysis in adipocytes, producing elevated FFAs that induce direct cellular toxicity, inflammatory responses, and mitochondrial function inhibition [14–16]. Furthermore, specific bile acids, including chenodeoxycholic acid, deoxycholic acid, and their taurine conjugates, have been identified as vital DAMPs capable of activating signals of the NOD -, LRR -, and pyrin domain-containing protein 3 (NLRP3) inflammasome [17,18]. Therefore, identifying and quantifying a specific subset of potential metabolite-derived DAMPs that reveal the underlying mechanisms and metabolic networks would be highly useful for AP diagnosis.

Metabolomics combined with machine learning (ML) methods has been applied to diagnosis and prediction in many fields of biomedical science [19–22]. Of note, the chemical labeling-based pseudotargeted metabolomics method offers a robust strategy for the comprehensive and sensitive profiling of metabolites [23,24]. The pseudotargeted metabolomics method combines the superiorities of the untargeted and targeted methods, providing improved data quality and a wider linear range compared to untargeted methods while also offering a better holistic overview of the metabolome independent of the need for authentic standards [25]. Additionally, chemical labeling helps overcome drawbacks such as low ionization efficiency and lack of characteristic MRM transitions, further facilitating the accuracy and sensitivity of the analysis [26,27]. Eventually, it provides high-quality metabolomics data for ML algorithms to interpret data, construct prediction models, and identify key contributors. That stated, the combination of ML and metabolomics presents remarkable advantages.

Here, we aim to create a novel framework combining potential DAMPs profiles obtained from the pseudotargeted metabolomics method and ML models for AP prediction. This integrated framework provides an accessible way to make clinical decision-support tools with remarkable performance properties. It is also expected to present some clinical implications for targeted interventions for AP and other inflammatory diseases.

2. Material and methods

2.1. Subject information

All the subjects were recruited from different campuses of the Affiliated Hospital of Nanjing University of Chinese Medicine, including the Main Campus (MC) and the Zidong Campus (ZDC). The diagnosis of AP was made according to the revised Atlanta classification and definitions [28]. The exclusion criteria were as follows: age < 18 or > 75 years, cancer, cirrhosis, pregnancy, primary sclerosing cholangitis, renal transplantation, and active viral hepatitis. Finally, serum from 54 AP patients and 30 healthy individuals from the MC were included as the discovery set, and data from 27 AP patients and 10 healthy individuals from ZDC were included as the external validation set. Written informed consent was obtained from all participants. The research was approved by the Medical Ethics Committee of Affiliated Hospital of Nanjing University of Chinese Medicine (Approval number: 2024NL-286–01) and complied with the World Medical Association Declaration of Helsinki. All samples were taken through a clinical standard procedure.

2.2. Chemicals and solvents

Standards compounds of FFAs and BAs (listed in Supplemental Table 1) and 2-(1-Piperazinyl) pyrimidine (PP) 95 % were all from Macklin (Shanghai, China). 1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide (EDC), 1-Hydroxy-7-azabenzotriazole (HOAT), and two internal standards (IS) (cholic acid-d4 and palmitic acid-^{13}C) were from Aladdin (Shanghai, China). LC-MS grade solvents and formic acid were from Merck (Darmstadt, Germany). Optimal-grade water was produced by a Milli-Q purification system (Millipore, Watford, UK).

2.3. Sample preparation and chemical labeling method

Serum samples collected from all participants were obtained using the PP-based derivatization method to facilitate subsequent metabolomic analysis. In brief, aliquots of 50 μL serum were added to 200 μL precooled acetonitrile and 1 μL of IS solvent, then vortexed to mix, and the protein was precipitated at −20°C for 1 h. The resulting mixture was centrifuged (13000 rpm, 4 °C, 15 min), and 200 μL of the supernatant was transferred for further derivatization. The derivatization method was as follows: 20 μL of PP solution (10 mg/ML) was added to the above supernatant and vortex mixed for 1 min. Then 10 μL of EDC solution (30 mg/ML) and 10 μL of HOAT solution (30 mg/ML) were added sequentially and mixed well, followed by 1 h of incubation at 25 ℃. Each PP-labeled sample was evaporated to dryness and redissolved with 50 % (v/v) acetonitrile/water equally. Finally, the resulting mixture was centrifuged (13000 rpm, 4 °C, 10 min) for further LC-MS/MS analysis. Pooled quality control (QC) samples were prepared by mixing an equal volume (10 μL) of each sample.

2.4. Pseudotargeted metabolomic profiling of potential DAMPs

Metabolomic profiling of potential DAMPs was conducted utilizing a pseudotargeted workflow combining untargeted and targeted LC-MS/MS analysis. The global characterization of the derived products was performed on an ExionLC TM AD UPLC system coupled with an X500B quadrupole time-of-flight (QTOF) system (AB SCIEX, USA) with electrospray ionization (ESI) in information-dependent acquisition (IDA) mode. A Waters Acquity BEH C18 column (2.1 × 150 mm, 1.8 μm) was applied. The mobile phase consisted of an aqueous solution with 0.1 % formic acid (v/v) (A) and acetonitrile (B). The gradient was performed with aqua/acetonitrile (95:5 v/v at 0 min, 90:10 v/v at 2.0 min, 25:75 v/v at 6.0 min, 3:97 v/v at 18.0 min, 3:97 v/v at 21.0 min, 95:5 v/v at 21.1 min, and 95:5 v/v at 25.0 min). The MS parameters settings were as follows: ion spray voltage, 5500 V (ESI+); ion source temperature, 550 °C; ion source gas 1, 45 psi; gas 2, 45 psi; curtain gas, 35 psi; collision-activated dissociation gas, 7; declustering potential, 80 V; collision energy (CE), 10 eV (MS mode) or 35 eV (MS2 mode); collision energy spread, 15 eV.

Pseudotargeted measurements of derived potential DAMPs were performed in scheduled high-resolution multiple reaction monitoring (MRM^{HR}) mode. For optimization, the CE voltage was set at 15, 25, and 35 V in positive ion mode for all candidate ion pairs to optimize; MRM^{HR} transitions were multiplexed according to the candidate elution time, and retention time (Rt) tolerance was set as 90 s. Other operating parameters were the same as those in the IDA settings. The peak area of metabolites detected was normalized by the respective IS (Cholic acid-d4 for BAs and Palmitic acid-^{13}C for FFAs). Data was carried out using SCIEX OS acquisition software.

2.5. Method evaluation

Validation of the analytical characteristics of the established pseudotargeted method, including linearity, repeatability, and stability. To test linearity, calibration was conducted based on the analysis of a derivatized 2^{10}-fold dilution series of QC solutions (1, 1/2, 1/4, 1/8, 1/16, 1/32, 1/64, 1/128, 1/256, 1/512, and 1/1024). Calibration curves were established by employing a linear regression with a weighting factor of $1/x^2$. The repeatability was assessed by analyzing three levels of QC samples, including low, medium, and high concentrations. The intraday and interday repeatability were estimated using the coefficient

of variation (CV, %) of each level QC sample with six replicates. Autosampler storage stability was evaluated by monitoring the QC samples in the analytical sequence over 24 h.

2.6. Model development and comparison

Five ML model classifiers were considered and achieved in R 4.4.0, namely, SVM (support vector machine), RF (random forest), GBM (gradient boosting machine), KNN (k-nearest neighbor), and LASSO (lasso regression machine), to predict AP in patients. The features with more than 50 % missing data were excluded from the analysis. Subsequently, the discovery dataset was then randomly divided into a training dataset (70 %) and an internal test dataset (30 %). To optimize the prediction model, a grid search combined with manual fine-tuning was applied to obtain the final hyperparameters. To enhance the model's generalization and mitigate overfitting, repeated k-fold cross-validation was adopted for hyperparameter tuning and model validation. A 10-fold repeat cross-validation was used here. The external validation dataset was used to validate the optimal model classifier.

Several commonly used evaluation metrics, such as the area under the receiver-operating characteristic curve (AUROC), sensitivity, specificity, accuracy, and F1 score, were used to evaluate the reliability of these models.

2.7. Model interpretation

The Shapley Additive exPlanations (SHAP) is a game-theoretic approach to unlock the ML results. SHAP allows us to estimate the importance of each feature by applying a test sets-based game theory approach to measure the impact of genus characteristics on the prediction scores [29]. The SHAP value quantifies the impact of the variable on the prediction in terms of both direction and magnitude. It is computed

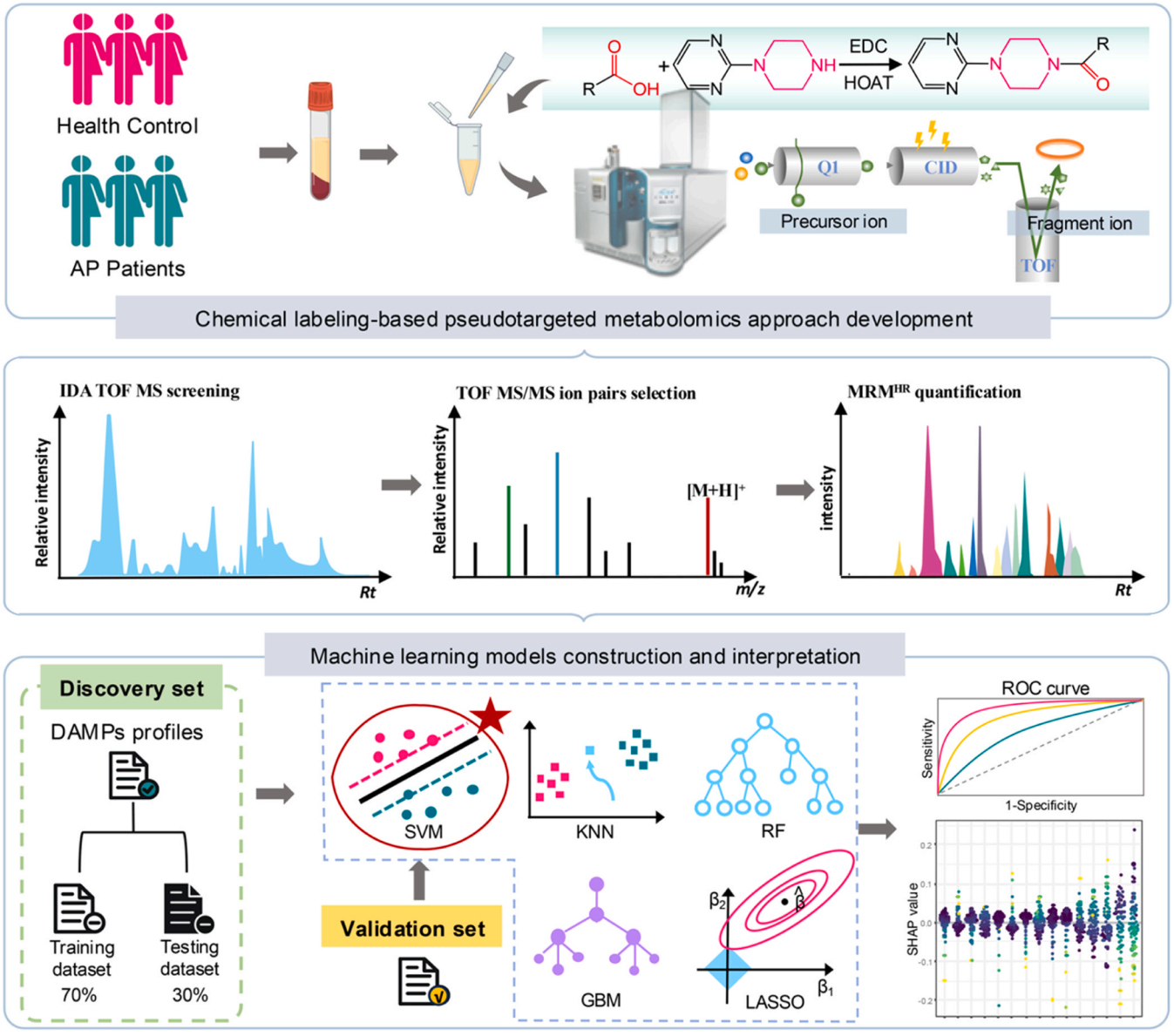

Fig. 1. The overall workflow of the study design. Serum samples from individuals underwent chemical labeling-based pseudotargeted metabolomics analysis. Using the metabolomics data and ML algorithms, a prediction model for AP (acute pancreatitis) was created and interpreted by using a discovery set and validated with an external validation set. IDA, data-independent acquisition; TOF, time of flight; MS, mass spectrometry; MRMHR, high-resolution multiple reaction monitoring; DAMPs, damage-associated molecular pattern molecules; KNN, k-nearest neighbor; RF, random forest; GBM, gradient boosting machine; SVM, support vector machine; LASSO, lasso regression machine; AUROC, the area under the receiver-operating characteristic curve; SHAP, Shapley Additive exPlanations.

by considering the prediction results for each possible combination of features. This comprehensive analysis provides valuable information about the contribution of each genus characteristic to the overall predictions.

3. Results and discussion

3.1. Study design and sample clinical characteristics

The overall workflow of this study was illustrated in Fig. 1. Among the 84 individuals from MC, the average age was 50.0 ± 13.55 and 56.6 ± 15.2 in each group. No significant differences were present between the groups based on the white blood cell counts, C-reactive protein level, or serum amylase level (P = 0.244, 0.449, and 0.057, respectively; Mann-Whitney *U* test). The serum lipase level in the AP group was significantly different from that in the healthy control group (P = 0.027; Mann-Whitney *U* test). The clinical characteristics of all participants were summarized in Supplemental Table 2. Next, the metabolite-derived DAMPs profile (FFAs and BAs) of serum samples was obtained using a chemical derivatization-based pseudotargeted metabolomics approach. Then, the optimal SVM model, out of five ML algorithms, was constructed to investigate the association between the DAMPs panels and clinical phenotypes, distinguishing AP patients from healthy controls, which was validated in an external validation set. In addition, SHAP analysis was utilized to specify the importance of each model feature and its basis for decision-making.

In this study, we employed the chemical labeling-based pseudotargeted metabolomics method to delineate the metabolite-derived DAMPs profile in AP patients and to identify specific DAMPs with diagnostic potential. Collectively, our study underscored the distinctive merits of applying ML-based metabolomics in the early diagnosis of AP, providing a framework for future clinical guidance.

3.2. Method development

3.2.1. Overview of a chemical labeling-based Pseudotargeted LC-MS/MS Method

AP is associated with a rapid and uncontrolled inflammatory response, potentially progressing to SIRS and MOF. The primary means of clinical management of AP are adequate fluid resuscitation and the prevention of MOF; no treatment can alter the course of the AP. Despite the urgency, the diagnosis and therapeutic intervention of AP continue to pose major challenges due to the complex and heterogeneous clinical and pathological landscapes. Increasing evidence suggests that DAMPs are initial causal or driving factors in mediating sterile inflammation [30]. Typically, in AP, metabolite-derived DAMPs are supposed to show changes in distribution, concentration, or physicochemical properties in response to metabolic aberrations under non-steady-state conditions. Considering the major etiologic factors (biliary and hyperlipidemia) of AP and valuable insights from prior studies, FFAs and BAs were selected as target pools to enhance diagnostic precision and inform clinical management strategies in this study. As known, FFAs and BAs are both common carboxyl compounds, but their analysis remains a challenge due to the following issues: (1) lack of characteristic fragment ions; (2) low ionization efficiency; (3) difficulty in obtaining sufficient standards due to the presence of multiple isomers; (4) wide concentration range in samples. The workflow scheme of a chemical labeling-based pseudotargeted LC-MS/MS method was established to confront the above, depicted in Supplemental Fig. S1. It mainly contains four key steps:

Step 1: Sample chemical labeling. Herein, PP was chosen as the suitable labeling probe based on the subunits as follows: (1) the piperazine moiety has a great reactivity toward the carboxyl group [31]; (2) the piperazine moiety with high proton affinity, which facilitates the ionization process, can increase the sensitivity; (3) the pyrimidine ring facilitates the ionization process, increasing the MS signal [26]. EDC and HOAT were used as condensation agents to cross-link the carboxyl group with the amine group. The scheme of the derivatization mechanism was shown in Fig. 2A.

Step 2: Metabolite Data Acquisition by UPLC-QTOF MS/MS with IDA full scan mode. Subsequent analysis of the PP-labeled sample was performed on UPLC-QTOF MS/MS with IDA full scan mode. (1) The raw data were analyzed in SCIEX OS 2.2.0, where the MS and MS/MS spectra were checked, and potential carboxyl metabolites were annotated to determine whether the MS/MS spectra of the derivatives contained characteristic fragment ions (*m/z* 122.0716 and 165.1135). (2) A house-built database of theoretically possible FFAs and BAs and their derivatizations was established, listing *m/z* and molecular formula information. Further confirmation was performed by comparing the *m/z* of the potential FFAs and BAs with the database ($\triangle m/z < 10$ ppm). Finally, FFAs with total carbon atoms varying from 7 to 22 and up to six double bonds (C = C) were included. (3) As reported, certain correlation patterns exist between derived FFA structures and chromatographic retention, which have been applied to facilitate metabolite identification and quantification without reference standards [27,32]. Here, the retention behavior of PP-labeled FFAs and the classical carbon number rule were explored using a mixture of six saturated FFA standards and applied to the RT calibration of screened metabolites (Supplemental Fig. S2).

Taking stearic acid as an example, the MS/MS spectra of PP-labeled stearic acid in positive ion mode were shown in Fig. 2B. The characteristic fragment ions of PP-labeled stearic acid (*m/z* 431.3741) were at *m/z* 122.0716 and 165.1135, which were both introduced by derivatization reagents. Upon comparison with the in-house database, it can be deduced that *m/z* 431.3741 was the PP-labeled stearic acid derivative. It was worth noting that the method was still not able to determine the position of the C = C bond. Therefore, the exact structure of the unsaturated fatty acid could not be confirmed. In total, 40 FFAs and 9 BAs were putatively identified based on the screening strategy described above (Supplemental Table 3).

Step 3: *Conduct pseudotargeted analysis based on UPLC-QTOF MS/MS with MRM^{HR} mode.* The selection of ion pairs and the optimization of the relevant parameters are necessary preliminary steps. A unique fragment ion can be applied as a quantitative ion with improved sensitivity to support the identification of potential fatty acids in biological samples. Herein, the precursor ions were chosen as $[M+H]^+$, and the fragment with the highest intensity in the MS/MS spectrum was chosen as a characteristic fragment ion of the precursor ion (characteristic ions introduced by derivatization, *m/z* 122.0716 and 165.1135). Then, optimal characteristic ion pairs were selected from different CEs (15, 25, and 35 V) (Fig. 2C). Based on the optimal ion-pair list in Supplemental Table 4, the pseudotargeted method can eventually be established and validated. Representative extracted ion chromatograms acquired from the QC sample after derivatization are shown in Fig. 3.

3.2.2. Method validation

Linearity, repeatability, and stability were validated for the established pseudotargeted metabolomics approach. The linearities of calibration curves for all analytes were satisfactory with correlation coefficients (R^2) greater than 0.990 among the validated concentration range of dilution series of QC solutions. The stability was evaluated by calculating the change in QC samples during the analytical sequence; the results showed that CVs (n = 8) were all < 30 % (Supplemental Table 5). The intra- and inter-day repeatability was evaluated by analyzing three-level QC samples. The results showed that CVs were all < 15 % in all QC samples, but due to the wide range of concentrations, several analytes were not well detected in low-level QC samples, which were within acceptable limits (Supplemental Table 6).

3.3. Exploratory analysis of the dataset

To characterize the serum DAMPs profiles of AP, metabolomic analysis was performed comparing AP patients with healthy controls in the discovery set. As shown in the principal component analysis (PCA)

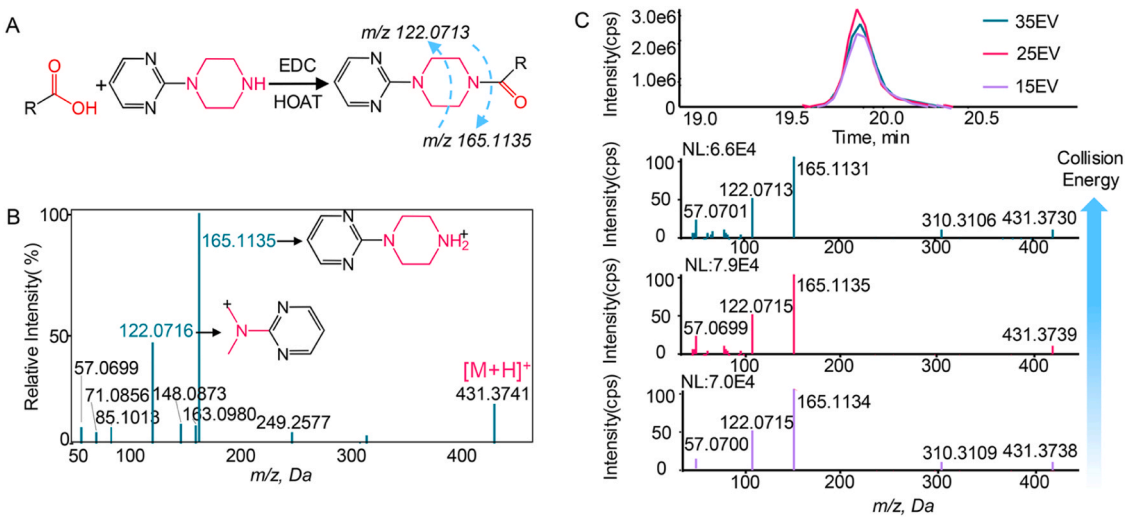

Fig. 2. Development of chemical labeling-based pseudotargeted LC-MS/MS method. (A) Illustration of the reaction of PP (2-(1-Piperazinyl) pyrimidine) derivatization and potential DAMPs containing carboxyl groups. (B) Fragmentation patterns of PP-labeled stearic acid. (C) Collision Energy optimization of PP-labeled stearic acid ion pairs.

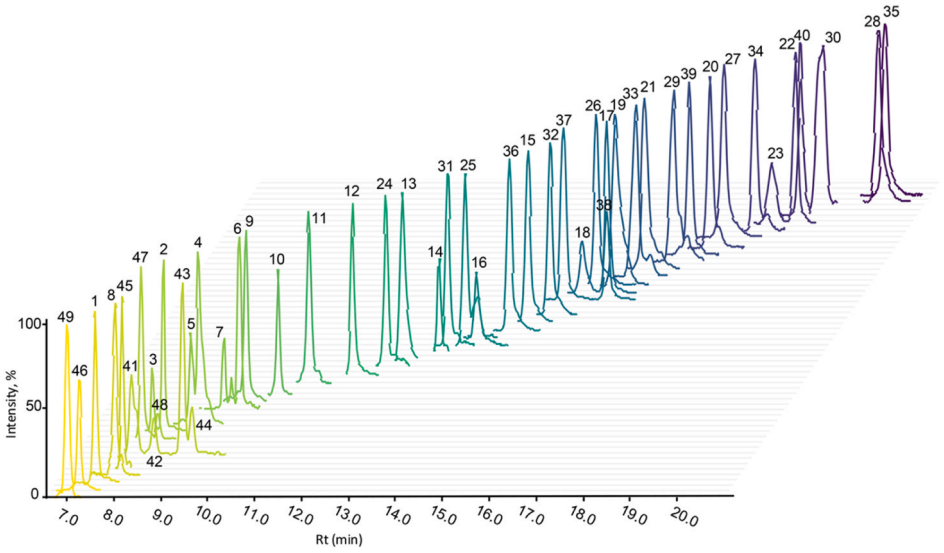

Fig. 3. Representative extracted ion chromatograms of PP-derivatized potential DAMPs.

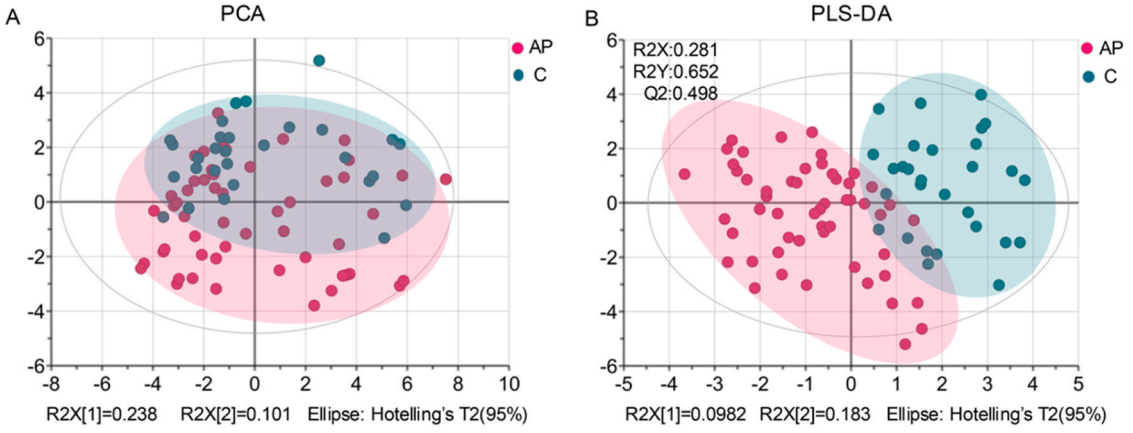

Fig. 4. Exploratory analysis of the dataset. (A) PCA (principal component analysis) score plot of the serum pseudotargeted metabolomics data comparing the acute pancreatitis group (AP, colored in pink) and healthy controls (C, colored in green). (B) score plot of the PLS-DA (partial-least squares discrimination analysis) model. The evaluation criteria (R2X, R2Y, Q2) are displayed.

scores plot (Fig. 4A), the DAMPs profiling of AP could not be completely distinguished from that of healthy controls, and the first two principal components only explained 33.9 % of the variance in the data, implicating that the relationships between features were complex. Furthermore, the partial-least squares discrimination analysis (PLS-DA) model was fitted. The score plots and cross-validated quality metrics for model predictive power (Q2) and explained variance (R2X and R2Y) are shown in Fig. 4B. Taken together, there were certain differences in the DAMPs profiles between the two groups, but the predictive performance of PCA and PLS-DA was limited.

3.4. Machine learning for prediction

Next, we leveraged the acquired DAMPs profiles to develop high-performance diagnostic approaches. ML models have been utilized to reveal potential complex correlations between multiple features within the metabolomics method and disease states and then create robust predictive models [33]. Herein, five ML algorithms, including SVM, RF, GBM, KNN, and LASSO, were used to construct models for predicting the clinical states. As shown in Fig. 5A, the accuracy, sensitivity, specificity, and recall of each model were computed and compared in the discovery set. While all models performed well, the SVM model achieved the best performance in multiple evaluation metrics: accuracy 0.92, precision 1.00, recall 0.88, and F1 score 0.93 (Supplemental Table 7). Additionally, the SVM model exhibited superior predictive performance with an AUROC of 0.944, indicating a high level of accuracy in prediction (Fig. 5B). In order to verify the reliability of the model, an external set, including 27 AP patients and 10 healthy individuals, was used to validate the classification model produced from the SVM analysis. The SVM model achieved an AUROC of 0.907, with an accuracy of 0.87 and a precision of 0.84 (Supplemental Fig. S3). Together, the SVM was ultimately selected as the optimal model.

The application of ML algorithms in our study facilitated the intricate relationships between the DAMPs features and disease status, surpassing the traditional chemometric models such as PCA and PLS-DA, which were found to be less robust in our dataset. This highlighted the complexity of the metabolic DAMPs profile of AP and validated our approach of utilizing ML for comprehensive data interpretation and differentiation between AP patients and healthy controls.

3.5. SHAP-based model interpretation

SHAP analysis was utilized to identify the potential DAMPs that contributed the most to the prediction of AP using the pseudotargeted metabolomics data set for serum. The SHAP value, derived from coalitional game theory, depicts the average impact of each feature. Fig. 5C illustrates the SHAP values and the corresponding distribution of data points for the 15 most influential features. Each point represents a sample, and the color scale from purple to yellow indicates the magnitude of the sample feature values. The effects of the top 15 features in terms of importance on the predicted results are shown in Fig. 5D. Specifically, stearic acid (C18:0), heptadecanoic acid (C17:0), and eicosadienoic acid (C20:2), which belong to FFAs, exhibit key contributions to the predictive results. FFAs have been studied to act as

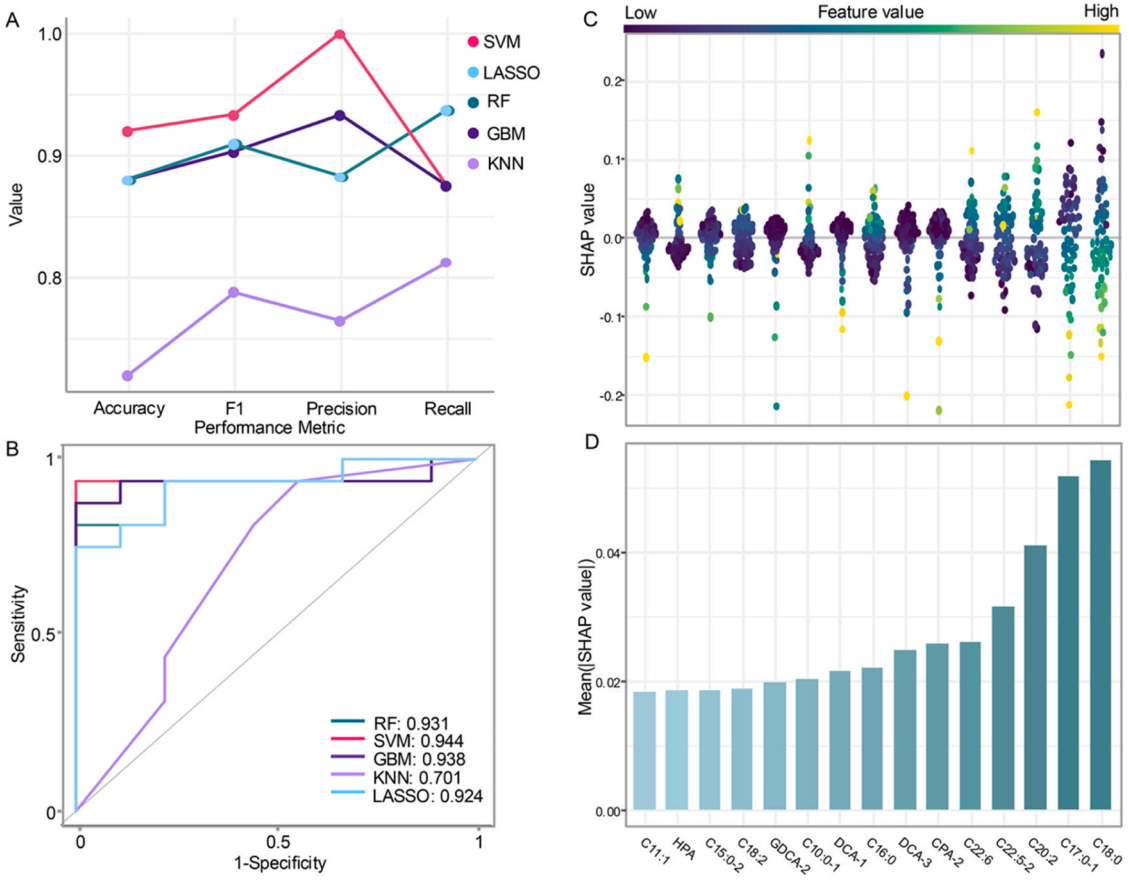

Fig. 5. Evaluation of machine learning model performance and interpretation of feature impact. (A) Line graph of accuracy, precision, sensitivity, and F1 score. (B) ROC (receiver-operating characteristic curve) plot for SVM (support vector machine), RF (random forest), GBM (gradient boosting machine), KNN (k-nearest neighbor), and LASSO (lasso regression machine). (C) Beeswarm plot of SHAP values distribution across all samples. The color assigned to each spectrum (purple for low eigenvalues and yellow for high eigenvalues) corresponds to the eigenvalues, depicting the intensity of features in that spectrum. (D) The column plot depicts the mean Shapley value and each feature's average impact on the model's output.

signalling molecules that augment the inflammatory response and subsequent lipotoxicity in AP. Moreover, FFAs can inhibit mitochondrial complexes in pancreatic acinar cells and induce the pathological elevation of intracellular Ca^{2+} concentration, cytokine release, and tissue injury. The pro- or anti-inflammatory effects of individual FFAs may vary with disease context, concentration, and overall metabolic profile [15]. Moreover, our study identified stearic acid (C18:0) as a significant predictor, aligning with literature suggesting its pro-inflammatory properties [34,35].

The limitations inherent in a cross-sectional study design preclude insights into possible longitudinal changes in AP. Future longitudinal studies are essential for a more nuanced understanding of metabolome differences across disease stages. Furthermore, it is imperative to expand the pool of participants to explore the metabolomics of specific AP clinical subtypes and their correlation with disease severity and timing of intervention more deeply. Future studies should also emphasize extending the identification of metabolite-derived DAMPs contributing to AP, potentially enhancing diagnostic accuracy and clinical relevance.

4. Conclusion

In summary, our study applied a chemical-labeling-based pseudotargeted metabolomics method to delineate metabolite-derived DAMPs profiles in AP. Subsequently, we utilized DAMPs profiles as ML predictors to construct predictive models capable of distinguishing between AP patients and healthy controls. Experimental results indicated that the SVM model demonstrated a robust classification performance in this study. This research not only provides an approach to improve clinical diagnosis from the perspective of the DAMPs profile but also enlightens the great possibilities of DAMPs as predictors of severity and complication management, recurrence prediction, and timing of surgery. Furthermore, the framework highlights the distinct advantages of ML-based pseudotargeted metabolomics approaches for decision guidance and can be extrapolated for exploring other pathological states.

CRediT authorship contribution statement

Gong Guan-Wen: Writing – review & editing, Supervision, Resources, Funding acquisition. **Xin Gui-Zhong:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Zheng Jia-Yi:** Writing – review & editing, Conceptualization. **Liu Li-Fang:** Writing – review & editing, Supervision, Project administration. **Fang Can:** Software, Methodology, Investigation. **Leng Hong-Xu:** Methodology, Investigation. **Cai Wen-Lu:** Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Conceptualization.

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

This work was supported by the National Natural Science Foundation of China [No.82274064], Young and Middle-aged Academic Leader of Cyan and Blue Project in Jiangsu universities [Gui-Zhong Xin], and the Special research project of Jiangsu Provincial Administration of Traditional Chinese Medicine [No. ZT202108].

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jpba.2025.116874.

Data availability

Data will be made available on request.

References

[1] P.J. Lee, G.I. Papachristou, New insights into acute pancreatitis, Nat. Rev. Gastroenterol. Hepatol. 16 (2019) 479–496.
[2] W. Zhou, S. Dong, Z. Chen, X. Li, W. Jiang, New challenges for microRNAs in acute pancreatitis: progress and treatment, J. Transl. Med. 20 (2022) 192.
[3] L. Boxhoorn, R.P. Voermans, S.A. Bouwense, M.J. Bruno, R.C. Verdonk, M. A. Boermeester, H.C. Van Santvoort, M.G. Besselink, Acute pancreatitis, Lancet 396 (2020) 726–734.
[4] N.J. Schepers, O.J. Bakker, M.G. Besselink, U.A. Ali, T.L. Bollen, H.G. Gooszen, H. C. van Santvoort, M.J. Bruno, Impact of characteristics of organ failure and infected necrosis on mortality in necrotising pancreatitis, Gut 68 (2019) 1044–1051.
[5] D. Lou, K. Shi, H.-P. Li, Q. Zhu, L. Hu, J. Luo, R. Yang, F. Liu, Quantitative metabolic analysis of plasma extracellular vesicles for the diagnosis of severe acute pancreatitis, J. Nanobiotechnology 20 (2022) 52.
[6] A. Türkoğlu, A. Böyük, M.H. Tanrıverdi, E. Gündüz, A. Dusak, İ. Kaplan, M. Gümüş, The potential role of BMI, plasma leptin, nesfatin-1 and ghrelin levels in the early detection of pancreatic necrosis and severe acute pancreatitis: A prospective cohort study, Int. J. Surg. 12 (2014) 1310–1313.
[7] D.X. Jin, R. Lacson, M. Eskian, J. McNabb-Baltar, P.A. Banks, S.R. Kayden, A. S. Raja, R. Khorasani, Prospective validation of a prediction model for the diagnosis of acute pancreatitis, JAMA Netw. Open 7 (2024) e2419014.
[8] M. Ma, W. Jiang, R. Zhou, DAMPs and DAMP-sensing receptors in inflammation and diseases, Immunity 57 (2024) 752–771.
[9] T. Gong, L. Liu, W. Jiang, R. Zhou, DAMP-sensing receptors in sterile inflammation and inflammatory diseases, Nat. Rev. Immunol. 20 (2020) 95–112.
[10] N. Kang, Z. Ji, Y. Li, J. Gao, X. Wu, X. Zhang, Q. Duan, C. Zhu, Y. Xu, L. Wen, X. Shi, W. Liu, Metabolite-derived damage-associated molecular patterns in immunological diseases, FEBS J. 291 (2024) 2051–2067.
[11] M. Inoue, T. Nakahama, R. Yamasaki, T. Shibuya, J.I. Kim, H. Todo, Y. Xing, Y. Kato, E. Morii, Y. Kawahara, An aicardi-goutières syndrome–causative point mutation in adar1 gene invokes multiorgan inflammation and late-onset encephalopathy in mice, J. Immunol. 207 (2021) 3016–3027.
[12] A.P. West, G.S. Shadel, Mitochondrial DNA in innate immune responses and inflammatory pathology, Nat. Rev. Immunol. 17 (2017) 363–375.
[13] W. Liu, Y. Ren, T. Wang, M. Wang, Y. Xu, J. Zhang, J. Bi, Z. Wu, Y. Zhang, R. Wu, Blocking CIRP protects against acute pancreatitis by improving mitochondrial function and suppressing pyroptosis in acinar cells, Cell Death Discov. 10 (2024) 1–14.
[14] X. Gu, Z. Huang, X. Ying, X. Liu, K. Ruan, S. Hua, X. Zhang, H. Jin, Q. Liu, J. Yang, Ferroptosis exacerbates hyperlipidemic acute pancreatitis by enhancing lipid peroxidation and modulating the immune microenvironment, Cell Death Discov. 10 (2024) 1–10.
[15] A.E. Phillips, A.S. Wilson, P.J. Greer, A. Hinton, S. Culp, P. Paragomi, I. Pothoulakis, V. Singh, P.J. Lee, I. Lahooti, D.C. Whitcomb, G.I. Papachristou, Relationship of circulating levels of long-chain fatty acids to persistent organ failure in acute pancreatitis, Am. J. Physiol. Gastrointest. Liver Physiol. 325 (2023) G279–G285.
[16] Y. Zhang, W. He, C. He, J. Wan, X. Lin, X. Zheng, L. Li, X. Li, X. Yang, B. Yu, X. Xian, Y. Zhu, Y. Wang, G. Liu, N. Lu, Large triglyceride-rich lipoproteins in hypertriglyceridemia are associated with the severity of acute pancreatitis in experimental mice, Cell Death Dis. 10 (2019) 1–17.
[17] H. Hao, L. Cao, C. Jiang, Y. Che, S. Zhang, S. Takahashi, G. Wang, F.J. Gonzalez, Farnesoid X receptor regulation of the NLRP3 inflammasome underlies cholestasis-associated sepsisNLRP3, Cell Metab. 25 (2017) 856–867, e5.
[18] Q. Zhu, C. Yuan, X. Dong, Y. Wang, B. Li, B. Tu, W. Chen, X. Xu, W. Gong, W. Xiao, Y. Ding, L. Hu, W. Li, G. Lu, Bile acid metabolomics identifies chenodeoxycholic acid as a therapeutic agent for pancreatic necrosis, Cell Rep. Med. 4 (2023) 101304.
[19] Y. Deng, Y. Yao, Y. Wang, T. Yu, W. Cai, D. Zhou, F. Yin, W. Liu, Y. Liu, C. Xie, J. Guan, Y. Hu, P. Huang, W. Li, An end-to-end deep learning method for mass spectrometry data analysis to reveal disease-specific metabolic profiles, Nat. Commun. 15 (2024) 7136.
[20] Y. Chen, B. Wang, Y. Zhao, X. Shao, M. Wang, F. Ma, L. Yang, M. Nie, P. Jin, K. Yao, H. Song, S. Lou, H. Wang, T. Yang, Y. Tian, P. Han, Z. Hu, Metabolomic machine learning predictor for diagnosis and prognosis of gastric cancer, Nat. Commun. 15 (2024) 1–13.
[21] E. Alladio, F. Trapani, L. Castellino, M. Massano, D. Di Corcia, A. Salomone, E. Berrino, R. Ponzone, C. Marchiò, A. Sapino, M. Vincenti, Enhancing breast cancer screening with urinary biomarkers and random forest supervised classification: a comprehensive investigation, J. Pharm. Biomed. Anal. 244 (2024) 116113.
[22] J. Rong, G. Sun, J. Zhu, Y. Zhu, Z. Chen, Combination of plasma-based lipidomics and machine learning provides a useful diagnostic tool for ovarian cancer, J. Pharm. Biomed. Anal. 253 (2025) 116559.
[23] X. Xiong, X. Gu, X. Li, K. Jia, Y. Wei, R. Zhao, A chemical derivatization-based pseudotargeted liquid chromatography-tandem mass spectrometry method for sensitive and high coverage determination of bile acids in human serum, Anal. Chim. Acta 1287 (2024) 342119.

[24] E. Ivanovová, B. Pisklákova, D. Dobešová, A. Kvasnička, D. Friedecký, Novel LC-MS tools for diagnosing inborn errors of metabolism, Microchem. J. 170 (2021) 106654.
[25] F. Zheng, X. Zhao, Z. Zeng, L. Wang, W. Lv, Q. Wang, G. Xu, Development of a plasma pseudotargeted metabolomics method based on ultra-high-performance liquid chromatography–mass spectrometry, Nat. Protoc. 15 (2020) 2519–2537.
[26] R. Jiang, Y. Jiao, P. Zhang, Y. Liu, X. Wang, Y. Huang, Z. Zhang, F. Xu, Twin derivatization strategy for high-coverage quantification of free fatty acids by liquid chromatography–tandem mass spectrometry, Anal. Chem. 89 (2017) 12223–12230.
[27] T. Tang, P. Zhang, S. Li, D. Xu, W. Li, Y. Tian, Y. Jiao, Z. Zhang, F. Xu, Absolute quantification of acylcarnitines using integrated Tmt-PP derivatization-based LC–MS/MS and quantitative analysis of multi-components by a single marker strategy, Anal. Chem. 93 (2021) 12973–12980.
[28] P.A. Banks, T.L. Bollen, C. Dervenis, H.G. Gooszen, C.D. Johnson, M.G. Sarr, G. G. Tsiotos, S.S. Vege, Classification of acute pancreatitis—2012: revision of the Atlanta classification and definitions by international consensus, Gut 62 (2013) 102–111.
[29] S. Lundberg, S.I. Lee, A unified approach to interpreting model predictions, Adv. Neural Inf. Process. Syst. (2017) 4766–4775.
[30] Y. Huang, W. Jiang, R. Zhou, DAMP sensing and sterile inflammation: intracellular, intercellular and inter-organ pathways, Nat. Rev. Immunol. 24 (2024) 703–719.
[31] Y. Xu, L. Zhang, H. Lu, P. Yang, Mass spectrometry analysis of phosphopeptides after peptide carboxy group derivatization, Anal. Chem. 80 (2008) 8324–8328.
[32] Q. Zhu, T. Zhang, L. Qin, X. Li, S. Zheng, Y. Feng, Method to calculate the retention index in hydrophilic interaction liquid chromatography using normal fatty acid derivatives as calibrants, Anal. Chem. 91 (2019) 6057–6063.
[33] J.D. Zhang, C. Xue, V.B. Kolachalama, W.A. Donald, Interpretable machine learning on metabolomics data reveals biomarkers for parkinson's disease, ACS Cent. Sci. 9 (2023) 1035–1045.
[34] K.A. Harvey, C.L. Walker, T.M. Pavlina, Z. Xu, G.P. Zaloga, R.A. Siddiqui, Long-chain saturated fatty acids induce pro-inflammatory responses and impact endothelial cell growth, Clin. Nutr. 29 (2010) 492–500.
[35] J. Zeng, Y. Zhang, J. Hao, Y. Sun, S. Liu, D.A. Bernlohr, E.R. Sauter, M.P. Cleary, J. Suttles, B. Li, Stearic acid induces CD11c expression in proinflammatory macrophages via epidermal fatty acid binding protein, J. Immunol. Baltim. Md 1950 200 (2018) 3407–3419.