AI-Assisted Identification of Primary and Secondary Metabolomic Markers for Post-Operative Delirium

Vladimir Lomivorotov

v.lomivorotov@gmail.com

  Penn State Hershey Medical Center, Penn State College of Medicine

Vladimir Ivanisenko

  Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk State University, Kurchatov Genomic Center of ICG SB RAS, Novosibirsk, Russia

Aelita-Luiza Makarova

  Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

Artem Rogachev

  Novosibirsk State University, N. N. Vorozhtsov Institute of Organic Chemistry, Novosibirsk, Russia
  https://orcid.org/0000-0002-3338-8529

Nikita Basov

  Novosibirsk State University, N. N. Vorozhtsov Institute of Organic Chemistry, Novosibirsk, Russia
  https://orcid.org/0000-0001-6390-5796

Evgeniy Gaisler

  Novosibirsk State University, N. N. Vorozhtsov Institute of Organic Chemistry, Novosibirsk, Russia

Irina Kuzmicheva

  Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk State University, Novosibirsk, Russia

Pavel Demenkov

  Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk State University, Kurchatov Genomic Center of ICG SB RAS, Novosibirsk, Russia

Artur Venzel

  Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk State University, Kurchatov Genomic Center of ICG SB RAS, Novosibirsk, Russia
  https://orcid.org/0000-0002-7419-5168

Timofey Ivanisenko

  Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk State University, Kurchatov Genomic Center of ICG SB RAS, Novosibirsk, Russia

Evgenia Antropova

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

Margarita Naidanova

Novosibirsk State University, Novosibirsk, Russia

Nikolay Kolchanov

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk State University, Novosibirsk, Russia

Alexey Kochetov

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk State University, Novosibirsk, Russia

Victoria Plesko

E. Meshalkin National Medical Research Center, Novosibirsk, Russia; Department of Anesthesiology and Intensive Care

Gleb Moroz

E. Meshalkin National Medical Research Center, Novosibirsk, Russia; Department of Anesthesiology and Intensive Care

Andrey Pokrovsky

Novosibirsk State University, Novosibirsk, Russia

Article

Additional Declarations: There is **NO** Competing Interest.

Vladimir A. Ivanisenko, Artem D. Rogachev, Aelita-Luiza A. Makarova, Nikita V. Basov, Evgeniy V. Gaisler, Irina N. Kuzmicheva, Pavel S. Demenkov, Artur S. Venzel, Timofey V. Ivanisenko, Evgenia A. Antropova, Margarita A. Naidanova, Nikolay A. Kolchanov, Alexey V. Kochetov, Victoria V. Plesko, Gleb B. Moroz, Andrey G. Pokrovsky, Vladimir V. Lomivorotov*

*corresponding author

AI-Assisted Identification of Primary and Secondary Metabolomic Markers for Post-Operative Delirium

Abstract
Despite considerable investigative efforts, the molecular mechanisms of postoperative delirium (POD) remain unresolved. The present investigation employs innovative methodologies for identifying potential primary and secondary metabolic markers of POD by analyzing serum metabolomic profiles utilizing the genetic algorithm and artificial neural networks. The primary metabolomic markers constitute a combination of metabolites that optimally distinguish between POD and non-POD groups of patients. Our analysis revealed L-lactic acid, inositol, and methylcysteine as the most salient primary markers, upon which the prediction accuracy of POD manifestation achieved AUC=99%. The secondary metabolomic markers represent metabolites that exhibit perturbed correlational patterns within the POD group. We identified 54 metabolites as the secondary markers of POD, incorporating neurotransmitters such as gamma-aminobutyric acid (GABA), serotonin. These findings imply a systemic disruption in metabolic processes in patients with POD. The deployment of gene network reconstruction techniques facilitated the postulation of hypotheses describing the role of established genomic POD markers in the molecular-genetic mechanisms of metabolic pathways dysregulation, involving the identified primary and secondary metabolomic markers. This study not only expands the understanding of POD pathogenesis but also introduces a novel technology for bioinformatic analysis of metabolomic data which could aid in uncovering potential primary and secondary markers in diverse research domains.

Main

Postoperative delirium (POD) is an acute state typified by cognition and attention deficits, disorganized thoughts, and disorientation, and is a frequent implication of assorted surgical procedures. Accumulating evidence associates POD with detrimental outcomes, such as prolonged hospital stays, augmented risk of complications, and escalated mortality rates[1,2] The propensity to POD manifestation escalates concurrently with age, as exemplified by a study reporting a 52% incidence of POD in cardiac surgery patients aged 60 years or older[3]. Risk factors for POD also include duration of the general anesthesia during surgery, neurotransmitter imbalance, neuroinflammation, metabolic syndrome, and others[4,5].

Metabolomics, by providing a snapshot of current biochemical activity, has been instrumental in diverse areas, such as profiling disease biomarkers[6,7], monitoring disease progression[8], elucidating xenobiotic metabolism[9] and assessing drug toxicity[10]. This sets metabolomics apart from genomics and proteomics, which only offer potential scenarios of conditions development[11].

The levels of individual metabolites and the interplay between various metabolites can serve as indicators of physiological and pathological transformations within an organism[12]. The stratification of metabolic profiles into predefined classes based on individual metabolites or the sets of metabolites forms the crux of extant computational

1

biomarker discovery methods[13]. Such indicators referred to as primary biomarkers could significantly contribute to diagnostic and therapeutic strategies.

Perturbations in the interplay among metabolites can offer insightful explanations of metabolic pathway dysfunction across an array of pathologies at a systemic level[14]. Metabolites that exhibit disrupted coordinated variability with their counterparts may be designated as secondary biomarkers[15,16]. These secondary biomarkers can not only expand our knowledge of the molecular mechanisms underlying physiological or pathophysiological processes, but also improve the predictive capacity of primary biomarkers[17].

Evaluation of metabolites levels and analysis of the correlations between metabolite concentrations is an applicable strategy to encapsulate the intricate relationships between metabolites[18]. Xiao et al. put forward a method designed to analyze a network of pairwise correlations among metabolites, premised on the Generalized Singular Value Decomposition (GSVD) algorithm, computed separately for POD and non-POD groups[19]. This approach enables discovery of clusters of metabolites predicated on mutual correlations, which differ between patient and control groups. Correlation analysis has also been exploited to establish the linkage between metabolites and phospholipid species, grounded on metabolomic and lipidomic data derived from Prostate Cancer Tissue[20].

In previous investigations, our team executed liquid chromatography with tandem mass spectrometry (LC-MS/MS) analysis of cerebrospinal fluid (CSF) and blood plasma samples collected from patients afflicted with high-grade glioma[21]. Consequently, we uncovered correlations between the metabolic profiles of blood plasma and CSF.

The identification of interrelated metabolites from the analysis of metabolomic data is challenging due to the systemic control of metabolic processes. Generally, the pairwise correlations observed between metabolites are insignificant, even when the pairs are part of the same metabolic pathway[22–24]. This phenomenon could be attributed to the inherent stochasticity of metabolic processes[25–28]. Thus, explicit relationships between metabolites may be hidden, adding another layer of complexity to the analysis.

Machine learning techniques, notably those encompassing artificial intelligence, are increasingly being harnessed for the analysis of high-dimensional omics data [29]. Autoencoders (AEs) are a class of unsupervised neural network architectures devised for dimensionality reduction that can capture non-linear relationships[30]. Autoencoder structure typically contains an input layer of neurons, one or more hidden layers, and an output layer, replicating the input data. The goal of autoencoder is to minimize the reconstruction error between input and output data.

Autoencoders have been demonstrated to adeptly manage the dimensionality reduction of original data by establishing non-linear relationships between features in the input data. Several variants of autoencoder architecture have been devised, such as convolutional, regularized, variational, sparse, stacked, deep, and generative[31]. In particular, variational autoencoder-derived latent representations of metabolomic data have been employed to analyze the groups of patients with conditions such as type 2 diabetes, acute myeloid leukemia, and schizophrenia[32]. Likewise, a supervised autoencoder (SAE) approach has been utilized for classification of clinical metabolomic data[33].

Denoising autoencoders are a subclass of neural network models that are a viable method for generating lower-dimensionality latent representations of salient [34]. For instance, a deep learning-empowered denoising autoencoder has been deployed to gauge peak quality in liquid chromatography – high resolution mass spectrometry (LC−MS) data, as well as to forecast peaks by eliminating noise from the original peak profiles[35]. Similarly, a denoising autoencoder (SERDA) has been used to rectify errors in a large-scale metabolomic data generated via gas chromatography–mass spectrometry[36]. Furthermore, a normalization autoencoder (NormAE), a brainchild of nonlinear AEs and adversarial learning, has been

utilized to expunge batch effects in LC-MS-based metabolomics data[37]. Additionally, a number of methodologies for analyzing metabolic pathways and gene networks have been proposed to decode the molecular genetic mechanisms underlying observed metabolomic data[13]. The MetaboAnalyst 5.0 web tool, for instance, provides a comprehensive toolkit for the analysis of metabolomic data in the context of metabolic pathways[38].

Gene networks can confer valuable insights regarding genetic regulation of identified metabolic pathways, thereby forming the basis for integration of metabolomic and genomic data. Previously, we designed ANDSystem, a computer tool for reconstruction of gene networks, based on automated extraction of knowledge from scientific publications and factual databases[39–42]. ANDSystem has been applied to diverse problems, ranging from reconstruction of host-virus interaction gene networks[43,44], gene prioritization[45,46], to identification of new potential pharmacological targets[47]. Notably, utilizing ANDSystem to evaluate serum metabolomic profiles derived from COVID-19 patients, facilitated the identification of potential regulatory pathways and the key SARS-CoV-2 proteins implicated in disruption of metabolic processes during the course of infection[48].

In this study, we performed a metabolomic analysis of blood plasma acquired from two distinct patient groups who underwent cardiac surgery. One group developed postoperative delirium, whilst the other group did not exhibit this complication. Blood samples were procured prior to surgery, as we examined the data regarding preoperative metabolomic profiles. The metabolomic profiles derived from these samples revealed potential primary and secondary biomarkers.

The primary biomarkers were identified as composite metabolites that markedly differentiated patients with postoperative delirium from those without this condition. The genetic algorithm was employed to discern these combinations and several groups consisting of 3 to 4 metabolites demonstrated comparable ability to distinguish patients exhibiting delirium from those who did not. Remarkably, two metabolites, L-lactic acid and Inositol, were present in nearly all combinations, with inclusion rates of 100% and 95%, respectively. Other metabolites, Methylcysteine and Adenine, were included in 40% and 12% of combinations, respectively. In total, 37 distinct metabolites were involved in the combinations.

For identification of the secondary markers, we suggest a method predicated on detection of anomalies within data, utilizing a denoising autoencoder[49]. In general, utilizing autoencoders for anomaly detection has been substantiated by multiple studies[49–53]. The secondary markers represent metabolites, for which the denoising autoencoder trained on metabolomic profiles from non-POD patients failed to replicate metabolite concentrations for patients in the POD group. Such aberrations in the autoencoder's outputs can be attributed to the disruption of encoded intricate non-linear relationships between metabolites concentrations. These relationships allow to reduce dimensionality of metabolomic profile input within the neural network's hidden layer without substantial loss of prediction accuracy in the output.

We herein suggest a novel approach to identify the secondary metabolomic biomarkers, deploying a digital patient model predicated on the use of denoising autoencoder applied to metabolomic data. The autoencoder detects metabolites whose interconnections with other metabolites, specific to the non-POD group, were disrupted in patients manifesting delirium. We identified 54 metabolites underscoring the systemic nature of metabolic aberrations in delirium patients. Interestingly, there were overlaps between the primary and secondary biomarkers, as evidenced by 8 shared metabolites, including inositol, Gamma-Aminobutyric Acid (GABA), biotin, and others. Of particular interest, the primary metabolomic markers of POD included neurotransmitters such as GABA, serotonin and its precursors L-Tryptophan and 5-Hydroxy-L-tryptophan. Given that some metabolites do not

cross the blood-brain barrier (BBB), we postulate that BBB permeability may be compromised due to neuroinflammatory processes in patients with POD. Notably, the outflow of CSF serotonin into blood plasma was insufficient for its classification as a primary marker, yet it was significant enough for inclusion among the secondary markers.

Upon performing an overrepresentation analysis of Kyoto Encyclopedia of Genes and Genomes (KEGG) metabolic pathways based on the entire list of primary biomarker metabolites, no statistically significant results were obtained. Nevertheless, when the primary and secondary biomarkers were combined, metabolic pathways registering lowest p-values for primary biomarkers, such as "Tryptophan metabolism" and "Pyrimidine metabolism" emerged as statistically significant. Consequently, the secondary biomarkers amplified the statistical significance of enriched metabolic pathways, as determined by analysis of the primary markers.

Utilizing ANDSystem for gene network reconstruction, we inferred the potential roles of established POD genetic markers in regulation of the identified metabolic pathways. Several genetic markers of POD, including IFNG, TNFA, LEP, and IL6 had a substantial regulatory impact on metabolic pathways.

Results

Preliminary Statistical Analysis

Following the preprocessing of LC-MS/MS metabolomic data of preoperative blood plasma samples, 210 metabolites were selected for evaluation (Supplementary Table S1). Based on post-surgical observations, all patients were classified into POD and non-POD groups. An initial comparative statistical analysis was carried out on the obtained metabolomic data of both patient groups for each of the 210 metabolites using the Mann-Whitney test (Supplementary Table S2). This analysis failed to reveal statistically significant differences between the two patient groups. The lack of detection of individual metabolomic markers could potentially be attributed to the complexity of the molecular genetic mechanisms underlying predisposition to POD, which likely involve the interplay of multiple metabolites.

Identification of the Primary Metabolomic Markers

Due to the incapacity of the basic statistical analysis to identify individual metabolomic markers, we next employed the methods that could facilitate the identification of not just distinct metabolites, but their combinations. For this purpose, we utilized a genetic algorithm method as described by Lu et al.[54]. The fitness of the individual in this method was assessed using an XGBoost model, which was provided with different combinations of metabolite concentrations generated by the "mutations" and "recombinations" facilitated by the genetic algorithm. This strategy yielded combinations of metabolites that optimally differentiated between patients with and without POD based on the concentration values of these metabolites. Each of these combinations comprised 3 to 4 metabolites. The accuracy of all combinations, as estimated via cross-validation, was found to be 83.3%.

The aggregated list of metabolites encompassed within these combinations amounted to 37 unique entities (Supplementary Table S3). It is noteworthy that L-lactic acid was a constituent in all combinations, while Inositol featured in 95%, Methylcysteine in 40%, and Adenine in a mere 12%. The area under the ROC curve (AUC) for the metabolite triad of L-lactic acid, Inositol, and Methylcysteine was 0.989 (Fig. 1). Consequently, these metabolites may be considered as pivotal biomarkers for the onset of POD.
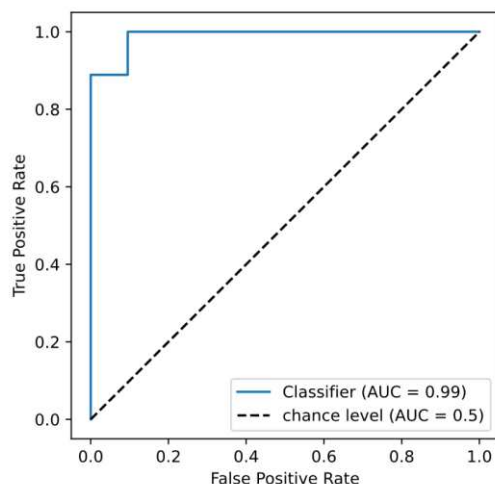
4

Fig. 1. Classification accuracy for distinguishing between postoperative delirium (POD) and non-POD patient groups.

We examined overrepresented KEGG metabolic pathways associated with the consolidated list of primary biomarkers using MetaboAnalyst 5.0. The KEGG pathways "Pyrimidine metabolism", "Tryptophan metabolism", and "Purine metabolism" were enriched with primary markers, however, the FDR values were insignificant (Table 1).

Table 1. Analysis of overrepresented KEGG metabolic pathways associated with the set of primary metabolomic markers.

KEGG pathway	Total	Metabolomic Markers (Hits)	p-value	FDR
Pyrimidine metabolism	39	3	0.006	0.256
Tryptophan metabolism	41	3	0.006	0.256
Purine metabolism	70	3	0.028	0.616

The enrichment analysis aims to identify metabolic pathways that exhibit systemic disturbances. However, indicators of the metabolic pathway disruptions can comprise not only increased or decreased concentrations of the individual metabolites, but also the disturbances of coordination or correlation between the concentrations of metabolites implicated in the metabolic pathway[14]. Metabolites exhibiting disrupted correlations with other metabolites in the context of a pathological process may be termed as the secondary metabolomic markers[16].

Identification of the Secondary Metabolomic Markers

To disclose the secondary metabolomic markers, we utilized an anomaly detection methodology via denoising autoencoder models, as delineated by Sakurada et al.[49]. This application of autoencoder, inherently proficient at deciphering multiple nonlinear correlations amongst features, holds a distinct advantage over the linear pairwise correlation methods such as principal component analysis (PCA). The topology of an autoencoder includes the input layer, hidden layers, and the output layer, in which neuron values reconstruct the input neuron values. We herein employed a three-layer denoising

autoencoder. The metabolite concentrations in the non-POD patient group served as the input data for autoencoder training.

We further adopted the strategy suggested by Sakurada et al.[49] to identify the secondary markers. This involved training the autoencoder model on the established dataset, followed by model application to a new dataset harboring anomalies that were absent in the training set. Features were considered as anomalous if their reconstruction employing a denoising autoencoder yielded significant discrepancies at the output layer values compared to the input layer values. We examined various values for the number of neurons in the hidden layer (Supplementary Table S4). A total of 100 autoencoder models were trained at different parameter values. The trained models were subsequently applied to the metabolomic profile data of POD patients. Metabolites whose concentrations were predicted by autoencoder with statistically significant deviations from the input data were classified as the secondary markers.

Despite achieving nearly identical training accuracy across the different parameters, application of these models yielded varying sets of the secondary markers. However, certain metabolites recurred frequently. We deemed metabolites resilient to the model parameter changes as the most crucial markers. The list of metabolites, ranked by the number of autoencoder models that identified them as the secondary markers, is presented in Supplementary Table S5. We conducted an overrepresentation analysis of KEGG metabolic pathways for 26 metabolites uncovered by 100 autoencoder models, however, the FDR values were insignificant (Table 2).

Table 2. Overrepresented KEGG metabolic pathways for the set of secondary markers identified by 100 autoencoder models.

KEGG Metabolic pathway	Total	Hits	P-value	FDR
beta-Alanine metabolism	21	2	0.006	0.403
Glycine, serine and threonine metabolism	33	3	0.015	0.403

Of great interest is a set of metabolites that were identified in fewer numbers of models. A list of metabolites identified by at least 97 models included serotonin and was selected as the secondary markers set. The enrichment analysis of KEGG metabolic pathways for the resulting list of 54 secondary markers revealed the same metabolic pathways with statistical significance (Table 3).

Table 3. Overrepresented KEGG metabolic pathways for the set of the secondary markers identified by 97 autoencoder models.

KEGG Metabolic pathway	Total	Hits	P-value	FDR
Glycine, serine and threonine metabolism	33	4	2.03E-4	0.0163
beta-Alanine metabolism	21	3	8.99E-4	0.036

The overlap between the primary and secondary metabolomic markers is noteworthy, with 8 metabolites identified as both primary and secondary markers (Table 4). This intriguing intersection suggests that these dual-status entities may hold a privileged position in the metabolic dysregulation associated with postoperative delirium. For instance, inositol

was identified as both primary and secondary marker and was previously proposed as a prognostic marker of POD.

Table 4. Metabolites serving as both the primary and secondary metabolomic markers.

KEGG Id	Metabolite	Primary marker importance
C00137	Inositol	87
C00120	Biotin	3
C01089	3-hydroxybutyric acid	1
C00318	Carnitine	1
C00334	Gamma-Aminobutyric acid	1
C00258	Glyceric acid	1
C00106	Uracil	1
C00438	Ureidosuccinic acid	1

According to our definition, the secondary markers are characterized by disrupted inter-metabolite correlation relationships. We hypothesized that these relationships could primarily be disrupted in the metabolic processes involving the primary metabolomic markers. Consequently, the conducted overrepresentation analysis for the combined list of primary and secondary metabolomic markers would yield more accurate identification of the metabolic pathways disrupted in patients with POD.

To validate this hypothesis, we merged the lists of primary and secondary biomarkers to further analyze overrepresented KEGG metabolic pathways. Consistent with our expectations, the overrepresentation analysis of metabolic pathways for the combined list of primary and secondary metabolomic markers revealed statistically significant results. After multiple comparisons adjustment "Glycine, serine, and threonine metabolism", "Aminoacyl-tRNA biosynthesis", and others were present among the significant metabolic pathways (Table 5). It is worth noting that the "Glycine, serine, and threonine metabolism" pathway previously exhibited statistical significance in the overrepresentation analysis for the set of secondary metabolomic markers (Table 3).

Table 5. Overrepresented KEGG metabolic pathways for the combined set of primary and secondary metabolomic markers. Metabolites that are primary and secondary markers are marked in bold.

KEGG pathway	Total	FDR	Primary Metabolomic Markers (Hits)	Secondary Metabolomic Markers (Hits)
Glycine, serine and threonine metabolism	33	0.0021	Betaine; **Glyceric acid**	L-Serine; Choline; Betaine aldehyde; Guanidoacetic acid; Cystathionine; L-Threonine; **Glyceric acid**
Aminoacyl-tRNA biosynthesis	48	0.0028	L-Leucine; L-Tryptophan; L-Tyrosine	L-Histidine; L-Arginine; L-Serine; L-Valine; L-Threonine; L-Proline

				Serotonin; 3-Hydroxyanthranilic acid; L-Kynurenine; Acetyl-CoA
Tryptophan metabolism	41	0.0239	L-Tryptophan; 5-Hydroxy-L-tryptophan	Serotonin; 3-Hydroxyanthranilic acid; L-Kynurenine; Acetyl-CoA
Arginine biosynthesis	14	0.0351	N-Acetylglutamic acid	L-Arginine; Citrulline; Ornithine
Butanoate metabolism	15	0.0372	**3-hydroxybutyric acid; Gamma-Aminobutyric acid**; D-2-Hydroxyglutaric acid	**3-hydroxybutyric acid**; Acetyl-CoA; **Gamma-Aminobutyric acid**
Valine, leucine and isoleucine biosynthesis	8	0.0375	L-Leucine	L-Threonine; L-Valine
Arginine and proline metabolism	38	0.0375	**Gamma-Aminobutyric acid**	L-Arginine; Guanidoacetic acid; **Gamma-Aminobutyric acid**; Spermine; Ornithine; L-Proline
Pyrimidine metabolism	39	0.0376	Cytidine; Thymine; **Ureidosuccinic acid; Uracil**	Uridine; **Ureidosuccinic acid**; **Uracil**; Deoxyribose 1-phosphate
Ubiquinone and other terpenoid-quinone biosynthesis	9	0.0394	L-Tyrosine; Homogentisic acid	Phenyllactic acid

In the assessed KEGG metabolic pathways, "Glycine, serine and threonine metabolism" pathway emerged as the most statistically significant. This pathway incorporated two primary and seven secondary markers. A comparable proportion of primary to secondary markers was discernible among the metabolites of "Aminoacyl-tRNA biosynthesis" pathway.

Reconstruction of Gene Networks of Overrepresented KEGG Metabolic Pathways Regulation by Genetic Markers of POD

At the next stage we extracted the genetic markers of POD from existing scientific reports (Supplementary Table S6). A compilation of 45 genes that were cited as POD markers and their role in the pathogenesis of POD extensively deliberated, were selected for analysis. Many of these genetic markers of POD are inflammatory factors, and their escalated levels have been documented in the plasma of POD patients in the preoperative and/or early postoperative period[55,56]. For instance, elevated plasma concentrations of interleukin 6 (IL6), tumor necrosis factor alpha (TNFA), and vascular endothelial growth factor (VEGF) have been detected in the early postoperative period in elderly patients who exhibited POD[57].

We proposed that POD genetic markers could play a substantial role in the mechanisms of POD pathogenesis, potentially through their involvement in metabolic pathway disruption. To validate this hypothesis, we reconstructed molecular genetic

regulatory pathways using ANDSystem, according to the Templates shown in Table 12. Each of the reconstructed regulatory pathways starts with proteins encoded by POD genetic markers, involves intermediate components – human proteins and genes – and ends with metabolic pathway enzymes. We considered a set of interactions between regulatory pathway participants, including regulation of gene expression, regulation of protein activity/degradation/transport, protein-protein interactions, and others. Consequently, these pathways can elucidate numerous regulatory events through which POD genetic markers could modulate the activity of metabolic pathway enzymes. We aimed to discover potential regulatory pathways for "Glycine, serine, and threonine metabolism" and "Aminoacyl-tRNA biosynthesis" by genetic markers of POD, as these metabolic pathways were identified as overrepresented with FDR<0.01. The gene networks representing regulation of the enzymes of "Glycine, serine, and threonine metabolism" and "Aminoacyl-tRNA biosynthesis" metabolic pathways are illustrated in Fig. 3.
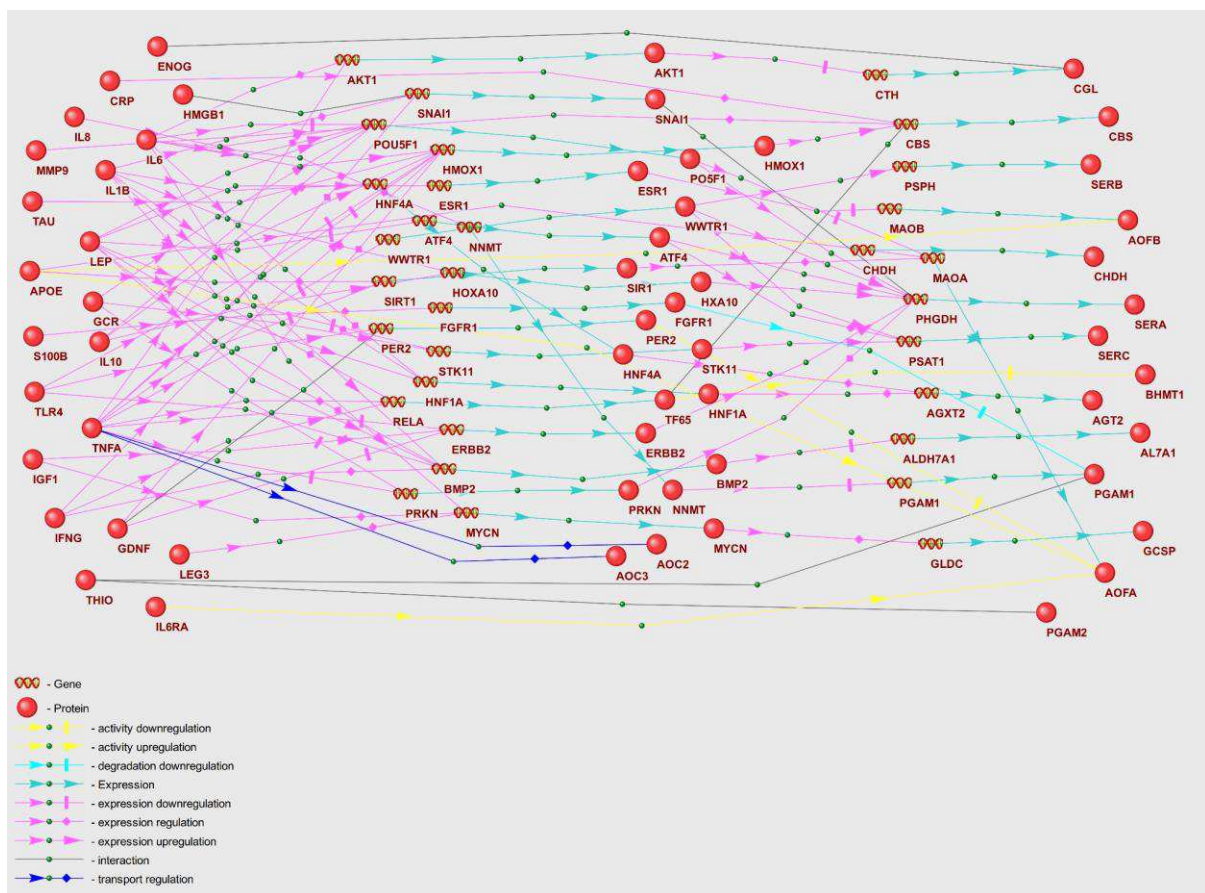

Fig. 2. Gene network of regulation of "Glycine, serine, and threonine metabolism" enzymes by POD genetic markers.
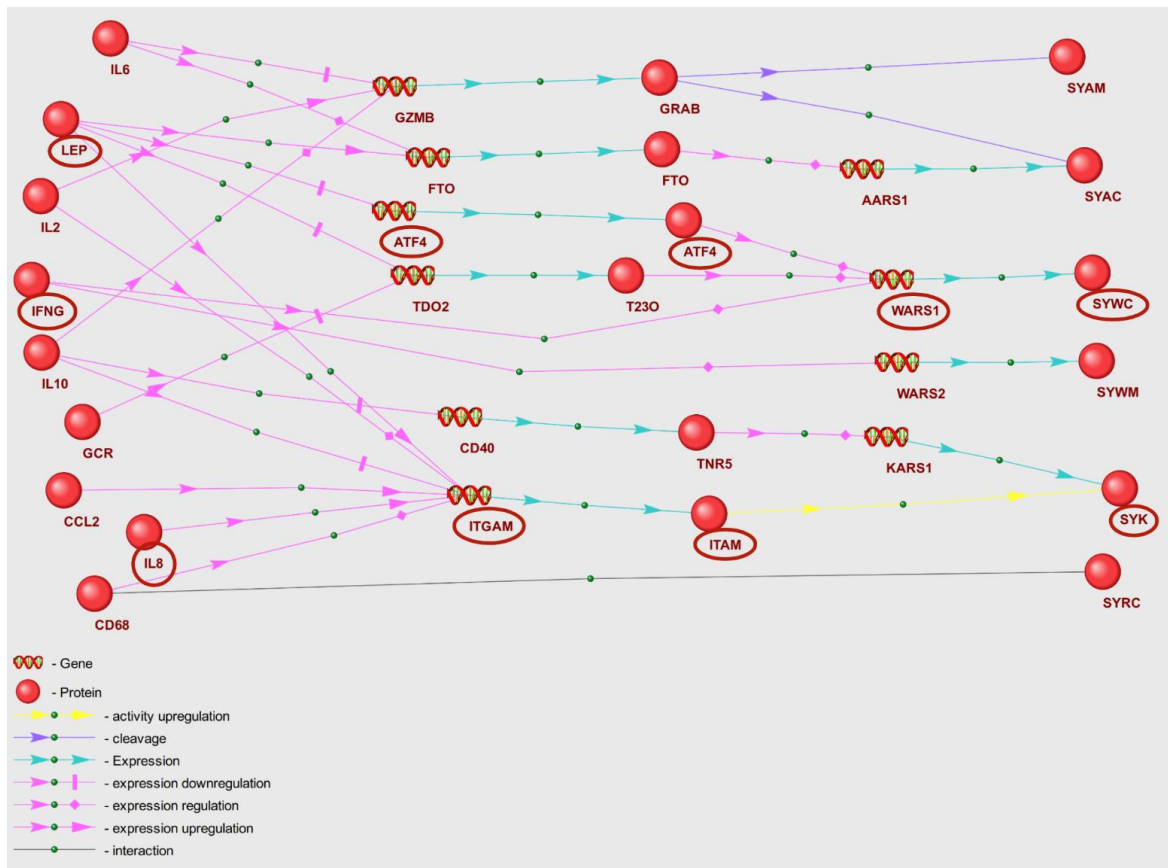
Fig. 3. Gene network of regulation of "Aminoacyl-tRNA biosynthesis" metabolic pathway enzymes by POD genetic markers.

A graph of a gene network includes linear chains and numerous branches, as well as chain convergences, that denote interactions between entities. The objects outlined in red ovals on Fig. 3 illustrate the participants of the example regulatory pathways within the gene network. IL8, the genetic marker of POD, modulates the activity of tyrosine protein kinase SYK along a linear chain through upregulation of the expression of intermediate gene *ITGAM*. Through a branched segment of gene network, tryptophan t-RNA ligase (*WARS1*) expression is directly upregulated by interferon gamma (IFNG), while leptin (LEP) modulates *WARS1* expression through downregulation of intermediate gene *ATF4*. As a result, the gene network graphs facilitate the elucidation of the extent to which POD genetic markers exert regulatory influences on the enzymes of each of the overrepresented KEGG pathways.

Analysis of the gene networks (Figs. 2, 3) reveals that 21 genetic markers regulate enzymes of the "Glycine, serine, and threonine metabolism" pathway and 9 genetic markers regulate enzymes of "Aminoacyl-tRNA biosynthesis". Table 6 contains characteristics of the regulatory pathways, including the number of POD genetic markers that modulate enzymes involved in KEGG metabolic pathways (N1), along with the number of enzymes under regulation (N2). In Table 6, certain genetic markers are represented in more than one template for the gene network of each metabolic pathway (a single marker may be included in several templates). Among the metabolic pathways whose enzymes are most regulated, the "Glycine, serine, and threonine metabolism" pathway emerged as the most regulated. Comparable statistics can be calculated for the regulatory pathway Templates. The regulation of enzymes by genetic markers is executed to the greatest extent through pathways built according to the double regulation of expression template (P4).

Table 6. Quantitative indicators of POD genetic markers (N1) and their corresponding regulated metabolic pathway enzymes (N2) as per regulatory pathway templates.

KEGG pathway	Regulatory pathways									
	P1		P2		P3		P4		P5	
	N1	N2	N1	N2	N1	N2	N1	N2	N1	N2
Glycine, serine, and threonine metabolism	2	3	3	5	3	3	15	13	5	3
Aminoacyl-tRNA biosynthesis	1	1	0	0	1	2	10	3	5	2

In delineating the influence of distinct genetic markers on the regulation of metabolic pathways, we quantified the number of enzymes regulated by each of the genetic markers (Fig. 4). It is noteworthy that 16 out of 23 markers exhibited a preferential affinity for specific metabolic pathways. For instance, TNFA specifically regulated the enzymes of "Glycine, serine, and threonine metabolism" pathway. Alternatively, genetic markers LEP, IL6, IL10, IFNG, IL8, GCR, and TAU exerted regulation over the both overrepresented metabolic pathways.
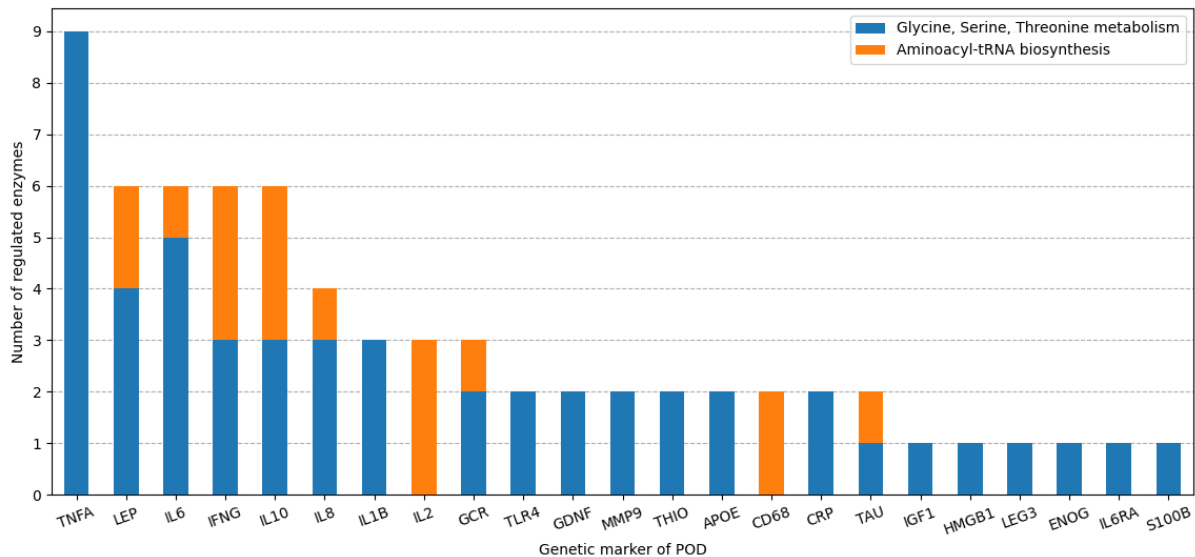

Fig. 4. The number of regulated enzymes in KEGG metabolic pathways for each of the POD genetic markers.

It is critical to underscore that gene regulatory networks were reconstructed using the ANDSystem, which is based on the text-mining approaches, and the statistics derived may be incomplete, as the reconstruction leveraged automated tools within the ANDSystem knowledge base. A characteristic of text mining applications is their ability mostly to discern regulatory interactions that are well-studied and comprehensively documented in the literature.

Reconstruction of the Gene Regulatory Networks for Metabolomic Markers not Present in Overrepresented Metabolic Pathways

It is worth noting that not all of the POD metabolomic markers identified by our study were found to participate in the overrepresented metabolic pathways. Among these are L-lactic acid and Inositol that assume significant roles as the primary markers. To detail the regulation of these metabolites by POD genetic markers, we reconstructed gene networks of

regulation of the enzymes involved in lactate and myo-inositol metabolism. We retrieved enzymes implicated in lactate and myo-inositol metabolism from the KEGG database. To locate these enzymes, we queried the KEGG Pathway database (https://www.genome.jp/kegg/pathway.html) using the identifiers C00186 (lactate) and C00137 (myo-inositol) and obtained 12 and 7 metabolic pathways for these metabolites, respectively. Only human metabolic pathways with the 'hsa' prefix were considered. The search list for subsequent analysis comprised enzymes that directly catalyze reactions involving the target metabolite (myo-inositol or lactate), as well as enzymes catalyzing upstream and downstream reactions. We accounted for one to two intermediate upstream reactions and a single downstream reaction link and identified nine enzymes associated with lactate metabolism, as well as 27 enzymes associated with myo-inositol metabolism (Tables 7, 8). To reconstruct the regulatory pathways that elucidate the mechanisms of POD genetic marker influence on the identified enzymes, we employed the same Templates that were used for analysis of the overrepresented metabolic pathways (Table 12). Notably, the final links of the Templates were enzymes involved in lactate or myo-inositol metabolism pathways (Tables 7, 8).

Gene regulatory networks including the enzymes involved in lactate and myo-inositol metabolism are depicted in Figs. 5 and 6, respectively. To describe these gene networks, we compiled tables containing POD genetic markers and the enzymes regulated by them along with the names of KEGG metabolic pathways (Tables 7, 8).
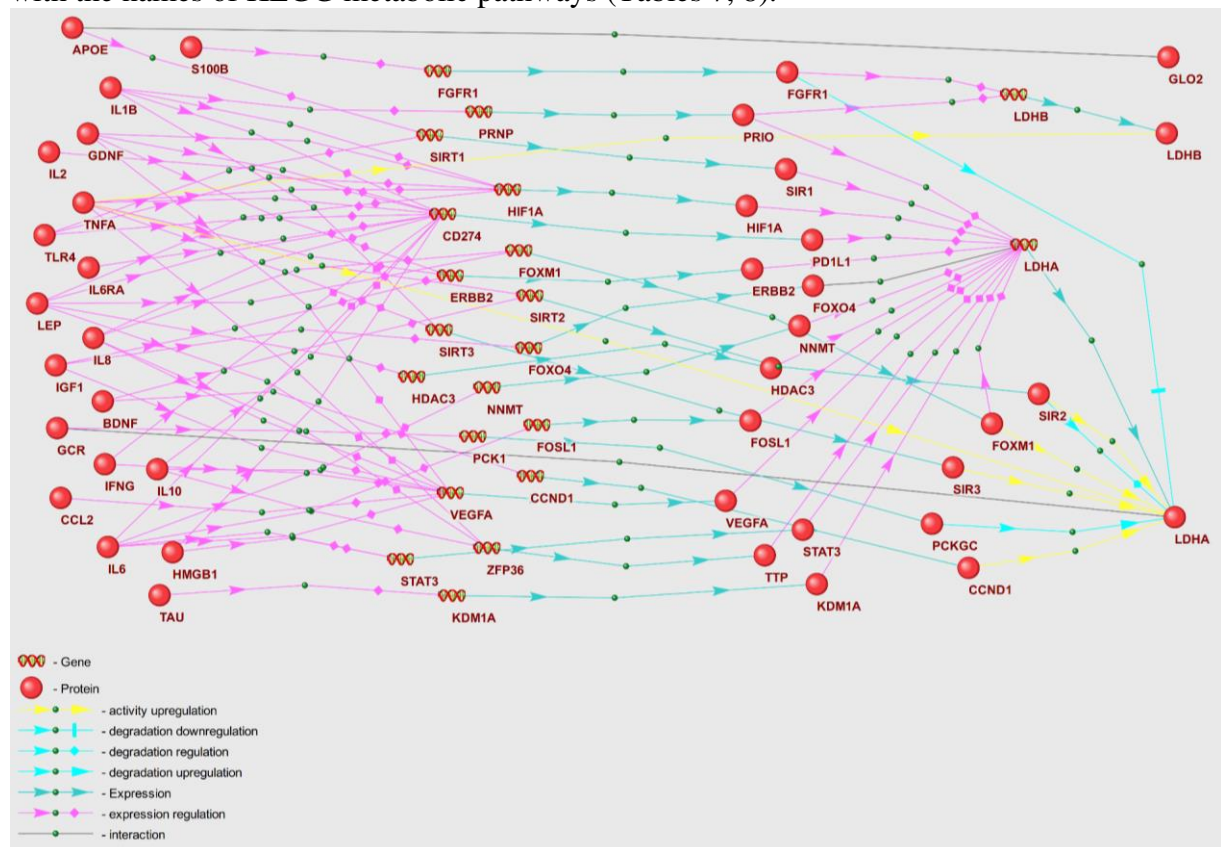

Fig. 5. Gene network of regulation of lactate metabolism enzymes by POD genetic markers.

Table 7. Descriptive features of gene network of regulation of lactate metabolic pathway enzymes.

№	KEGG lactate metabolism pathways	Enzymes in the metabolic pathway	Regulatory genetic markers

1	Glycolysis / Gluconeogenesis: hsa00010;	LDHA	APOE, CCL2, GDNF, HMGB1, IGF1, IFNG, IL1B, IL2, IL6, IL6RA, IL8, IL10, LEP, TLR4, TNFA
		LDHB	IL1B, S100B, TNFA
2	HIF-1 signaling pathway: hsa04066	LDHC, LDHAL6A, LDHAL6B	-
3	Pyruvate metabolism: hsa00620	GLO2	APOE
		LDHA	APOE, CCL2, GDNF, HMGB1, IGF1, IFNG, IL1B, IL2, IL6, IL6RA, IL8, IL10, LEP, TLR4, TNFA
		LDHB	IL1B, S100B, TNFA
		GLUL, GRHPR, HAGH, LDHC, LDHD, LDHAL6A, LDHAL6B	-

As illustrated in Table 7, lactate dehydrogenase chain A (LDHA) is under the most profound regulatory control, with 15 POD genetic markers exerting influence. A total of 19 genetic markers take part in modulation of the enzymes pertinent to lactate metabolic pathways. Concurrently, APOE, IL1B, and TNFA each regulate two corresponding enzymes implicated in lactate metabolism pathways.
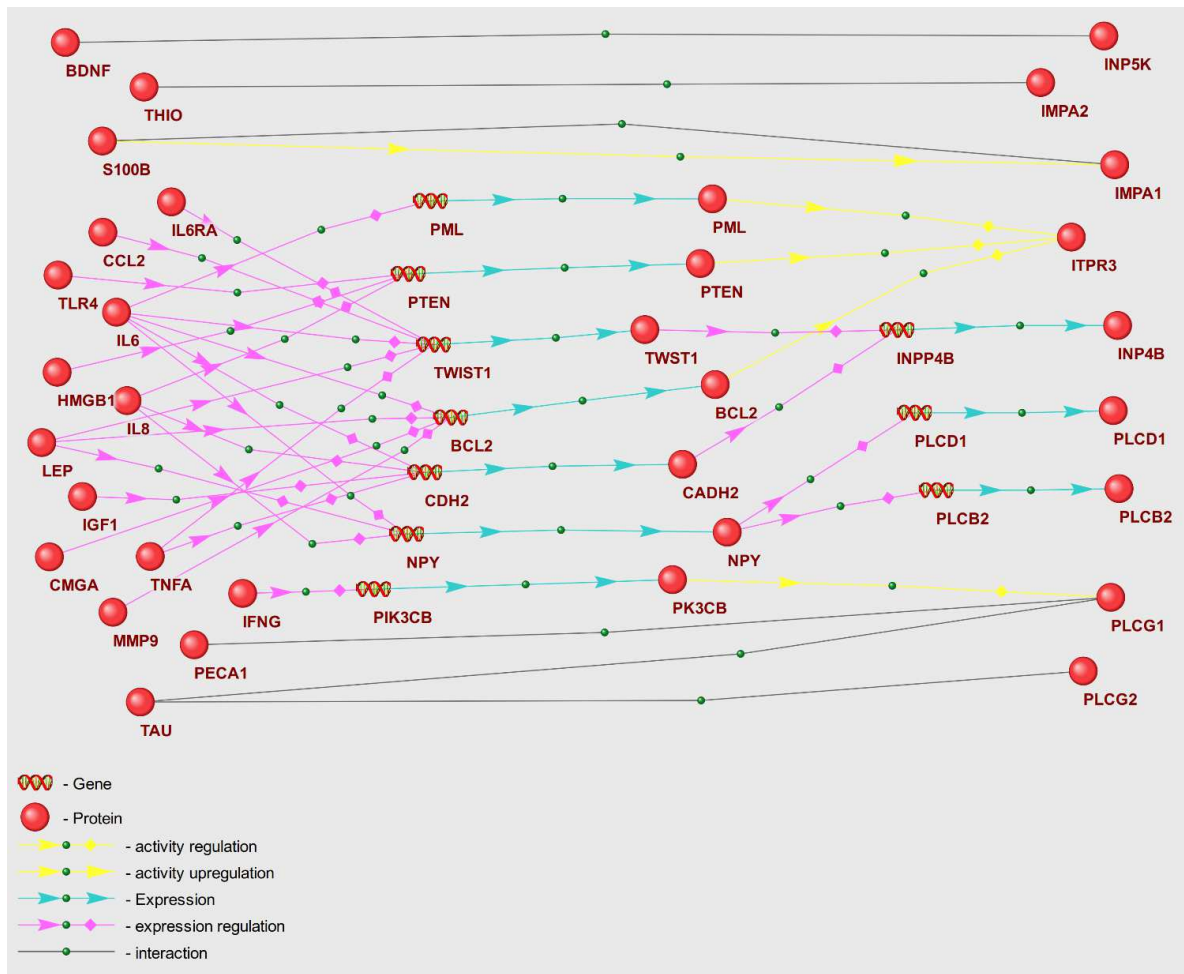
Fig. 6. Gene network of regulation of inositol metabolism enzymes by POD genetic markers.

As per the data shown in Table 8, the enzymes Inositol polyphosphate 4-phosphatase type II (INP4B) and Inositol 1,4,5-trisphosphate receptor type 3 (ITPR3) are primarily regulated, with six POD genetic markers directing regulatory links towards them. Fourteen POD genetic markers are implicated in the regulation of enzymes participating in inositol metabolism. Importantly, IL6, IL8, and LEP act as regulators of four enzymes – INP4B, ITPR3, PLCB2, and PLCD1 – involved in inositol metabolism pathways.

Table 8. Characteristics of gene regulatory network of inositol metabolic pathway enzymes.

№	KEGG inositol metabolism pathways	Enzymes in the metabolic pathway	Regulatory genetic markers
1	Inositol phosphate metabolism: hsa00562	IMPA1	S100B
		IMPA2	THIO
		INP4B	CCL2, IGF1, IL6, IL6RA, IL8, LEP
		INP5K	BDNF
		ITPR3	HMGB1, IL6, IL8, LEP, MMP9, TLR4
		CDIPT, INP1, INP4A, ISYNA1, MIOX, MTM1	-
2	Phosphatidylinositol signaling system:	IMPA1	S100B
		IMPA2	THIO

14

	hsa04070	INP4B	CCL2, IGF1, IL6, IL6RA, IL8, LEP
		PLCB2, PLCD1	IL6, IL8, LEP
		PLCG1	PECA1
		PLCG2	TAU
		BPNT2, CDIPT, INP1, INP4A, INP5A, MTM1, PLCs	-
3	Ascorbate and aldarate metabolism: hsa00053	MIOX	-

Among the set of metabolites, gamma-aminobutyric acid (GABA) merits specific consideration. Our LS-MS/MS metabolomic analysis revealed a decreased GABA concentration in the POD patients group. The gene network delineating the regulation of GABA metabolism enzymes by POD genetic markers is depicted in Fig. 7, with principal gene network components shown in Table 9. To reconstruct the gene network, we extracted 13 enzymes implicated in GABA metabolism from the KEGG database.
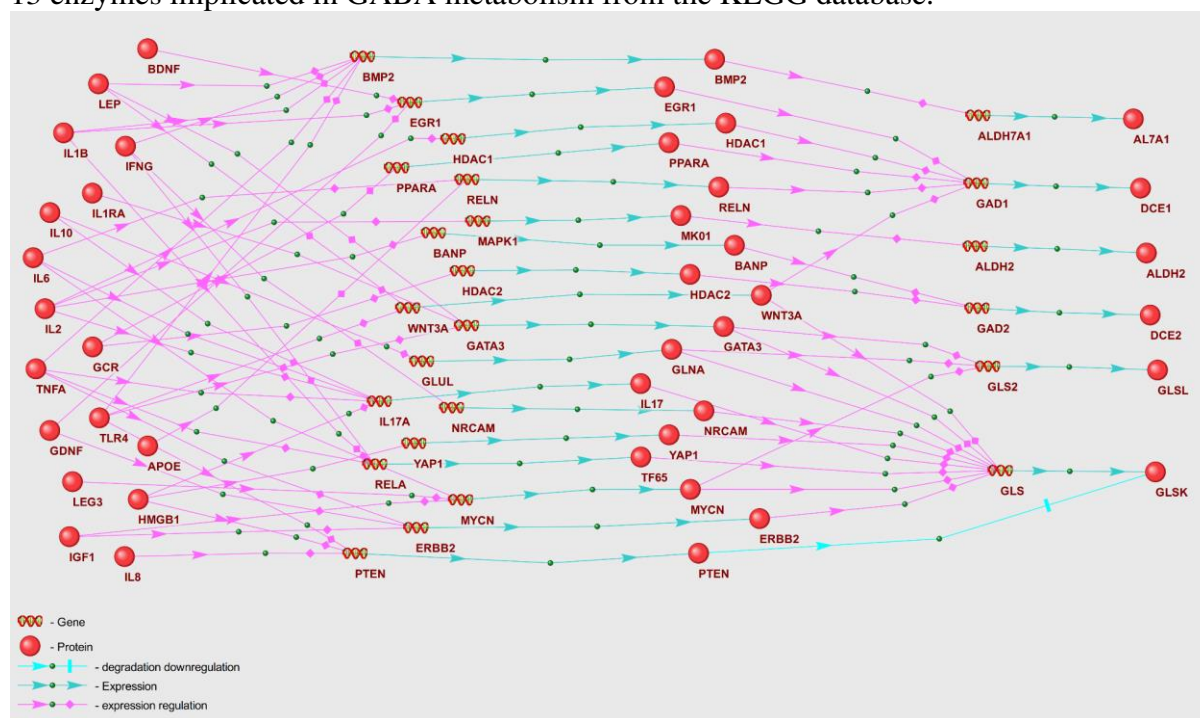

Fig. 7. Gene network of regulation of GABA metabolism enzymes by POD genetic markers.

As shown in Table 9, glutaminase (GLSK) emerges as the primary target of regulatory effects, with 14 POD genetic markers directing their influence towards it. 17 distinct POD genetic markers are implicated in regulation of GABA metabolism enzymes. Among these are glucocorticoid receptor (GCR) and interleukin 2 (IL2), each modulating four distinct enzymes involved in GABA metabolism.

Table 9. Characteristics of gene regulatory network of GABA metabolic pathway enzymes.

№	KEGG GABA metabolism pathways	Enzymes in the metabolic pathway	Regulatory genetic markers
1	GABAergic synapse: hsa04727	GLSK	GDNF, GCR, HMGB1, IGF1, IFNG, IL1B, IL1RA, IL2, IL6, IL8, IL10, LEG3, LEP, TLR4

		GLSL	GCR, IGF1, IL1RA, IL6, IL8, IL10, LEG3, LEP, TLR4
		GAD1	APOE, BDNF, GCR, GDNF, IL1B, IL2, IL6
		GAD2	GCR, IL2
		GABT	-
2	Arginine and proline metabolism: hsa00330	ALDH2	IL2
		ALDH1B1, ALDH3A2, ALDH7A1, ALDH9A1, CNDP1, CNDP2, GATM	-
3	Butanoate metabolism: hsa00650	GAD1	APOE, BDNF, GCR, GDNF, IL1B, IL2, IL6
		GAD2	GCR, IL2
		GABT	-

It is critical to highlight that gamma-aminobutyric acid possesses an inherent inability to cross the blood-brain barrier (BBB). Consequently, any correlation between GABA concentrations in blood plasma and cerebrospinal fluid may be tenuous. To account for the observed relationships between blood plasma and cerebrospinal fluid metabolites, we propose two potential mechanistic explanations. Firstly, we posit that the reconstructed gene regulatory networks may function similarly across various tissue types, including brain cells. Regulatory influences may thus drive congruent directional changes in metabolite concentration in cerebrospinal fluid and blood plasma, despite potential disparities in the absolute concentration values. We therefore hypothesize that POD genetic markers may induce changes in metabolite concentrations through mechanisms that are analogous in the brain and cerebrospinal fluid. Alternatively, the BBB in POD patients might have experienced localized disruption, thus fostering partial permeability to metabolites.

A noteworthy outcome of our analyses is the identification of serotonin as a secondary metabolomic marker, as outlined in Supplementary Table S5. Given that serotonin was not identified within the overrepresented metabolic pathways, we reconstructed the gene network through which POD genetic markers could potentially modulate serotonin levels. Towards this, we considered nine enzymes derived from two KEGG serotonin metabolism pathways (Fig. 8). The key constituents of this gene network are enumerated in Table 10.
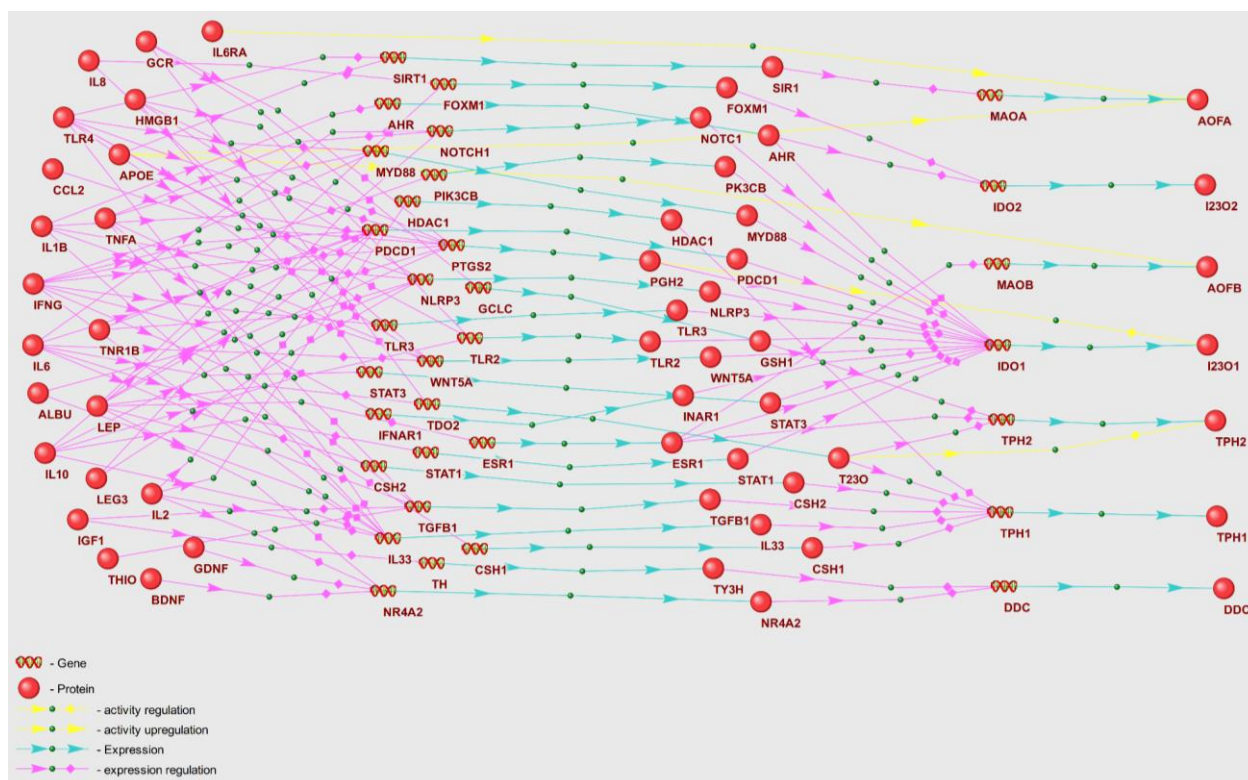
Fig. 8. Gene network of regulation of serotonin metabolism enzymes by POD genetic markers.

Table 10. Characteristics of gene regulatory network of the enzymes in serotonin metabolic pathways.

№	KEGG serotonin metabolism pathways	Enzymes in the metabolic pathway	Regulatory genetic markers
1	Serotonergic synapse: hsa04726	TPH1	ALBU, GCR, IGF1, IL1B, IL2, IL6, IL10, LEP, THIO, TLR4, TNFA
		DDC	BDNF, GDNF, IGF1, IL2, LEP
		TPH2	GCR, LEP
		AOFA	APOE, TLR4
		AOFB	APOE, LEP
2	Tryptophan metabolism: hsa00380	I23O1	ALBU, CCL2, GCR, HMGB1, IFNG, IL1B, IL2, IL6, IL8, IL10, LEG3, LEP, TLR4, TNFA, TNR1B
		DDC	BDNF, GDNF, IGF1, IL2, LEP
		I23O2	IL2, LEP
		AOFA	APOE, TLR4
		AOFB	APOE, LEP
		INMT, SNAT	-

In accordance with Table 10, I23O1 enzyme that is responsible for catalyzing oxidation of L-tryptophan receives the highest number of regulatory impacts from 15 POD genetic markers. A cumulative total of 21 POD genetic markers contribute to the regulation

of serotonin metabolism enzymes. Remarkably, leptin (LEP) exerts regulatory control over six enzymes involved in serotonin metabolism – AOFB, DDC, IDO1, IDO2, TPH1, and TPH2.

Discussion

The understanding of the molecular mechanisms underlying postoperative delirium remains incomplete, despite ongoing research efforts. To date, identification of the prognostic metabolomic markers of POD following major surgical operations remains elusive. This challenge is largely attributed to the complex systemic perturbations at the genetic and metabolomic levels entrenched in POD pathology, which hamper the delineation of discernible patterns using conventional statistical methodologies. In this study, we sought to identify the potential primary and secondary metabolomic biomarkers of POD from metabolomic data. We employed an integrative approach, involving statistical methodologies, the genetic algorithm, and artificial neural networks.

Acknowledging the complex pathophysiological processes inherent to POD, we defined two categories of metabolomic markers to provide a comprehensive description of changes in the metabolomic status of POD patients compared to non-POD patients. Primary markers, within the context of our study, are distinguished by a unidirectional alteration in concentration in blood plasma of patients with POD. Secondary metabolomic markers are typified by disruptions in inter-metabolite correlation relationships.

It is critical to underscore that many of the identified metabolites are incapable of crossing the blood-brain barrier (BBB) from blood plasma to directly influence the central nervous system. Furthermore, the concentrations of such metabolites in blood plasma and cerebrospinal fluid may not correlate. To interpret the results of blood plasma metabolomic analysis that allow us to hypothesize the mechanisms of disturbances in the brain, we postulated two assumptions. Our primary assumption is predicated on the expectation that the reconstructed gene regulatory networks function comparably across diverse tissues. Consequently, it can be anticipated that in various tissues POD genetic markers exert similar effects on metabolite concentration despite potential disparities in the values of metabolite concentrations across different tissues. An alternative hypothesis posits that inflammatory processes in POD could potentiate increased BBB permeability to various metabolites[58,59].

Metabolomic Markers of Postoperative Delirium

Our findings reveal that lactate and inositol are amongst the most prominent potential primary metabolomic markers for postoperative delirium (Supplementary Table S3). Through the use of gene networks, we elucidate the potential regulatory mechanisms regulating the metabolism of lactate and inositol, both of which are significant metabolomic markers of POD (Tables 7, 8).

An elevated concentration of lactate detected in patients' plasma prior to surgery may serve as a risk factor for POD. Lactate plays a crucial role in energy metabolism and can impact tissue perfusion[60]. The concentration of lactate in patients' plasma at the preoperative and early postoperative periods may function as a prognostic metabolomic marker of POD. Lactic acid, as well as glucose, is vital for providing energy to brain cells. During anaerobic conditions and hypoxia, lactic acid is involved in cellular metabolic pathways and serves as a substrate for tissue perfusion[61]. Lactate has been recognized as a marker of ischemia, hypoxia, and CNS damage[67,82]. The concentrations of glucose and lactate in blood plasma may reflect the stress response to surgical intervention. Decreased levels of glucose and lactate were observed in patients prior to cerebral aneurysm-related surgery[62]. In addition, S-methylcysteine emerged as one of the most important primary metabolomic markers (Supplementary Table S3). Guo et al. identified S-methylcysteine, eicosapentaenoic acid,

linolenic acid and linoleic acid as the metabolomic indicators associated with an elevated risk of POD[63].

Secondary metabolomic markers are characterized by a disrupted correlation with other metabolites, as opposed to a shift towards an increase or decrease in the POD group, as seen with primary markers. Consequently, not all secondary markers demonstrated statistically significant differences in concentration when comparing POD and non-POD groups. Our analysis intriguingly identified neurotransmitters as the POD potential markers. Specifically, serotonin was amongst the secondary metabolomic markers. It was reported that serotonin in the cerebrospinal fluid of the brain is a well-established marker of POD[64,65]. Plasma and CSF serotonin concentrations may not correlate due to the inability of the serotonin molecule to cross the blood-brain barrier. Nonetheless, we propose that disruptions in serotonin metabolism, which identified this neurotransmitter as a secondary marker, may be mediated by the influence of POD genetic markers sharing similar characteristics across various tissues, including brain tissues. Accordingly, the detection of other metabolites that are incapable of crossing the blood-brain barrier may be interpreted as potential markers of POD. Another example of a neurotransmitter marker is GABA, which was identified as both a primary and secondary metabolomic marker of POD.

Overrepresented Metabolic Pathways

Through the analysis of KEGG metabolic pathways, overrepresented within both the primary and secondary markers, a set of statistically significant metabolic pathways were discerned (Tables 3, 4). Notably, the analysis conducted exclusively on the primary marker set did not yield statistically significant results, likely due to the intricate nature of metabolic pathway dysfunction. This complexity encompasses not only directional shifts in metabolite concentrations but also the coordination in the variability of metabolite pathway participants. Consequently, the need for advancing bioinformatic methods for identifying secondary metabolomic markers is underscored.

The overrepresented metabolic pathways identified herein, are frequently related to amino acid metabolism. It is established that mechanisms such as amino acid metabolism pathway disorders, fatty acids, and the activation of alternative energy metabolic processes, are implicated in the pathogenesis of POD[66,84]. Guo et al., in their study of patients who underwent arthroplasty following a hip fracture, identified oxidative stress disorders, disruptions in energy metabolism and amino acid metabolism pathways in patients with POD[68]. According to the report, alterations in fatty acid and amino acid metabolism could take part in the pathophysiology of POD.

Tripp et al. conducted a computational analysis of blood plasma metabolomic profiles in elderly surgical patients. Based on metabolomic data, they conducted an overrepresentation analysis of KEGG metabolic pathways, which revealed alterations in the "Valine, leucine, and isoleucine biosynthesis" pathway in patients with POD on the day prior to surgery, and metabolites of the "Citrate cycle" pathway on the second day postoperatively[69]. Interestingly, we also identified the "Valine, leucine, and isoleucine biosynthesis" pathway as overrepresented within the primary and secondary metabolomic markers set.

Takahashi et al. constructed a protein-protein interaction network and identified two sets of human genes associated with POD development[70]. Their overrepresentation analysis revealed KEGG metabolic pathways enriched with POD genetic markers, including "neuroactive ligand-receptor interaction", cAMP, IL17, and TNF signaling pathways. Among the Gene Ontology terms enriched with POD genetic markers were "glutamate receptor activity", "neurotransmitter receptor activity", "G protein-coupled serotonin receptor

activity". In line with our findings, the metabolic pathways including neurotransmitter metabolism were among the overrepresented metabolic pathways.

Gene networks regulating metabolic pathways with POD genetic markers

Recently, several genetic markers for postoperative delirium, implicated extensively in the pathogenesis of POD, have been identified[71]. These markers are supposed to take part in initiating molecular pathways linked to the pathophysiology of delirium, including neuroinflammation, blood-brain barrier disruption, neurotransmitter imbalance, oxidative stress, neuroendocrine dysregulation, and neuronal connectivity disorganization[72]. By reconstructing and analyzing gene networks, we propose that these genetic markers, when dysregulated in the POD group, can instigate systemic disturbances in metabolic pathways. This is achieved through abnormal regulation of enzyme expression, activity, and stability (Figs. 2, 3). Among the analyzed POD genetic markers (Supplementary Table S6), TNFA is involved in regulation of the greatest number of enzymes within the overrepresented metabolic pathways (Fig. 4). Consequently, these genetic markers may serve as the key factors in metabolic pathway dysfunction identified via overrepresentation analysis.

As illustrated in the gene network graph (Fig. 3), interleukin-6 could potentially regulate SYAC and SYAM enzymes of the "Aminoacyl-tRNA biosynthesis" metabolic pathway, with the involvement of granzyme B and FTO protein. In the presence of IL6 and inflammation, *GZMB* gene expression may be decreased[73,74]. Granzyme B enzyme can cleave histidyl-, isoleucyl-, alanyl-tRNA synthetases[75]. Concurrently, via the STAT3 signaling pathway, IL6 is capable of enhancing *FTO* gene expression. Notably, FTO deficiency can result in diminished expression of *LRS* (leucyl-tRNA synthetase, SYLC) and other aminoacyl-tRNA synthetase family enzymes[76]. These enzymes catalyze the binding of amino acids to aminoacyl-tRNA, thereby directly influencing the primary metabolomic markers identified as alanine, leucine, and isoleucine.

Although our metabolomic analysis of blood plasma revealed these amino acids to be decreased in the POD group (Supplementary Table S1), this reduction aligns with the anticipated regulatory effect of *IL6* overexpression as suggested by the gene network. Preoperative elevation of IL6 was observed in POD patients[77]. Another regulatory pathway example is the IFNG→Tryptophanyl-tRNA synthetase pathway (Fig. 3). As per the gene network, IFNG directly enhances *WARS1* (Tryptophanyl-tRNA ligase 1) expression, and that upregulation may result in a decrease in tryptophan content[78]. Our data identified tryptophan as a primary metabolomic marker, and this metabolite was found to be decreased in the POD group. It is crucial to consider the whole set of regulatory effects on the target gene described by the gene network when assessing the impact of a specific regulatory pathway. The interference of multiple regulatory pathways, which together form the gene network, necessitates the use of mathematical modeling methods. Therefore, in the future studies, we are planning to interpret regulatory pathways at a qualitative level.

In an investigation of the regulatory pathways that potentially disrupt the function of enzymes in the "Glycine, serine, and threonine metabolism" pathway, IFNG→BHMT1, IL1B→BHMT1, TNFA→BHMT1 emerge as important contributors (Fig. 2). The gene network suggests that IFNG, IL1B, and TNFA can stimulate *RELA* gene expression, which encodes the TF65 transcription factor. TF65 in turn can inhibit the transcriptional activity of the *BHMT1* (Betaine-homocysteine S-methyltransferase 1) promoter. Of particular note, the BHMT1-mediated conversion involves secondary metabolomic markers L-homocysteine and L-methionine[79].

Investigating the potential influence of POD genetic markers on the deviant regulation of metabolomic markers that were not incorporated into the primary metabolic profile (PMP), we reconstructed gene networks for metabolic pathways housing the relevant enzymes. This

enabled us to postulate the principal regulators of these relevant enzymes within the set of POD genetic markers. As per our prior demonstration (Table 7), the regulation of inositol was primarily influenced by IL6, IL8, and LEP, while lactate regulation was mainly driven by APOE, IL1B, and TNFA.

For instance, the gene network in Fig. 5 indicates that IL6 regulates the expression of *CD274, FOSL, NNMT, STAT3,* and *VEGFA*, whose protein products subsequently regulate the expression of *LDHA* (L-lactate dehydrogenase A chain). LDHA mediates the reversible conversion of lactate to pyruvate. It is noteworthy that associations between *STAT3* and *VEGFA* elevated expression and POD development were identified in prior studies[57,83]. The enzyme NNMT catalyzes the transmethylation of nicotinamide. Intriguingly, the *FOSL* (Fos-related antigen 1) and *CD274* (Programmed cell death 1 ligand 1) genes are linked with apoptosis. It was established that the apoptotic processes in brain neurons constitute a part of the pathophysiological manifestations of postoperative delirium[84].

From the gene networks of GABA and serotonin regulation by genetic markers (Figs. 7, 8), it is evident that serotonin is subject to more regulatory influence than GABA. A unique attribute of the serotonin gene network is the multifaceted regulation of the I23O1 enzyme which catalyzes the oxidation reaction of L-tryptophan. This enzyme was reported to be associated with POD, due to its contribution to a decrease in L-tryptophan levels and a concurrent increase in kynurenine levels[85]. Notably, our analytical results identified L-tryptophan and kynurenine as the potential POD metabolomic markers (Supplementary Tables S3, S5).

In the analysis of pathologies such as postoperative delirium using blood plasma metabolomic data, a pertinent question regards the extent to which alterations in the plasma metabolomic profile can represent changes in the cerebrospinal fluid (CSF) and metabolomic status of the brain. Studies investigating correlations between changes in blood plasma and CSF metabolomic profiles are therefore of considerable relevance[21]. In this study, we hypothesize that certain metabolic pathways linked to POD development are disrupted in both blood plasma and CSF. This could particularly be evidenced by detection of elevated oxidative stress in cerebrospinal fluid, which can significantly impact the functioning of the nervous system. Preoperative low CSF saturation was previously suggested as a potential prognostic marker of POD[86]. In another study, based on CSF metabolomic analysis in patients undergoing arthroplasty, Pan et al. identified spermidine, putrescine, and glutamine as potential prognostic markers of POD[87]. The authors proposed the ratio of spermidine and putrescine concentrations to beta-amyloid 42 (Aβ42) as a potential marker of postoperative delirium.

Notably, changes in the content and correlation of metabolites in blood plasma can reveal disruptions in the pathways of amino acid biosynthesis, energy metabolism, and oxidative stress. These disturbances can, in turn, contribute to the pathophysiology of POD[69]. According to the reconstructed gene networks, disruptions in the functioning of metabolic pathways leading to alterations in metabolomic profiles in the POD group can primarily be associated with the modified functions of POD genetic markers.

The molecular mechanisms of postoperative delirium remain incompletely resolved, and the task of delineating reliable prognostic metabolomic markers post various surgical operations continues to present a formidable challenge – most predominantly in cardiac surgery. The complex nature of metabolomic data analysis in postoperative delirium studies can be attributed to the multifaceted systemic perturbations at the genetic and metabolomic levels inherent in this pathology, consequently impeding the attainment of robust results via conventional statistical methodologies.

In this study, we put forward several innovative computational strategies for discerning potential primary and secondary metabolic markers of POD, grounded in the

examination of blood plasma metabolomic data from patients. These strategies encompass the deployment of the genetic algorithm coupled with artificial neural networks. Employing these methodologies to analyze blood plasma metabolomic data from patients, harvested cardiac surgery, facilitated the identification of a constellation of primary and secondary potential metabolomic markers of POD. L-Lactic acid, Inositol, Methylcysteine, and Adenine emerged as the pivotal primary markers, indicative of the metabolic concentration shifts in the POD group. A set of 54 metabolites, inclusive of neurotransmitters such as GABA and serotonin, were categorized as the secondary markers, characterized by their disrupted correlations with other metabolites in the POD group.

Analysis of metabolic pathway overrepresentation, along with the reconstruction of gene regulatory networks, afforded us to postulate the interplay between perturbations in metabolic pathway functionality associated with a predisposition to POD and established genetic markers of POD, such as TNFA, LEP, IL6, and others. It is plausible to conjecture that functional disruptions in genetic POD markers could have an expansive influence on a spectrum of pathophysiological mechanisms, and that these genes have the potential to impact the metabolomic profile of patients. Collectively, the findings underscore the systemic essence of metabolic disturbances within patients exhibiting POD.

Further experimental validation of the proposed hypotheses concerning the role of POD genetic markers in the aberrant regulation of metabolic pathways could potentially emerge as a promising trajectory in the exploration of POD pathogenesis. The suggested approach of bioinformatic analysis of metabolomic data utilizing the genetic algorithm and artificial neural network methodologies holds perspective for the identification of primary and secondary potential markers across a wide amount of metabolomic studies.

Materials and Methods
Study Subjects
This study included patients aged 65 and above who underwent cardiac surgery involving cardiopulmonary bypass (CPB). Exclusion criteria included emergency intervention, aortic surgery, significant carotid artery stenosis, Parkinson's disease, liver cirrhosis (Child-Pugh Class B or C), and the consumption of anticholinergic drugs, antidepressants, antiepileptic medications, and chemotherapy drugs. Patient recruitment transpired between June 2019 and January 2021, resulting in inclusion of 39 patients in the study. Table 11 presents biometric data of the participants. Within five days post-surgery, the presence of postoperative delirium was ascertained using the CAM-ICU test (Confusion Assessment Method for the Intensive Care Unit). The initial test was conducted 6-8 hours post-surgery, followed by twice daily assessments. A positive CAM-ICU test at any time point was indicative of postoperative delirium.

Table 11. Age and gender characteristics of the patients.

Group	Sex (M/F)	Min. age	Max. age	Avg. age	Median	Standard Deviation
Non-POD	11/16	65	75	69.6	70	3.0
POD	5/7	65	79	69.7	69.5	4.3

Compliance with Ethical Standards
The Expert Committee granted approval for this study, according to protocol №10 from 16 April 2019. The Ethics Committee of the E.N. Meshalkin National Medical Research Center (Novosibirsk, Russia) granted approval for this study from 26 April 2019 with

protocol №24, ensuring adherence to ethical guidelines. The WHO International Clinical Trials Registry Platform (ICTRP) registered this study at ClinicalTrials.gov ID NCT03931499.

Blood Sampling and Preparation

Venous blood samples were collected from patients 24 hours after cardiac surgery, with a 9 ml BD Vacutainer® KEDTA tube, containing potassium EDTA as an anticoagulant. Plasma was obtained via centrifugation at 2000 g and 4°C for a duration of 15 minutes, then portioned and cryopreserved at -80°C until its deployment for analysis.

Adhering to the protocol proffered by Li et al.[80], all plasma samples were processed in aggregate. Each of 100 µl aliquots of blood plasma was mixed with 400 µl cooled methanol/acetonitrile solution (1:1 ratio), vigorously shaken, and subsequently centrifuged at 16,000 rpm and +4°C for 15 minutes. The supernatant was transferred into a glass vial insert for subsequent analysis. In parallel, two quality control samples were prepared by mixing equal volumes of blood plasma derived from the groups with and without POD.

LC-MS/MS Analysis

HPLC-MS/MS analysis was conducted using a Shimadzu LC-20AD Prominence chromatograph, which featured a binary gradient pump, a cooling autosampler SIL-20AC thermostated at 10 °C, and a CTO-10ASvp column oven maintained at 40 °C. A reversed-phase column octadecylsilyl sorbent ProntoSIL 120-5 C18 AQUA (LLC Econova, Novosibirsk, Russia), was employed for the LC separation. The mobile phase A consisted of water with 0.1% HCOOH (v/v), while mobile phase B was pure acetonitrile. The elution gradient proceeded as follows: from 0 to 2 minutes, the B content was 3%; from 2 to 7 minutes, it increased to 10%; from 7 to 10 minutes, it reached 90%; from 10 to 12 minutes, it was increased to 100% and was maintained until 16.5 minutes. After the run, the column was equilibrated for 3.4 minutes with 3% B. The flow rate was 200 µL/min, and a 2 µL sample injection volume was used.

For mass selective detection, a 6500 QTRAP mass spectrometer (AB SCIEX, USA) was utilized, equipped with an electrospray ionization source operating in both positive and negative modes. The mass spectrometer was set to operate in multiple reaction monitoring (MRM) mode for metabolite detection with the following parameters: the ion spray voltages (IS) were 5500 V for positive mode and -4500 V for negative mode. The gas dryer temperature (TEM) was set at 475°C, collision-induced dissociation gas (CAD) was configured as "Medium," and the sprayer gas (GS1), dryer gas (GS2), and curtain gas (CUR) were set to 35, 35, and 30 psi, respectively. The declustering potential (DP) was ±93V, the entrance potential (EP) was ±10V, and the collision cell exit potential (CXP) was ±20V for positive and negative ion modes, respectively. Additionally, the time for polarity switching (settling time) was established at 5 ms, and the dwell time was 3 ms for each MRM transition. The precursor ion and fragment ion transitions, metabolite names, dwell times, and the appropriate collision energies for both positive and negative ion modes were adapted from Yuan et al.[81], with several metabolite transitions added by our research group. The instrument was controlled and data were collected using Analyst 1.6.2 software (AB SCIEX), and chromatograms were processed using MultiQuant 2.1 software (AB SCIEX). The obtained peak area values were used for following statistical analysis.

Data Preprocessing

The raw data were preprocessed to fill in missing values for metabolite content within the analyzed samples. In instances where the number of samples with missing values did not exceed 5% of the total value number across 39 patients, the median value derived from

remaining samples was substituted as the metabolite content value. This approach is validated by robustness of the median to outliers. The ensuing values were further log-transformed.

Data Analysis
Statistical Analysis

The significance of difference between mean metabolite levels in POD and non-POD patients was computed using the Mann-Whitney test, implemented in the SciPy package v1.8.0 (https://scipy.org/). The multiple hypothesis testing was enacted using the Benjamini-Yekutieli procedure from the statsmodels Python package v0.13.231 (https://www.statsmodels.org/stable/index.html).

Genetic Algorithm

To identify the combinations of metabolites that differ most effectively between POD and non-POD patient groups, we employed the genetic algorithm, as described by Lu et al., 2017[54]. This algorithm was executed through the application of a Python module (https://pygad.readthedocs.io). The fitness function was formulated in the following manner: (1) an XGBoost classification model (https://xgboost.readthedocs.io/) was constructed utilizing the input list of metabolites, (2) the precision of this model was evaluated by a 5-fold cross-validation methodology, with selection of the minimum accuracy value. The function returned the calculated model precision. The XGBoost model was trained with the following parameters: n_estimators = 1000, learning_rate = 0.05, and other parameters set to their default values.

Denoising Autoencoder (DAE)

The DAE architecture comprised three fully interconnected layers (input, hidden, and output), each utilizing the Rectified Linear Unit (ReLU) activation function. The input and output layers contained 210 neurons, which corresponds to the number of metabolites analyzed. The hidden layer, depending upon the specific list, contained from 50 to 149 neurons. The optimizer applied was Adam, with a learning rate of 0.001. The model was trained over 20 epochs, with a batch size of 8 and Mean Squared Error (MSE) as the loss metric. Data was split into training (80%) and testing (20%) sets. The DAE model was implemented using the PyTorch package (https://pytorch.org/). During the training phase, the DAE input contained metabolomic profile with noise added to the metabolite concentration values. The DAE aimed to encode the input data and predict the original data prior to addition of noise.

Generation of Noisy Metabolomic Profile

A noisy metabolomic profile was generated by introducing a random number, derived from a normal distribution to each metabolite concentration value. The formula applied was as follows:

$$CN_n = CO_n + d \cdot e,$$

where CN_n denotes the noisy concentration value of the n-th metabolite, CO_n represents the original concentration value of the n-th metabolite, e is a member of a normal distribution with the parameters N (Ex = 0, $\sigma = CO_n$), and d is the noise level factor.

The coefficient d was assigned a value of 0.25 for distinct DAE models. This procedure permitted the generation of 500 noisy profiles from a single original metabolomic profile.

Identification of the Secondary Metabolomic Markers

To explore secondary metabolomic markers, we adopted a technique for anomaly detection by employing a denoising autoencoder, as delineated by Sakurada et al., 2014[49]. The algorithm comprised the following steps: initially, the denoising autoencoder was trained using a set of metabolomic profiles from patients without Postoperative Delirium (non-POD group). Subsequently, the trained model was utilized to examine the metabolomic profiles pertaining to the POD patients group. Metabolites manifesting the anomalous concentration values as per the denoising autoencoder were classified as the secondary metabolic markers.

To evaluate the anomaly degree of each metabolite in the POD patient group, we contrasted the distribution of denoising autoencoder errors when analyzing the POD and non-POD groups. The error was gauged as the absolute value of the discrepancy between the metabolite concentration values supplied to the denoising autoencoder input and those procured at the output. Crucially, the original metabolomic profiles, devoid of noise, were supplied to the denoising autoencoder input. Disparities between error distributions were assessed via the Mann-Whitney test. Data was deemed to contain anomalies if the two distributions exhibited statistically significant differences. Multiple hypothesis testing was enacted using the False Discovery Rate (FDR) Benjamini-Hochberg procedure.

The architecture of the denoising autoencoder involves the hidden layer with a lower dimensionality than either the input or output layers. It can be hypothesized that the manifestation of anomalies in the metabolite concentration data, as identified via the denoising autoencoder, is associated with a disruption of relationships (correlation links) between metabolites in POD patients. Such disruptions could be linked to dysfunctions in metabolic processes, engendering a predisposition to POD.

MetaboAnalyst

The overrepresentation analysis of KEGG metabolic pathways was evaluated using the enrichment analysis instrument implemented in the MetaboAnalyst 5.0 web tool[38] (https://www.metaboanalyst.ca/).

Reconstruction of Gene Networks

Gene networks reconstruction was carried out using ANDVisio program, which serves as a graphical user interface for the ANDSystem cognitive system. Using the "Pathway Wizard" module of the ANDVisio program, we reconstructed molecular genetic pathways of regulation of identified metabolic pathways enzymes by POD genetic markers. Regulatory pathways were built according to five Template types presented in Table 12. These Templates include various combinations of molecular genetic interactions, such as regulation of gene expression, protein-protein interactions, regulation of protein activity, degradation, catalysis, and transport.

Table 12. Templates for molecular genetic pathways of metabolic pathway enzymes regulation by POD genetic markers.

№	Template name	Regulatory pathway
P1	Protein-protein interactions	**PODp** –protein-protein interaction→ **EnzP**
P2	Protein function regulation	**PODp** –regulation of activity/degradation/catalyze/transport→ **EnzP**
P3	Expression regulation	**PODp** –expression regulation→ **EnzG** –expression→ **EnzP**
P4	Double expression regulation	**PODp** –expression regulation→ **Hg** –expression→ **Hp** –expression regulation→ **EnzG** –expression→ **EnzP**

P5	Combined regulatory pathway	**PODp** –expression regulation→ **Hg** –expression→ **Hp** – regulation of activity/degradation/catalyze/transport→ **EnzP**

Abbreviations:
PODp – POD genetic markers proteins; Hg – human genes; Hp – human proteins; EnzG – genes encoding metabolic pathway enzymes; EnzP – metabolic pathway enzymes.

Data availability

The data is provided with this paper in supplementary information. The source code of AIMetabolicMarker is available at https://github.com/LCP-ICG/AIMetabolicMarker.

References

1. Gottesman, R. F. et al. Delirium after coronary artery bypass graft surgery and late mortality. *Ann. Neurol.* **67**, 338–344 (2010).
2. Pisani, M. A. *et al.* Days of Delirium Are Associated with 1-Year Mortality in an Older Intensive Care Unit Population. *Am. J. Respir. Crit. Care Med.* **180**, 1092–1097 (2009).
3. Rudolph, J. L. *et al.* Derivation and Validation of a Preoperative Prediction Rule for Delirium After Cardiac Surgery. *Circulation* **119**, 229–236 (2009).
4. Feinkohl, I. *et al.* Metabolic syndrome and the risk of postoperative delirium and postoperative cognitive dysfunction: a multi-centre cohort study. *Br. J. Anaesth.* **131**, 338–347 (2023).
5. Bellelli, G., Brathwaite, J. S. & Mazzola, P. Delirium: A Marker of Vulnerability in Older People. *Front. Aging Neurosci.* **13**, 626127 (2021).
6. Bartikoski, B. J. *et al.* A Review of Metabolomic Profiling in Rheumatoid Arthritis: Bringing New Insights in Disease Pathogenesis, Treatment and Comorbidities. *Metabolites* **12**, 394 (2022).
7. Wu, X., Wang, Z., Luo, L., Shu, D. & Wang, K. Metabolomics in hepatocellular carcinoma: From biomarker discovery to precision medicine. *Front. Med. Technol.* **4**, 1065506 (2023).
8. Gonzalez-Covarrubias, V., Martínez-Martínez, E. & del Bosque-Plata, L. The Potential of Metabolomics in Biomedical Applications. *Metabolites* **12**, 194 (2022).
9. Jin, Y., Chi, J., LoMonaco, K., Boon, A. & Gu, H. Recent review on selected xenobiotics and their impacts on gut microbiome and metabolome. *TrAC Trends in Analytical Chemistry* **166**, 117155 (2023).
10. Quintás, G., Castell, J. V. & Moreno-Torres, M. The assessment of the potential hepatotoxicity of new drugs by in vitro metabolomics. *Front. Pharmacol.* **14**, 1155271 (2023).
11. Sherlock, L. & Mok, K. H. Metabolomics and Its Applications to Personalized Medicine. *EKC 2019 Conference Proceedings*, 25–42 (2021). https://doi.org/10.1007/978-981-15-8350-6_3.
12. Jansen, J. J. *et al.* Between Metabolite Relationships: an essential aspect of metabolic change. *Metabolomics* **8**, 422–432 (2012).
13. Li, C., Gao, Z., Su, B., Xu, G. & Lin, X. Data analysis methods for defining biomarkers from omics data. *Anal. Bioanal. Chem.* **414**, 235–250 (2022).
14. Auyeung, A., Wang, H. C., Aravagiri, K. & Knezevic, N. N. Kynurenine Pathway Metabolites as Potential Biomarkers in Chronic Pain. *Pharmaceuticals* **16**, 681 (2023).
15. Lanpher, B., Brunetti-Pierri, N. & Lee, B. Inborn errors of metabolism: the flux from Mendelian to complex diseases. *Nat. Rev. Genet.* **7**, 449–459 (2006).
16. Ombrone, D., Giocaliere, E., Forni, G., Malvagia, S. & la Marca, G. Expanded newborn screening by mass spectrometry: New tests, future perspectives. *Mass Spectrom. Rev.* **35**, 71–84 (2016).

17. Younesi, S. *et al.* Incorporation of second-tier tests and secondary biomarkers to improve positive predictive value (PPV) rate in newborn metabolic screening program. *J. Clin. Lab. Anal.* **36**, e24471 (2022).

18. Cawley, A., Keen, B., Tou, K., Elbourne, M. & Keledjian, J. Biomarker ratios. *Drug Test . Anal.* **14**, 983–990 (2022).

19. Xiao, X. *et al.* Exploring metabolic pathway disruption in the subchronic phencyclidine model of schizophrenia with the Generalized Singular Value Decomposition. *BMC Syst. Biol.* **5**, 72 (2011).

20. Lima, A. R. *et al.* Comprehensive Metabolomics and Lipidomics Profiling of Prostate Cancer Tissue Reveals Metabolic Dysregulations Associated with Disease Development. *J. Proteome Res.* **21**, 727–739 (2022).

21. Rogachev, A. D. *et al.* Correlation of Metabolic Profiles of Plasma and Cerebrospinal Fluid of High-Grade Glioma Patients. *Metabolites* **11**, 133 (2021).

22. Camacho, D., de la Fuente, A. & Mendes, P. The origin of correlations in metabolomics data. *Metabolomics* **1**, 53–63 (2005).

23. Min Lee, J., Gianchandani, E. P., Eddy, J. A. & Papin, J. A. Dynamic Analysis of Integrated Signaling, Metabolic, and Regulatory Networks. *PLoS Comput. Biol.* **4**, e1000086 (2008).

24. Rosato, A. *et al.* From correlation to causation: analysis of metabolomics data using systems biology approaches. *Metabolomics* **14**, 37 (2018).

25. Tsimring, L. S. Noise in biology. *Rep. Prog. Phys.* **77**, 026601 (2014).

26. Kiviet, D. J. *et al.* Stochasticity of metabolism and growth at the single-cell level. *Nature* **514**, 376–379 (2014).

27. Borri, A., Palumbo, P. & Singh, A. Noise propagation in metabolic pathways: the role of growth-mediated feedback. *bioRxiv* 2020.03.21.001495, https://doi.org/10.1101/2020.03.21.001495.

28. Borri, A., Palumbo, P. & Singh, A. A general framework for noise propagation in a cascade of metabolic transformations. *IFAC-PapersOnLine* **55**, 121–126 (2022).

29. Gomes, B. & Ashley, E. A. Artificial Intelligence in Molecular Medicine. *New Engl. J. Med.* **388**, 2456–2465 (2023).

30. Kramer, M. A. Nonlinear principal component analysis using autoassociative neural networks. *AIChE J.* **37**, 233–243 (1991).

31. Pawar, K. & Attar, V. Z. Assessment of Autoencoder Architectures for Data Representation. *Springer*, 101–132 (2020). https://doi.org/10.1007/978-3-030-31756-0_4.

32. Gomari, D. P. *et al.* Variational autoencoders learn transferrable representations of metabolomics data. *Commun Biol* **5**, 645 (2022).

33. Chardin, D., Gille, C., Pourcher, T., Humbert, O. & Barlaud, M. Learning a confidence score and the latent space of a new supervised autoencoder for diagnosis and prognosis in clinical metabolomic studies. *BMC Bioinformatics* **23**, 361 (2022).

34. Ca, P. V., Edu, L. T., Lajoie, I., Ca, Y. B. & Ca, P.-A. M. Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion Pascal Vincent Hugo Larochelle Yoshua Bengio Pierre-Antoine Manzagol. *J. Mach. Learn. Res.* **11**, 3371-3408 (2010).

35. Ji, H. & Tian, J. Deep denoising autoencoder-assisted continuous scoring of peak quality in high-resolution LC−MS data. Chemometr. Intell. Lab. Syst. 231, 104694 (2022).

36. Zhang, A. Quality Control and Applications of Mass Spectrometry-based Metabolomics: From Cell Research to Large-scale Human Cohort Studies. *UC Davis* (2021). https://escholarship.org/uc/item/9f34h8n2

37. Rong, Z. *et al.* NormAE: Deep Adversarial Learning Model to Remove Batch Effects in Liquid Chromatography Mass Spectrometry-Based Metabolomics Data. *Anal. Chem.* **92**, 5082–5090 (2020).

38. Pang, Z. *et al.* MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res.* **49(W1)**, W388–W396 (2021).

39. Ivanisenko, V. A. *et al.* ANDSystem: an Associative Network Discovery System for automated literature mining in the field of biology. *BMC Syst. Biol.* **9**, S2 (2015).

40. Ivanisenko, V. A., Demenkov, P. S., Ivanisenko, T. V., Mishchenko, E. L. & Saik, O. V. A new version of the ANDSystem tool for automatic extraction of knowledge from scientific publications with expanded functionality for reconstruction of associative gene networks by considering tissue-specific gene expression. *BMC Bioinformatics* **20**, 34 (2019).

41. Ivanisenko, T. V. *et al.* ANDDigest: a new web-based module of ANDSystem for the search of knowledge in the scientific literature. *BMC Bioinformatics* **21**, 228 (2020).

42. Ivanisenko, T. V., Demenkov, P. S., Kolchanov, N. A. & Ivanisenko, V. A. The New Version of the ANDDigest Tool with Improved AI-Based Short Names Recognition. *Int. J. Mol. Sci.* **23**, 14934 (2022).

43. Saik, O. V., Ivanisenko, T. V., Demenkov, P. S. & Ivanisenko, V. A. Interactome of the hepatitis C virus: Literature mining with ANDSystem. *Virus Res.* **218**, 40–48 (2016).

44. Antropova, E. A. *et al.* Computer analysis of regulation of hepatocarcinoma marker genes hypermethylated by HCV proteins. *Vavilovskii Zhurnal Genet. Selektsii* **26**, 733–742 (2022).

45. Saik, O. V. *et al.* Prioritization of genes involved in endothelial cell apoptosis by their implication in lymphedema using an analysis of associative gene networks with ANDSystem. *BMC Med. Genomics* **12**, 47 (2019).

46. Yankina, M. A., Saik, O. V., Ivanisenko, V. A., Demenkov, P. S. & Khusnutdinova, E. K. Evaluation of Prioritization Methods of Extrinsic Apoptotic Signaling Pathway Genes for Retrieval of the New Candidates Associated with Major Depressive Disorder. *Russ. J. Genet.* **54**, 1366–1374 (2018).

47. Saik, O. V. *et al.* Novel candidate genes important for asthma and hypertension comorbidity revealed from associative gene networks. *BMC Med .Genomics* **11**, 15 (2018).

48. Ivanisenko, V. A. *et al.* Plasma metabolomics and gene regulatory networks analysis reveal the role of nonstructural SARS-CoV-2 viral proteins in metabolic dysregulation in COVID-19 patients. *Sci. Rep.* **12**, 19977 (2022).

49. Sakurada, M. & Yairi, T. Anomaly Detection Using Autoencoders with Nonlinear Dimensionality Reduction. In *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis (MLSDA'14). Association for Computing Machinery, New York, NY, USA*, 4–11 (2014), https://doi.org/10.1145/2689746.2689747.

50. An, J. & Cho, S. Variational Autoencoder based Anomaly Detection using Reconstruction Probability. (2015) via https://api.semanticscholar.org/CorpusID:36663713.

51. Zhou, C. & Paffenroth, R. C. Anomaly Detection with Robust Deep Autoencoders. in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 665–674 (ACM, 2017). https://doi.org/10.1145/3097983.3098052.

52. Ribeiro, M., Lazzaretti, A. E. & Lopes, H. S. A study of deep convolutional auto-encoders for anomaly detection in videos. *Pattern Recognit. Lett.* **105**, 13–22 (2018).

53. Morales-Forero, A. & Bassetto, S. Case Study: A Semi-Supervised Methodology for Anomaly Detection and Diagnosis. in *2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)* 1031–1037 (IEEE, 2019). https://doi.org/10.1109/IEEM44572.2019.8978509.

54. Lu, H. *et al.* A hybrid feature selection algorithm for gene expression data classification. *Neurocomputing* **256**, 56–62 (2017).

55. Kazmierski, J., Banys, A., Latek, J., Bourke, J. & Jaszewski, R. Raised IL-2 and TNF-α concentrations are associated with postoperative delirium in patients undergoing coronary-artery bypass graft surgery. *Int. Psychogeriatr.* **26**, 845–855 (2014).

56. Huang, H. *et al.* Early Serum Metabolism Profile of Post-operative Delirium in Elderly Patients Following Cardiac Surgery With Cardiopulmonary Bypass. *Front. Aging Neurosci.* **14**, 857902 (2022).

57. Vasunilashorn, S. M. *et al.* Cytokines and Postoperative Delirium in Older Patients Undergoing Major Elective Surgery. *J. Gerontol. A Biol. Sci. Med. Sci.* **70**, 1289–1295 (2015).

58. Taylor, J. *et al.* Postoperative delirium and changes in the blood–brain barrier, neuroinflammation, and cerebrospinal fluid lactate: a prospective cohort study. *Br. J. Anaesth.* **129**, 219–230 (2022).

59. Terrando, N. & Akassoglou, K. Breaking barriers in postoperative delirium. *Br. J. Anaesth.* **129**, 147–150 (2022).

60. Wang, J., Li, M., Wang, P. & Fang, P. Goal-directed therapy based on rScO2 monitoring in elderly patients with one-lung ventilation: a randomized trial on perioperative inflammation and postoperative delirium. *Trials* **23**, 687 (2022).

61. Smith, D. *et al.* Lactate: A Preferred Fuel for Human Brain Metabolism *in Vivo*. *J. Cereb. Blood Flow Metab.* **23**, 658–664 (2003).

62. Tang, C. *et al.* Neuroprotective effect of bispectral index-guided fast-track anesthesia using sevoflurane combined with dexmedetomidine for intracranial aneurysm embolization. *Neural Regen. Res.* **13**, 280 (2018).

63. Guo, Y. *et al.* Preoperative Serum Metabolites Are Associated With Postoperative Delirium in Elderly Hip-Fracture Patients. *J. Gerontol. A Biol. Sci. Med. Sci.* **72**, 1689–1696 (2017).

64. Watne, L. O. *et al.* Increased CSF levels of aromatic amino acids in hip fracture patients with delirium suggests higher monoaminergic activity. *BMC Geriatr.* **16**, 149 (2016).

65. Hall, R. J. *et al.* CSF biomarkers in delirium: a systematic review. *Int. J. Geriatr. Psychiatry.* **33**, 1479–1500 (2017).

66. Shahabi Raberi, V. *et al.* Postoperative Delirium and Dementia in Patients Undergoing Cardiac Surgery: A Review of Randomized Controlled Trials. *Galen Medical Journal* **12**, e3045 (2023).

67. Dienel, G. A. Lactate shuttling and lactate use as fuel after traumatic brain injury: metabolic considerations. *J. Cereb. Blood Flow. Metab.* **34**, 1736–1748 (2014).

68. Guo, R.-J., Xue, F.-S., Shao, L.-J.-Z. & Zheng, L. Establishing a predictive nomogram for postoperative delirium: Methodology is critical. *Injury* **50**, 618 (2019).

69. Tripp, B. A. *et al.* Targeted metabolomics analysis of postoperative delirium. *Sci. Rep.* **11**, 1521 (2021).

70. Takahashi, Y., Terada, T. & Muto, Y. Systems Level Analysis and Identification of Pathways and Key Genes Associated with Delirium. *Genes (Basel)* **11**, 1225 (2020).

71. Androsova, G., Krause, R., Winterer, G. & Schneider, R. Biomarkers of postoperative delirium and cognitive dysfunction. *Front. Aging Neurosci.* **7**, 112 (2015).

72. Maldonado, J. R. Neuropathogenesis of Delirium: Review of Current Etiologic Theories and Common Pathways. *Am. J. Geriatr. Psychiatry* **21**, 1190–1222 (2013).

73. Gulati, P. *et al.* Role for the obesity-related FTO gene in the cellular sensing of amino acids. *Proc. Natl Acad. Sci. USA* **110**, 2557–2562 (2013).

74. Cifaldi, L. *et al.* Inhibition of Natural Killer Cell Cytotoxicity by Interleukin-6: Implications for the Pathogenesis of Macrophage Activation Syndrome. *Arthritis Rheumatol.* **67**, 3037–3046 (2015).

75. Casciola-Rosen, L., Andrade, F., Ulanet, D., Wong, W. B. & Rosen, A. Cleavage by Granzyme B Is Strongly Predictive of Autoantigen Status. *J. Exp. Med.* **190**, 815–826 (1999).

76. Gulati, P. & Yeo, G. S. H. The biology of FTO: from nucleic acid demethylase to amino acid sensor. *Diabetologia* **56**, 2113–2121 (2013).

77. Capri, M. *et al.* Pre-Operative, High-IL-6 Blood Level is a Risk Factor of Post-Operative Delirium Onset in Old Patients. *Front. Endocrinol. (Lausanne)* **5**, 176 (2014).

78. Adam, I. *et al.* Upregulation of tryptophanyl-tRNA synthethase adapts human cancer cells to nutritional stress caused by tryptophan degradation. *Oncoimmunology* **7**, e1486353 (2018).

79. Breksa, A. P. & Garrow, T. A. Recombinant Human Liver Betaine-homocysteine *S*-Methyltransferase: Identification of Three Cysteine Residues Critical for Zinc Binding. *Biochemistry* **38**, 13991–13998 (1999).

80. Li, K., Naviaux, J.C., Bright, A.T. *et al*. A robust, single-injection method for targeted, broad-spectrum plasma metabolomics. *Metabolomics* **13**, 122 (2017).

81. Yuan, M., Breitkopf, S., Yang, X. *et al*. A positive/negative ion–switching, targeted mass spectrometry–based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. *Nat. Protoc.* **7**, 872–881 (2012).

82. Bergersen, L. H. Lactate transport and signaling in the brain: potential therapeutic targets and roles in body–brain interaction. *J. Cereb. Blood Flow Metab*. **35**, 176–185 (2015)

83. Jian, R. & He, X. TIPE2 knockdown exacerbates isoflurane-induced postoperative cognitive impairment in mice by inducing activation of STAT3 and NF-κB signaling pathways. *Transl. Neurosci.* **14**, 20220282 (2023).

84. Xu, Y., Ma, Q., Du, H., Yang, C. & Lin, G. Postoperative Delirium in Neurosurgical Patients: Recent Insights into the Pathogenesis. *Brain Sci.* **12**, 1371 (2022).

85. Phing, A. H. *et al.* Altered Tryptophan-Kynurenine Pathway in Delirium: A Review of the Current Literature. *Int. J. Mol. Sci.* **24**, 5580 (2023).

86. Schoen, J. *et al.* Preoperative regional cerebral oxygen saturation is a predictor of postoperative delirium in on-pump cardiac surgery patients: a prospective observational trial. *Crit. Care* **15**, R218 (2011).

87. Pan, X. *et al.* Cerebrospinal Fluid Spermidine, Glutamine and Putrescine Predict Postoperative Delirium Following Elective Orthopaedic Surgery. *Sci. Rep.* **9**, 4191 (2019).

Authors and Affiliations

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
Vladimir A. Ivanisenko, Aelita-Luiza A. Makarova, Irina N. Kuzmicheva, Pavel S. Demenkov, Artur S. Venzel, Timofey V. Ivanisenko, Evgenia A. Antropova, Nikolay A. Kolchanov, Alexey V. Kochetov

Novosibirsk State University, Novosibirsk, Russia
Vladimir A. Ivanisenko, Artem D. Rogachev, Nikita V. Basov, Evgeniy V. Gaisler, Irina N. Kuzmicheva, Pavel S. Demenkov, Artur S. Venzel, Timofey V. Ivanisenko, Margarita A. Naidanova, Nikolay A. Kolchanov, Alexey V. Kochetov, Andrey G. Pokrovsky

Kurchatov Genomic Center of ICG SB RAS, Novosibirsk, Russia
Vladimir A. Ivanisenko, Pavel S. Demenkov, Artur S. Venzel, Timofey V. Ivanisenko

N. N. Vorozhtsov Institute of Organic Chemistry, Novosibirsk, Russia
Artem D. Rogachev, Nikita V. Basov, Evgeniy V. Gaisler

E. Meshalkin National Medical Research Center, Novosibirsk, Russia; Department of Anesthesiology and Intensive Care
Vladimir V. Lomivorotov, Gleb B. Moroz, Victoria V. Plesko

Penn State Milton S. Hershey Medical Center, Hershey, PA, United States; Department of Anesthesiology and Perioperative Medicine
Vladimir V. Lomivorotov

Author contributions
V.V.L., V.A.I., A.G.P., N.A.K., A.V.K. designed the experiment and directed the study. V.V.L., A.D.R., N.V.B., V.V.P., G.B.M. performed the experiments. V.A.I., I.N.K., P.S.D., A.S.V., T.V.I. developed a bioinformatics method. V.A.I., A.-L.A.M., E.V.G., I.N.K., P.S.D., A.S.V., T.V.I., E.A.A., M.A.N. analyzed data. V.A.I., A.-L.A.M., A.D.R., P.S.D., A.S.V., T.V.I., V.V.L. wrote the manuscript.

Competing interests
The authors declare no competing interests.

Additional information

Corresponding Author
Correspondence to Vladimir V. Lomivorotov v.lomivorotov@gmail.com

Tables

Table 1. Analysis of overrepresented KEGG metabolic pathways associated with the set of primary metabolomic markers.

KEGG pathway	Total	Metabolomic Markers (Hits)	p-value	FDR
Pyrimidine metabolism	39	3	0.006	0.256
Tryptophan metabolism	41	3	0.006	0.256
Purine metabolism	70	3	0.028	0.616

Table 2. Overrepresented KEGG metabolic pathways for the set of secondary markers identified by 100 autoencoder models.

KEGG Metabolic pathway	Total	Hits	P-value	FDR
beta-Alanine metabolism	21	2	0.006	0.403
Glycine, serine and threonine metabolism	33	3	0.015	0.403

Table 3. Overrepresented KEGG metabolic pathways for the set of the secondary markers identified by 97 autoencoder models.

KEGG Metabolic pathway	Total	Hits	P-value	FDR
Glycine, serine and threonine metabolism	33	4	2.03E-4	0.0163
beta-Alanine metabolism	21	3	8.99E-4	0.036

Table 4. Metabolites serving as both the primary and secondary metabolomic markers.

KEGG Id	Metabolite	Primary marker importance
C00137	Inositol	87
C00120	Biotin	3
C01089	3-hydroxybutyric acid	1
C00318	Carnitine	1
C00334	Gamma-Aminobutyric acid	1
C00258	Glyceric acid	1
C00106	Uracil	1
C00438	Ureidosuccinic acid	1

Table 5. Overrepresented KEGG metabolic pathways for the combined set of primary and secondary metabolomic markers. Metabolites that are primary and secondary markers are marked in bold.

KEGG pathway	Total	FDR	Primary Metabolomic Markers (Hits)	Secondary Metabolomic Markers (Hits)
Glycine, serine and threonine metabolism	33	0.0021	Betaine; **Glyceric acid**	L-Serine; Choline; Betaine aldehyde; Guanidoacetic acid; Cystathionine; L-Threonine; **Glyceric acid**
Aminoacyl-tRNA biosynthesis	48	0.0028	L-Leucine; L-Tryptophan; L-Tyrosine	L-Histidine; L-Arginine; L-Serine; L-Valine; L-Threonine; L-Proline
Tryptophan metabolism	41	0.0239	L-Tryptophan; 5-Hydroxy-L-tryptophan	Serotonin; 3-Hydroxyanthranilic acid; L-Kynurenine; Acetyl-CoA
Arginine biosynthesis	14	0.0351	N-Acetylglutamic acid	L-Arginine; Citrulline; Ornithine

Butanoate metabolism	15	0.0372	**3-hydroxybutyric acid; Gamma-Aminobutyric acid**; D-2-Hydroxyglutaric acid	**3-hydroxybutyric acid**; Acetyl-CoA; **Gamma-Aminobutyric acid**
Valine, leucine and isoleucine biosynthesis	8	0.0375	L-Leucine	L-Threonine; L-Valine
Arginine and proline metabolism	38	0.0375	**Gamma-Aminobutyric acid**	L-Arginine; Guanidoacetic acid; **Gamma-Aminobutyric acid**; Spermine; Ornithine; L-Proline
Pyrimidine metabolism	39	0.0376	Cytidine; Thymine; **Ureidosuccinic acid**; **Uracil**	Uridine; **Ureidosuccinic acid**; **Uracil**; Deoxyribose 1-phosphate
Ubiquinone and other terpenoid-quinone biosynthesis	9	0.0394	L-Tyrosine; Homogentisic acid	Phenyllactic acid

Table 6. Quantitative indicators of POD genetic markers (N1) and their corresponding regulated metabolic pathway enzymes (N2) as per regulatory pathway templates.

KEGG pathway	Regulatory pathways									
	P1		P2		P3		P4		P5	
	N1	N2	N1	N2	N1	N2	N1	N2	N1	N2
Glycine, serine, and threonine metabolism	2	3	3	5	3	3	15	13	5	3
Aminoacyl-tRNA biosynthesis	1	1	0	0	1	2	10	3	5	2

Table 7. Descriptive features of gene network of regulation of lactate metabolic pathway enzymes.

№	KEGG lactate metabolism pathways	Enzymes in the metabolic pathway	Regulatory genetic markers
1	Glycolysis / Gluconeogenesis: hsa00010;	LDHA	APOE, CCL2, GDNF, HMGB1, IGF1, IFNG, IL1B, IL2, IL6, IL6RA, IL8, IL10, LEP, TLR4, TNFA
		LDHB	IL1B, S100B, TNFA
2	HIF-1 signaling pathway: hsa04066		

		LDHC, LDHAL6A, LDHAL6B	-
3	Pyruvate metabolism: hsa00620	GLO2	APOE
		LDHA	APOE, CCL2, GDNF, HMGB1, IGF1, IFNG, IL1B, IL2, IL6, IL6RA, IL8, IL10, LEP, TLR4, TNFA
		LDHB	IL1B, S100B, TNFA
		GLUL, GRHPR, HAGH, LDHC, LDHD, LDHAL6A, LDHAL6B	-

Table 8. Characteristics of gene regulatory network of inositol metabolic pathway enzymes.

№	KEGG inositol metabolism pathways	Enzymes in the metabolic pathway	Regulatory genetic markers
1	Inositol phosphate metabolism: hsa00562	IMPA1	S100B
		IMPA2	THIO
		INP4B	CCL2, IGF1, IL6, IL6RA, IL8, LEP
		INP5K	BDNF
		ITPR3	HMGB1, IL6, IL8, LEP, MMP9, TLR4
		CDIPT, INP1, INP4A, ISYNA1, MIOX, MTM1	-
2	Phosphatidylinositol signaling system: hsa04070	IMPA1	S100B
		IMPA2	THIO
		INP4B	CCL2, IGF1, IL6, IL6RA, IL8, LEP
		PLCB2, PLCD1	IL6, IL8, LEP
		PLCG1	PECA1
		PLCG2	TAU
		BPNT2, CDIPT, INP1, INP4A, INP5A, MTM1, PLCs	-
3	Ascorbate and aldarate metabolism: hsa00053	MIOX	-

Table 9. Characteristics of gene regulatory network of GABA metabolic pathway enzymes.

№	KEGG GABA metabolism pathways	Enzymes in the metabolic pathway	Regulatory genetic markers
1	GABAergic synapse: hsa04727	GLSK	GDNF, GCR, HMGB1, IGF1, IFNG, IL1B, IL1RA, IL2, IL6,

			IL8, IL10, LEG3, LEP, TLR4
		GLSL	GCR, IGF1, IL1RA, IL6, IL8, IL10, LEG3, LEP, TLR4
		GAD1	APOE, BDNF, GCR, GDNF, IL1B, IL2, IL6
		GAD2	GCR, IL2
		GABT	-
2	Arginine and proline metabolism: hsa00330	ALDH2	IL2
		ALDH1B1, ALDH3A2, ALDH7A1, ALDH9A1, CNDP1, CNDP2, GATM	-
3	Butanoate metabolism: hsa00650	GAD1	APOE, BDNF, GCR, GDNF, IL1B, IL2, IL6
		GAD2	GCR, IL2
		GABT	-

Table 10. Characteristics of gene regulatory network of the enzymes in serotonin metabolic pathways.

№	KEGG serotonin metabolism pathways	Enzymes in the metabolic pathway	Regulatory genetic markers
1	Serotonergic synapse: hsa04726	TPH1	ALBU, GCR, IGF1, IL1B, IL2, IL6, IL10, LEP, THIO, TLR4, TNFA
		DDC	BDNF, GDNF, IGF1, IL2, LEP
		TPH2	GCR, LEP
		AOFA	APOE, TLR4
		AOFB	APOE, LEP
2	Tryptophan metabolism: hsa00380	I23O1	ALBU, CCL2, GCR, HMGB1, IFNG, IL1B, IL2, IL6, IL8, IL10, LEG3, LEP, TLR4, TNFA, TNR1B
		DDC	BDNF, GDNF, IGF1, IL2, LEP
		I23O2	IL2, LEP
		AOFA	APOE, TLR4
		AOFB	APOE, LEP
		INMT, SNAT	-

Table 11. Age and gender characteristics of the patients.

Group	Sex (M/F)	Min. age	Max. age	Avg. age	Median	Standard Deviation
Non-POD	11/16	65	75	69.6	70	3.0
POD	5/7	65	79	69.7	69.5	4.3

Table 12. Templates for molecular genetic pathways of metabolic pathway enzymes regulation by POD genetic markers.

№	Template name	Regulatory pathway
P1	Protein-protein interactions	**PODp** –protein-protein interaction→ **EnzP**
P2	Protein function regulation	**PODp** –regulation of activity/degradation/catalyze/transport→ **EnzP**
P3	Expression regulation	**PODp** –expression regulation→ **EnzG** –expression→ **EnzP**
P4	Double expression regulation	**PODp** –expression regulation→ **Hg** –expression→ **Hp** –expression regulation→ **EnzG** –expression→ **EnzP**
P5	Combined regulatory pathway	**PODp** –expression regulation→ **Hg** –expression→ **Hp** – regulation of activity/degradation/catalyze/transport→ **EnzP**

Abbreviations:
PODp – POD genetic markers proteins; Hg – human genes; Hp – human proteins; EnzG – genes encoding metabolic pathway enzymes; EnzP – metabolic pathway enzymes.

Figure Legends

Fig. 1. Classification accuracy for distinguishing between postoperative delirium (POD) and non-POD patient groups.
Legend: blue line – classifier (AUC = 0.99), black dashed line – chance level(AUC = 0.5)

Fig. 2. Gene network of regulation of "Glycine, serine, and threonine metabolism" enzymes by POD genetic markers.
Legend: double-strand spiral – Gene; red sphere – Protein; yellow rectangle arrow – activity downregulation; yellow triangle arrow – activity upregulation; light blue rectangle arrow – degradation downregulation; turquoise triangle arrow – expression; pink rectangle arrow – expression downregulation; pink arrow with rhombus – expression regulation; pink triangle arrow – expression upregulation; black line – protein-protein interaction; dark blue arrow with rhombus – transport regulation.

Fig. 3. Gene network of regulation of "Aminoacyl-tRNA biosynthesis" metabolic pathway enzymes by POD genetic markers.
Legend: double-strand spiral – Gene; red sphere – Protein; yellow triangle arrow – activity upregulation; violet arrow – cleavage; turquoise triangle arrow – expression; pink rectangle arrow – expression downregulation; pink arrow with rhombus – expression regulation; pink triangle arrow – expression upregulation; black line – protein-protein interaction. The objects outlined in red ovals illustrate the participants of the example regulatory pathways within the gene network.

Fig. 4. The number of regulated enzymes in KEGG metabolic pathways for each of the POD genetic markers.
Legend: blue bars – enzymes of "Glycine, serine, and threonine metabolism" metabolic pathway; orange bars – enzymes of "Aminoacyl-tRNA biosynthesis" metabolic pathway.

Fig. 5. Gene network of regulation of lactate metabolism enzymes by POD genetic markers.
Legend: double-strand spiral – Gene; red sphere – Protein; yellow triangle arrow – activity upregulation; light blue rectangle arrow – degradation downregulation; light blue arrow with

rhombus – degradation regulation; light blue triangle arrow – degradation upregulation; turquoise triangle arrow – expression; pink arrow with rhombus – expression regulation; black line – protein-protein interaction.

Fig. 6. Gene network of regulation of inositol metabolism enzymes by POD genetic markers.
Legend: double-strand spiral – Gene; red sphere – Protein; yellow arrow with rhombus – activity regulation; yellow triangle arrow – activity upregulation; turquoise triangle arrow – expression; pink arrow with rhombus – expression regulation; black line – protein-protein interaction.

Fig. 7. Gene network of regulation of GABA metabolism enzymes by POD genetic markers.
Legend: double-strand spiral – Gene; red sphere – Protein; light blue rectangle arrow – degradation downregulation; turquoise triangle arrow – expression; pink arrow with rhombus – expression regulation.

Fig. 8. Gene network of regulation of serotonin metabolism enzymes by POD genetic markers.
Legend: double-strand spiral – Gene; red sphere – Protein; yellow arrow with rhombus – activity regulation; yellow triangle arrow – activity upregulation; turquoise triangle arrow – expression; pink arrow with rhombus – expression regulation.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- DeliriumSupplementaryTables.xlsx