scientific reports

Check for updates

OPEN

UHPLC-MS/MS-based untargeted metabolite profiling of Lyme neuroborreliosis

Ilari Kuukkanen[1,2]✉, Annukka Pietikäinen[2,3,4], Tiia Rissanen[5], Saija Hurme[5], Elisa Kortela[6], Mari J. Kanerva[7], Jarmo Oksi[7], Jukka Hytönen[2,3,4,8] & Maarit Karonen[1,2,8]

The diagnosis of Lyme neuroborreliosis (LNB) requires the demonstration of intrathecal synthesis of *Borrelia* antibodies in a patient's cerebrospinal fluid (CSF), which involves the invasive procedure of a lumbar puncture. This study serves as a feasibility study aimed at exploring the potential of using serum samples, which are easily obtainable routine clinical samples, for LNB diagnostics via advanced metabolomics techniques. Serum samples were collected from confirmed LNB patients before and after treatment, with post-treatment samples serving as controls. The objective of the study was to find stable biomarkers for acute LNB through untargeted metabolomics analysis using ultrahigh-performance liquid chromatography coupled with tandem mass spectrometry (UHPLC-MS/MS). The study focused on biomarkers that could be reliably detected in serum samples stored under typical clinical conditions, without the need for special handling, ensuring consistent detection over time. The analysis revealed 26,978 molecular features (MFs), of which 1,746 were statistically significant ($p < 0.001$). Further manual investigation into 91 of the most prominent MFs revealed 53 potential biomarkers for LNB, individually or in combination. The workflow developed provides a comprehensive platform for biomarker detection, with potential applications in both research and clinical settings for LNB and other infections. This minimally invasive diagnostic approach is promising, and additional validation and future studies are needed for it to be considered as a practical alternative or a complement to CSF-based diagnostics of LNB in everyday clinical practice.

Keywords *Borrelia burgdorferi*, Diagnostic biomarkers, Lyme neuroborreliosis, Metabolomics, Serum analysis, UHPLC-MS/MS

Abbreviations

AGC	Automated gain control
APAP	Acetaminophen
BEH	Ethylene bridged hybrid
CEF	Ceftriaxone
CSF	Cerebrospinal fluid
dd-MS^2	Data-dependent tandem mass spectrometry
DOX	Doxycycline
EFNS	European Federation of Neurological Societies
EM	Erythema migrans
HESI	Heated electrospray ionization
IS	Internal standard
IT	Injection time (i.e. the time ions can pass into the C-trap)
LB	Lyme borreliosis
LNB	Lyme neuroborreliosis
MF	Molecular feature
MS/MS	Tandem mass spectrometry

[1]Department of Chemistry, University of Turku, Turku, Finland. [2]TBD Turku, University of Turku, Turku, Finland. [3]Institute of Biomedicine, University of Turku, Turku, Finland. [4]Clinical Microbiology, Tyks Laboratories, Turku University Hospital, Turku, Finland. [5]Department of Biostatistics, University of Turku and Turku University Hospital, Turku, Finland. [6]HUS Inflammation Center, Helsinki, Finland. [7]Department of Infectious Diseases, Turku University Hospital and University of Turku, Turku, Finland. [8]Jukka Hytönen and Maarit Karonen contributed equally to this work. ✉email: ilari.j.kuukkanen@utu.fi

MSI	Metabolomics standards initiative
MSX	Multiplexing
PDA	Photo diode array
PTFE	Polytetrafluorethylene
RCH	Relative centrifugal force
RF	Average response factor
RP	Reversed phase
UHPLC	Ultrahigh-performance liquid chromatography

Lyme borreliosis (LB) is an infectious disease caused by a group of spirochete bacteria belonging to the *Borrelia burgdorferi* sensu lato complex (later referred to as *Borrelia*). These bacteria are spread in nature by ticks from genus *Ixodes* and are mobilized from the tick's gut to a new host during a blood meal. *Borrelia* is widely spread and is reported to be the most common tick-borne pathogen in the Northern Hemisphere[1]. The early localized LB typically manifests as a migrating local skin infection, erythema migrans (EM), around the tick's feeding site[2]. EM may be accompanied by non-specific and general symptoms such as fatigue, fever, arthralgia, headache, and myalgia[3,4]. Without proper treatment, the infection can disseminate to affect other organs from the site of the EM, including the heart, skin and the nervous system[5]. Lyme neuroborreliosis (LNB) is a disseminated form of LB in which *Borrelia* bacteria have spread from the initial tick bite site to the nervous system. Approximately 15% of diagnosed LB infections progress to LNB, which manifests as various symptoms and findings including lymphocytic meningitis, painful meningoradiculitis, cranial neuropathy, peripheral neuropathy, and, rarely, encephalitis[1,6–8]. The early diagnosis and prompt treatment of LB, in all its forms, is recommended for patient recovery and the prevention of long-term complications.

In the early phase, the diagnosis of LB is based on the clinical appearance of EM, while in the disseminated forms laboratory tests are critical. The primary approach in LB laboratory testing is the detection of *Borrelia*-specific antibodies in the patient's serum. The diagnosis of definite LNB, however, is based on combination of suggestive neurological symptoms, cerebrospinal fluid (CSF) pleocytosis, and demonstration of intrathecal synthesis of *Borrelia* antibodies found in the patient's CSF samples[6,8,9]. However, it's important to note that antibody levels in the patient's CSF can remain elevated for months after a successful antibiotic treatment, which may result in false LNB diagnoses in certain cases[10,11]. The false-positive assay results for LB can also occur due to cross-reacting antibodies due to other infections[12,13]. Therefore, the interpretation of laboratory results requires careful consideration of the clinical context and patient history. These difficulties highlight the necessity for more robust and dependable methodology in LB and LNB diagnostics.

Diagnostic metabolomics is a rapidly advancing field of omics that offers reliable and sensitive methods to investigate low molecular weight metabolites (< 1,500 Da) within biological systems. This approach incorporates both quantitative and qualitative analyses of biological samples to provide disease-specific metabolite fingerprints. Moreover, metabolite profiling can provide insights into metabolic pathways and cellular processes affected during a disease[14].

In metabolomics, ultrahigh-performance liquid chromatography integrated with tandem mass spectrometry (UHPLC-MS/MS) has emerged as a potent analytical platform due to its high sensitivity, selectivity, and applicability for both targeted and untargeted analysis[15–18]. LC-MS-based metabolomic approaches have been successfully applied in various biomedical research areas, such as infectious diseases and cancer[19,20], and in LB research[21–26].

The LB metabolomics studies conducted by Molins et al. in 2015 and 2017 showcased the efficacy of LC-MS-based metabolomic profiling in distinguishing early LB patients from healthy controls and from diseases of differential diagnostic relevance[22,23]. Building on the established methodologies, our study introduces a UHPLC-MS/MS-based analytical method tailored for untargeted metabolomics analysis, capable of accommodating the variability inherent in real-life samples. The method is particularly focused on detecting stable serum biomarkers for LNB infection. To achieve this, the study utilizes a unique set of serum samples collected as part of routine diagnostics from definite LNB patients at two critical time points: the acute pretreatment phase and the post-treatment convalescent phase, 12 months after antibiotic treatment. This distinctive study design enables a comprehensive analysis of metabolomic changes associated with disease improvement.

Materials and methods

Chemicals

Analytical grade ethanol, analytical grade methanol and the LC-MS grade formic acid were acquired from VWR International (Fontenay-Sous-Bois, Paris, France). The LC-MS grade acetonitrile was acquired from Fisher Scientific (Loughborough, Leicestershire, United Kingdom). Deuterated standards, DL-phenyl-d_5-alanine and L-tryptophan-(indole-d_5) were acquired from Cambridge Isotope Laboratories Inc. (Andover, Massachusetts, United States of America). Ultra-pure type I water was generated using the Merck Millipore Synergy UV system.

Serum samples

The LNB patients' serum samples for this study were collected as part of routine diagnostics and were included in our previous study[27], which compared the efficacy of intravenous ceftriaxone and oral doxycycline in the treatment of LNB. All individuals participating in the original study provided informed consent and ethical approval was granted by the National Committee on Medical Research Ethics in Finland. The diagnosis of LNB was confirmed according to the European Federation of Neurological Societies (EFNS) guidelines and criteria[8]. The present investigation adhered to the ethical principles outlined in the Declaration of Helsinki for medical research involving human material and data.

To maintain confidentiality, all samples were coded, ensuring that no personal identifiable information was handled, except for age and sex. The sample material comprised serum samples from 81 individual definite LNB patients. Among them, 68 patients had both acute pretreatment and post-treatment samples, with the post-treatment samples serving also as control samples for the pretreatment samples. Additionally, six patients with only pretreatment samples and seven with only post-treatment samples were included in the analysis. Consistent with standard clinical practice, the samples were stored at -20 °C, a common storage temperature for routine diagnostic samples, rather than -80 °C, which is primarily used when samples are collected for research purposes. The freezers and freezer rooms used for storage did not have a frost-free system. The sample collection and antibiotic treatments were conducted in Turku University Hospital, Turku, Finland (49 patients) and in Helsinki University Hospital, Helsinki, Finland (32 patients).

Sample preparation and analytical workflow

The sample preparation method was designed with a systematic approach to ensure the precise analysis of serum metabolites. By implementing controlled procedures, the method optimizes consistency and a reliable foundation for subsequent UHPLC-Orbitrap-MS/MS analysis, enhancing the precision of metabolomic profiling.

Random sampling was conducted, while ensuring that samples matching the same patient were analyzed during the same analysis time. Samples were left to melt slowly, while vortexing them at 750 rpm. After vortexing, 150 µL of serum was transferred to a new Eppendorf tube. Next, 50 µL of deuterated 7.125 µM L-tryptophan-(indole-d_5) aq. internal standard (IS) solution was added to the serum and vortexed for 15 min at 750 rpm. To precipitate macromolecules, mainly proteins, from the serum, 750 µL of cold methanol was added to the mixture and the solution was vortexed for 30 min at 750 rpm. Following the precipitation (Supplementary Information Fig. S1-S3 and Table S1), samples were centrifuged for 20 min at a relative centrifugal force (RCF) of 21,913 *g*. The supernatant (700 µL) was transferred to a new Eppendorf tube and dried under vacuum.

The dried samples were dissolved with 150 µL of deuterated 10 µM DL-phenyl-d_5-alanine aq. IS solution. The dissolved samples were vortexed for 15 min at 750 rpm before undergoing centrifuge filtration through 0.2 µm polytetrafluorethylene (PTFE) micro centrifugal filters for 20 min at an RCF of 9,056 *g*.

Following filtration, the samples were pipetted into a 700 µL 96-well plate and analyzed with UHPLC-Orbitrap-MS/MS preceding in silico metabolomics (Supplementary Information Table S2) and statistical analysis. Blank samples, prepared in the same manner as the serum samples, were used as controls to monitor the sample preparation process and assess matrix effects during the analysis.

UHPLC-Orbitrap-MS/MS analysis

The development of our UPHLC-Orbitrap-MS/MS method involved optimization of parameters, with a specific focus on the detection of low molecular weight serum metabolites. Employing positive ion mode, we fine-tuned the MS instrument to ensure accurate detection and quantification of metabolites. A mass range of 70−1,050 Da was found most suitable and was used for the metabolomic and statistical analysis.

Prior to the UHPLC-Orbitrap-MS/MS analysis, the order of serum samples was randomized, and the instrument underwent calibration using Thermo Scientific™ Pierce™ LTQ Velos ESI positive ion calibration solution. The serum samples were analyzed using an ultrahigh-resolution UHPLC-PDA-HESI-QOrbitrap-MS/MS platform, which comprised of an Acquity UPLC® system equipped with a photodiode-array (PDA) detector (Waters Corporation, Milford, MA, USA). This system was coupled to a Q Exactive™ hybrid quadrupole-Orbitrap™ mass spectrometer (Thermo Fisher Scientific GmbH, Bremen, Germany) via heated electrospray ionization (HESI) source.

The reversed phase (RP) column used in the study was an Acquity UPLC® BEH Phenyl 1.7 µm 2.1 × 100 mm column (Waters Corporation, Wexford, Ireland). Elution was carried out using acetonitrile (A) and 0.1% formic acid (B) as eluents at a flow rate of 0.5 mL × min^{-1}. The elution profile was as follows: 0.0–0.5 min, 0.1% A in B (isocratic gradient); 0.5–7.5 min, 90% A in B (linear gradient); 7.5–8.5 min, 90% A in B (isocratic gradient, column wash); 8.5–8.6 min, 0.1% A in B (linear gradient); 8.6–10.1 min, 0.1% A in B (isocratic gradient, column stabilization). The injection volume was 5 µL with a full loop overfill factor of 3. The UV data, in the range of 190–500 nm, was collected with the PDA detector during the 10.1-minute analysis, while high-resolution MS data was recorded between 0.0 and 7.5 min. The UHPLC flow was diverted to waste during the column wash and stabilization periods.

MS data was recorded in positive ion mode within mass ranges of 70−1,050. Full MS scan resolution was set at 70,000, with Automatic Gain Control™ (AGC) target value at 3×10^6 and a maximum injection time (IT) of 200 ms. Data-dependent MS/MS (dd-MS^2) scans were collected at a resolution of 17,500, with AGC target value of 1×10^5 and maximum IT of 50 ms. The dd-MS^2 scan utilized Top N technique, with maximum precursor multiplexing per scan (MSX) count of 1 and loop count of 5, resulting in a Top N value of 5, where five of the most intensive ions were selected for dd-MS^2 fragmentation. Additionally, a lock mass of m/z 214.08963 was utilized to enhance mass accuracy.

The settings for the HESI-source included a capillary temperature of 380˚C, a spray voltage of 3,800 V, a sheath gas (N_2) flow rate of 60 (arbitrary units), and an auxiliary gas (N_2) flow rate of 20 (arbitrary units). The S-lens average response factor (RF) level was set to 60 (arbitrary units), and the probe heater temperature was maintained at 300˚C. ISs, comprising both DL-phenyl-d_5-alanine and L-tryptophan-(indole-d_5), were analyzed in duplicate with every 10 injections.

In silico analysis of metabolomic data and molecular feature identification

Generated UHPLC-MS/MS data in RAW format was directly subjected to metabolomic analysis using Compound Discoverer 3 (Thermo Fisher Scientific Inc., Waltham, MA, USA; version 3.1.0.305). Analysis employed the built-in "Untargeted Metabolomics with Statistics Detect Unknowns with ID using Online

Databases and mzLogic" workflow. Compound Discoverer 3 is software designed for LC-MS data processing, with automated identification, quantification, annotation, and interpretation of the results. Manual valuations and identifications were done using Xcalibur™ software version 4.1.31.9 (Thermo Fisher Scientific GmbH, Bremen, Germany). The metabolomic data contained metabolite-specific information such as the compound-specific exact mass, retention times, and corresponding integrated peak areas for all the detected MFs. The metabolomic quantification file was then transformed into CSV format for further statistical analysis. The parameters used in the metabolomic analysis and the details regarding the MS databases employed for both manual and in silico identification can be found in Supplementary Information Table S2. The final identified MFs are categorized according to their identification confidence levels (see Supplementary Information Table S3). These classifications align with the 5-level system proposed by Schymanski et al. (2014), which is rooted in the Metabolomics Standards Initiative (MSI) framework by Sumner et al. (2007)[28,29]. The level 1 identification corresponds to validated identification (confirmed structure), level 2 to putative identification (MS/MS match to literature), level 3 to tentative structure (database/literature match to molecular formula), level 4 to matching molecular formula (including isotope distribution, charge state and adduct ion formation), and level 5 to a unique feature.

Statistical analysis

Differences between the pretreatment and post-treatment samples of all 26,978 MFs were summarized with descriptive statistics and studied with Wilcoxon signed-rank test, due to the non-normality of the distributions. The compounds which were statistically significant (level set at 0.001) in both datasets (raw and standardized) were examined more closely and visually with a volcano plot. For further analysis, 91 of these MFs were selected based on the largest difference between samples.

Linear mixed models for repeated measures were used for further analysis of all 91 MFs variables to evaluate the effects of patients' characteristics. The models included one within-factor (timepoint: pretreatment, 12 month post-treatment), and several between factors (sex, age, antibiotic treatment, hospital, acetaminophen use and LNB symptom duration). The compound symmetry covariance structure was used for time. Logarithmic transformations were used to achieve the normal distribution assumption of MF variables.

The normality of variables was evaluated visually and tested with the Shapiro-Wilk test. Tests were performed as two-sided with a significance level set at 0.05. The analyses were carried out using RStudio (version 2023.03.0.386) based on R (version 4.3.0; RStudio, PBC, Boston, MA, USA).

Results

We developed a rapid and easily adoptable UHPLC-MS/MS method with an efficient sample preparation protocol for the untargeted profiling of serum samples from patients with definite LNB. This analytical method was also designed to process a large number of serum samples within a single analysis timeframe, making it well-suited for high-throughput studies. To assess the sensitivity and repeatability of the method, ISs were utilized during both the method development and actual analyses. An overview of the integrated workflow is presented in Fig. 1.

Serum macromolecules pose a challenge to chromatographic separation and MS ionization during UHPLC-MS/MS analysis, potentially compromising sensitivity and accuracy. To mitigate this, we employed a precipitation process using cold methanol to achieve effective precipitation of macromolecules. During method development, we initially tested the commonly used 1:3 (v/v) ratio for the actual patients' serum samples. However, this ratio proved insufficient for our analytical platform, as it left residual proteins and possible other macromolecules in the extracts. When combined with the fact that acetonitrile was used as the organic eluent in our UHPLC-MS/MS analysis, these residuals caused significant system variability, including increased column pressure, diminished chromatographic performance, and interferences in the HESI. The efficiency of six different serum:methanol (v/v) ratios were evaluated to optimize the precipitation of macromolecules, mainly proteins, as detailed in Supplementary Information Fig. S1–S3 and Table S1. During the precipitation, an aqueous (aq.) IS, L-tryptophan-(indole-d_5), was added to samples to monitor the possible analyte loss. In the actual analyses of samples, this ISs served as an indicator for assessing the efficacy of the sample preparation and ensuring the integrity of the analytes throughout the preparation process.

The results of the optimization of macromolecule precipitation showed that serum:methanol ratios from 1:5 to 1:8 (v/v) produced the highest yields compared to the ratios of 1:3 and 1:4 (see Supplementary Information Table S1). The 1:5 (v/v) ratio produced extracts with less macromolecules, ensuring more stable system performance throughout the analysis and providing a balance between optimal MS results and practical sample handling. Therefore, considering effective precipitation, practicality and ease of sample handling, a decision was made to employ the 1:5 (v/v) ratio in the actual analyses.

To monitor the sample preparation process and to assess method performance, stable isotopically labeled L-tryptophan-(indole-d_5) and DL-phenyl-d_5-alanine were employed as ISs. These ISs helped to evaluate potential matrix effects, such as MS signal suppression or enhancement. The chemical structures, corresponding chromatograms and UV spectra, calibration curves, and repeatability measurements of the ISs are provided in the Supplementary Information (Fig. S4-S9). The results demonstrated that L-tryptophan-(indole-d_5) and DL-phenyl-d_5-alanine function as reliable ISs, meeting the criteria for stability, repeatability, and detectability. Furthermore, the MS parameters were optimized, particularly for small metabolites (see section "UHPLC-Orbitrap-MS/MS analysis"). This optimization ensured that the MS instrument correctly detected metabolites with precision.

Serum samples were obtained from 81 patients diagnosed with definite LNB. Patient ages ranged from 17 to 88 years, with 28 females and 53 males included in the study. The average age of patients was 57 years, with female patients averaging 63 years and male patients averaging 53 years. Among the female patients, 16 were treated with oral doxycycline (DOX), while 12 received intravenous ceftriaxone (CEF). Among male patients,

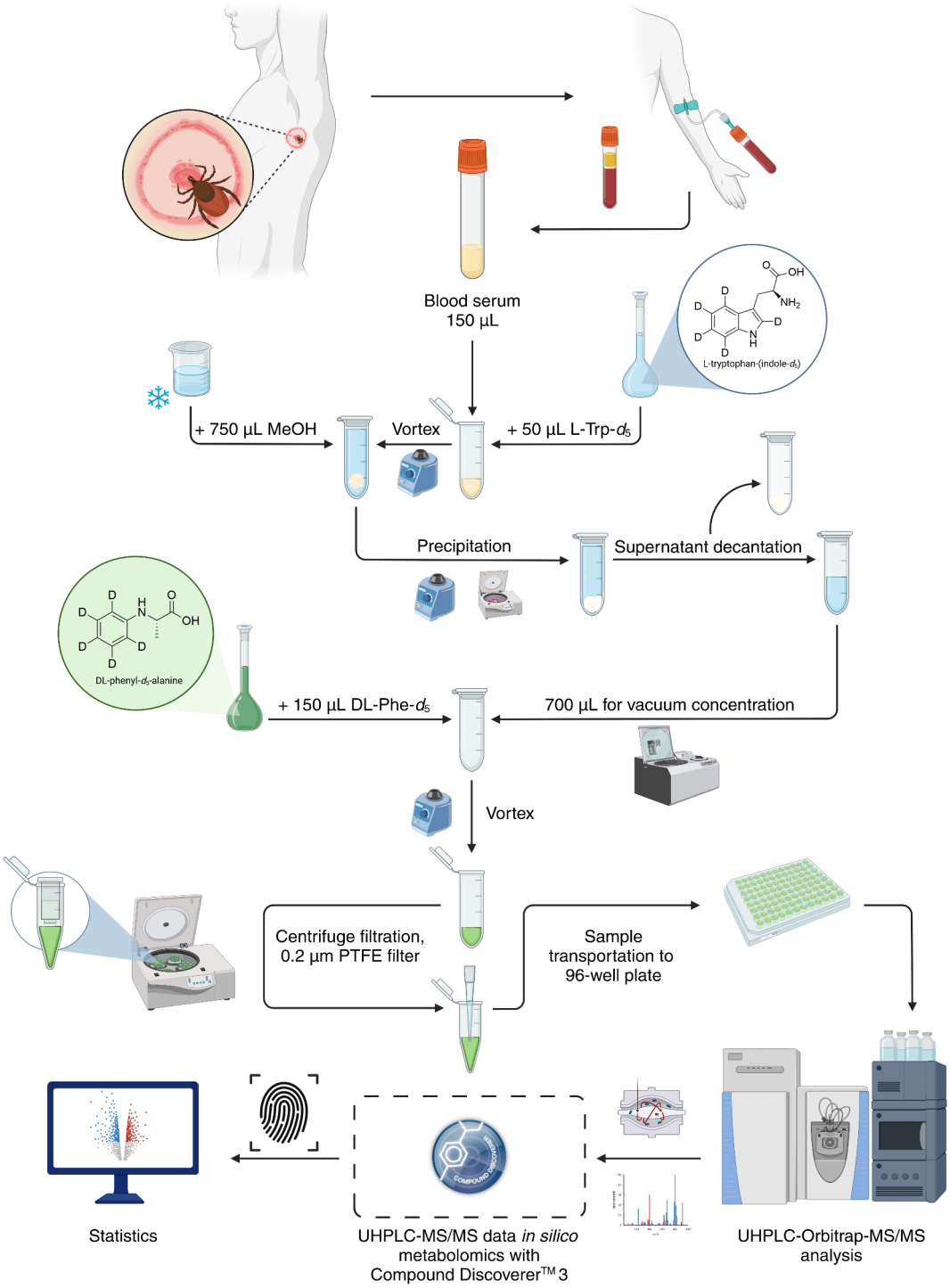

Fig. 1. Integrated workflow of the preparation of serum samples, ultrahigh-performance liquid chromatography combined with tandem mass spectrometry (UHPLC-MS/MS) analysis, in silico metabolomics, and statistics.

24 received DOX and 29 were administered CEF. In total, 40 patients underwent DOX treatment, while 41 received CEF. Ten of these patients had previously experienced one or more episodes of LB. The baseline details of the patients, along with their treatment strategies, some receiving oral DOX and others intravenous CEF, are presented in the previous study[27]. The distributions of patients' age, sex, treating university hospital, and antibiotic treatments in the analyzed samples are presented in Fig. 2. In addition to these, the use of APAP and the duration of symptoms experienced by the patients were also taken into account.

The pretreatment and post-treatment serum samples were analyzed using UHPLC-MS/MS. The in silico metabolomic analysis of the MS data revealed the presence of 26,798 MFs, i.e., peak with a specific *m/z* value

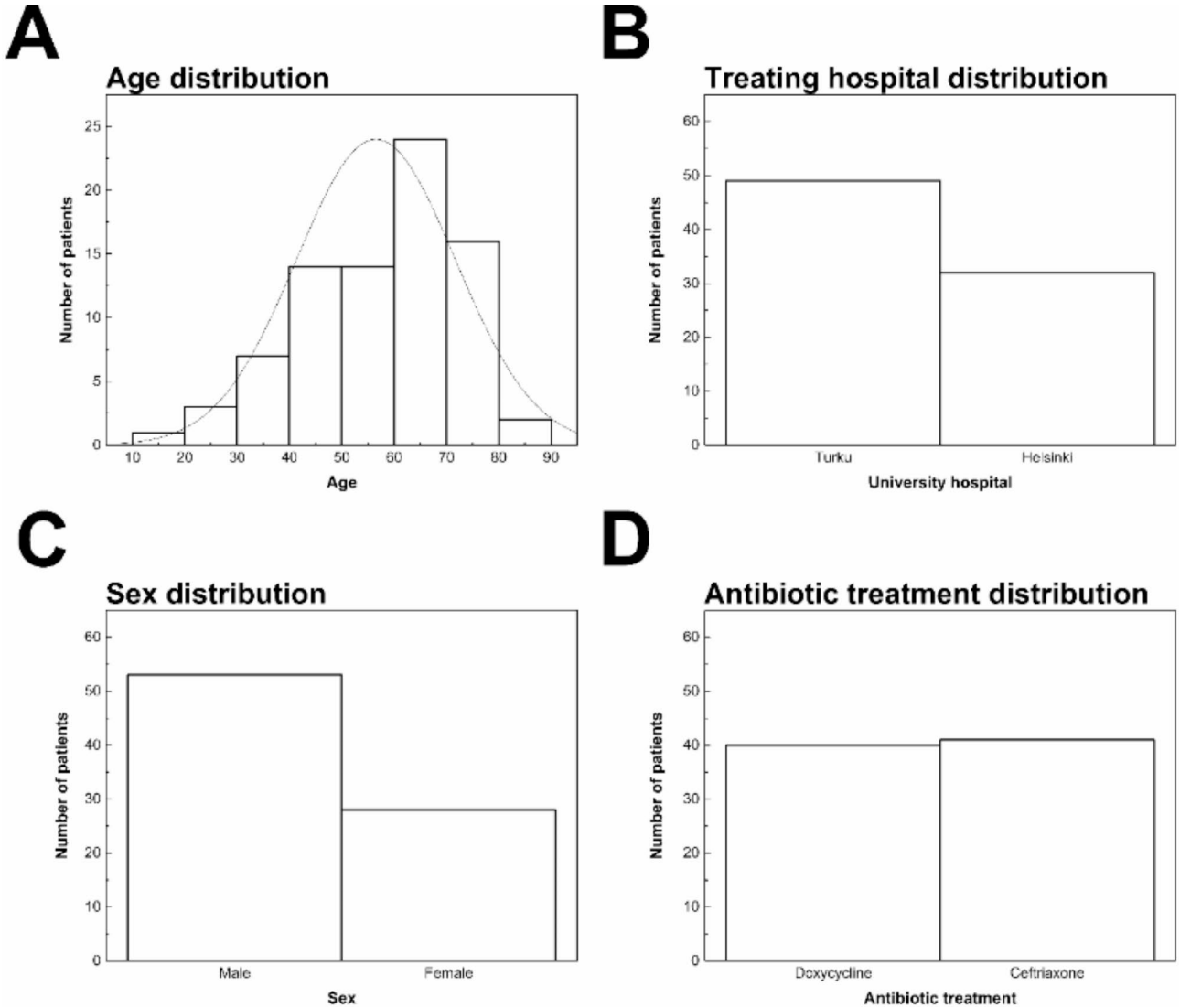

Fig. 2. Distribution of patients' age (**A**), treating university hospital (**B**), sex (**C**) and antibiotic treatment (**D**) among Lyme neuroborreliosis (LNB) patients (n = 81), whose serum samples were analyzed in this study[27].

and peak area. The built-in automated *in silico* identification algorithm provided a preliminary characterization for 5,735 MFs based on their exact mass, fragmentation patterns, and MS database matches. Due to the inherent sensitivity of metabolite profiling to technical variations in UHPLC-MS/MS analyses, as well as unique variations among individual patients, a systematic statistical approach was adopted.

The statistical analyses utilized the IS-standardized data and aimed to identify shared aspects of the metabolic dynamics associated with LNB. First, the analysis focused on comparing metabolomic profiles in the pretreatment and post-treatment samples of each individual patient. This approach aimed to detect alterations in the metabolic profiles between acute infection and convalescence thus revealing metabolite fingerprints of untreated LNB. Second, the statistical analyses sought to identify similar metabolomic profile changes across the different individuals, providing insights into commonalities and variations in the metabolic responses in LNB.

The statistical analysis highlighted 1746 prominent MFs ($p < 0.001$) from the detected 26,798 MFs. Closer manual inspection further narrowed down this selection to 91 MFs based on the most substantial changes in the detected peak areas. These changes, whether increases or decreases, were required to be at least 2-fold, with the alteration in median peak areas measuring a minimum of 5×10^5 (counts × seconds) between the pretreatment and the post-treatment group medians. The 1746 statistically significant MFs are presented in Fig. 3.

Further analysis was applied to the 91 most prominent MFs. There were only few MFs where patient's sex or received antibiotic treatment had a statistically significant effect. Specifically, only ten out of the 91 MFs (11.0%) were statistically significantly affected by the patient's sex, and four out of 91 MFs (4.4%) showed statistically significant impact of the antibiotic treatment. Patient age was influenced by eight out of the 91 MFs (8.8%), while the hospital of sample origin affected 32 out of the 91 MFs (35.2%). Treating hospital and patient age had a common association only in three MFs (3.3%). A total of 50 MFs (54.9%) were directly or indirectly linked

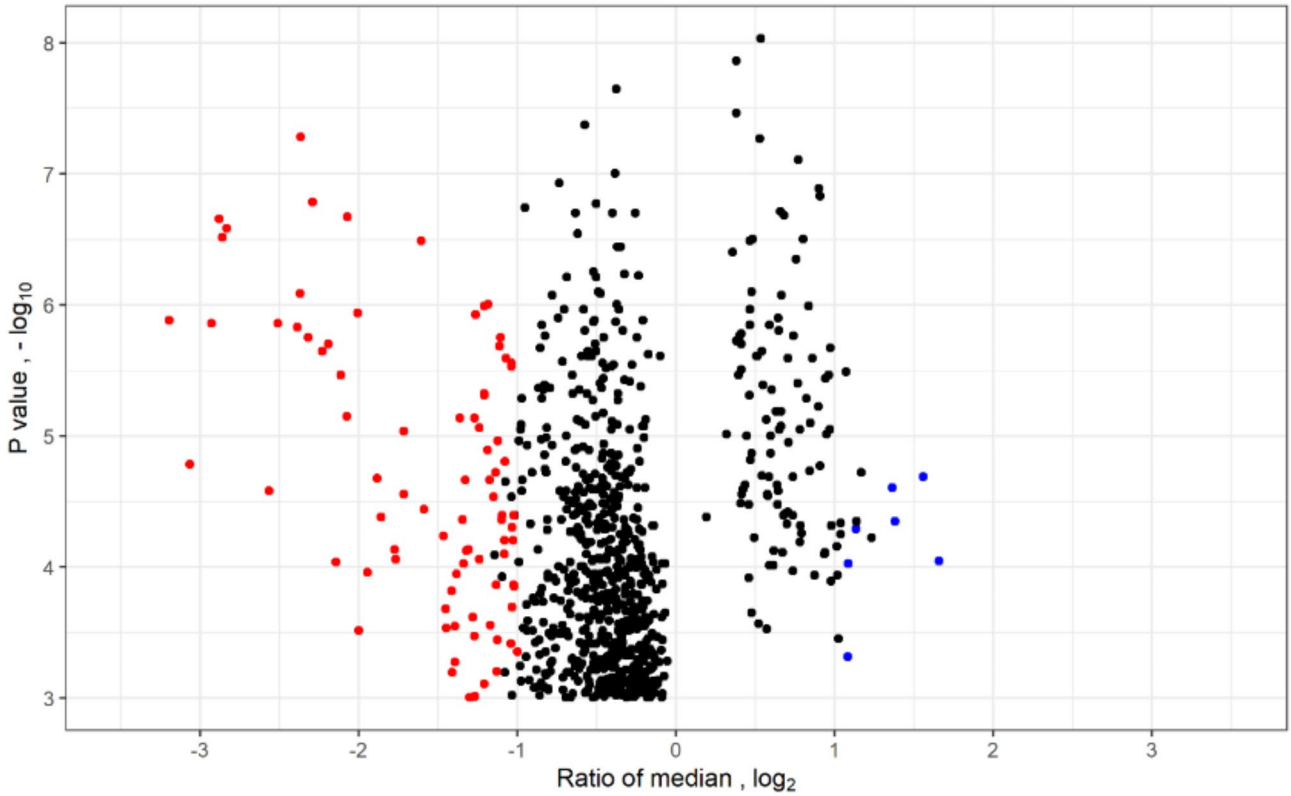

Fig. 3. Volcano plot of 1746 molecular features (MFs) with significant alterations ($p < 0.001$) out of the total 26,978 MFs in the metabolomic analysis data obtained via ultrahigh-performance liquid chromatography coupled with tandem mass spectrometry (UHPLC-MS/MS). Notably, 91 most prominent MFs are marked by red and blue. The 84 MFs highlighted in red signify a notable drop in peak areas, indicating a significant decrease in MF concentration, while those seven MFs marked in blue represent a rise in peak areas, indicating a substantial increase in MF concentration. These observations are particularly noteworthy, as they reflect the most substantial alterations between the acute pretreatment and post-treatment samples.

to the use of APAP (statistically significantly). In contrast, only one MF (1.1%) was significantly affected by the patient's estimation of LNB symptom duration prior to diagnosis.

Among the 91 most prominent MFs, some had automatic pre-identification matches and classifications through in silico metabolomic analysis. These putative identifications were verified manually to ensure correct identification. Additionally, the remaining MFs without automatic identification underwent manual investigation, where each MF was examined and characterized based on features such as exact mass, molecular formula derived from the exact mass, available MS/MS spectra, and potential matches from MS databases. Most of the compounds were polar in nature, with a median retention time of 2.14 min (range: 0.63–7.12 min). The metabolite identifications are presented in Supplementary Information Table S3.

The chemical identification of the 91 most prominent MFs resulted in 23 MFs with assigned putative structures, 49 MFs with a predicted molecular formula, and 42 MFs with unique features (exact *m/z* value). These MFs could be classified into four different groups: (i) MFs derived from medication used by the patient (17 MFs), (ii) MFs not directly derived from medications but, based on statistical analyses, strongly associated with medication use (21 MFs), (iii) known general metabolites, which could be identified based on the databases and literature (5 MFs), and (iv) unknown MFs which could serve as potential biomarkers for LNB in the future (48 MFs). For more detailed information about the MFs see Supplementary Information Table S3.

Discussion

Through UHPLC-MS/MS metabolomics, we explored the complex metabolic landscape of LNB, aiming to uncover new insights that may aid in the diagnosis of this infection. Our analysis revealed a substantial number (26,798) of MFs through in silico metabolomic analysis, necessitating a systematic approach to address technical and individual variations across samples and individuals. The statistical strategy, which compared metabolite profile changes between pretreatment and post-treatment samples within each patient and across different individuals, was effective in identifying stable but prominent MFs across samples. We identified 1,746 MFs of significant statistical relevance through this approach. This was followed by manual curation, which further refined the selection to 91 MFs.

It is understandable that the levels of substances used for medical purposes, such as the identified painkiller acetaminophen (APAP, paracetamol) and its metabolized forms, are different between the acute infection and

the convalescent stage. The metabolites in the first group were easily identified based on the MS databases used and the common phases I and II metabolisms causing well-known mass shifts, such as the sulfate conjugation of APAP[30–38]. In contrast, the second group of MFs stood out the most because of their strong association with APAP use among the LNB patients. These features represent metabolites that are directly or indirectly linked to APAP use prior to the sample collection and cannot currently be reliably used as biomarkers for LNB. Patients who used APAP were likely to have been in pain, potentially reflecting a specific type of disease presentation i.e., radiculitis. Consequently, the observed MFs associated with APAP use may relate to both the nature of the underlying infection and the effects of the medication itself.

The known metabolites in the third group showed the presence of primary metabolites that have been previously linked with i.e., acylcarnitine metabolism and arginine metabolism but may not be used as biomarkers specific for LNB[39–50]. For example, elevated L-homoarginine (**MF#7**) levels are potentially linked to cardiovascular health, such as improved endothelial function, reduced platelet aggregation, and increased insulin secretion[45,51,52]. Conversely, low serum L-homoarginine levels might predict cardiovascular mortality[46]. Additionally, L-homoarginine is associated with liver cirrhosis and hyperargininemia due to its role as an organ-specific inhibitor of human liver and bone alkaline phosphohydrolase[47].

Another identified top MF, L-acetylcarnitine (**MF#1**), possesses distinct neuroprotective, neuromodulatory, and neurotrophic functions. It easily penetrates plasma membranes and crosses the blood-brain barrier. Additionally, it modulates various neurotransmitter systems by increasing acetyl coenzyme A levels and choline acetyltransferase activity, potentially mitigating neurodegenerative processes[48]. In animal models, L-acetylcarnitine is found critical for hippocampal function and elicits rapid and enduring antidepressant-like effects through epigenetic histone acetylation mechanisms[49]. Additionally, reduced L-acetylcarnitine levels have been found in individuals with major depressive disorder[49]. Conversely, heightened blood L-acetylcarnitine levels may correlate with inflammation or infection. These elevations, particularly in short-chain acylcarnitines are attributed to liver release during infection or stress/trauma to support B and T-cell synthesis in combating infections or injuries. For example, individuals with sepsis or septicemia often exhibit significantly elevated L-acetylcarnitine levels, with those exceeding 20 µmol/L facing up to a fivefold increased mortality risk[44]. Fitzgerald et al. 2020 and 2021 reported the detection of L-acetylcarnitine in their studies, examining the host metabolic response to early LB and the metabolic response to post-treatment LB symptoms/syndrome, based on an authentic chemical standard and matching MS/MS pattern[25,26]. Our statistical analysis revealed that L-acetylcarnitine exhibited the most significant change in MF between pretreatment and post-treatment phases, showing a notable increase across individuals. This finding is consistent with the results reported by Fitzgerald et al.[25,26]. The third group also included three other tentative MF identifications shown in Table S3 [53–56].

The fourth group contains the unique MFs with potential for future biomarkers for LNB. These are so far unknown but could be used alone or in combination for rapid and less invasive diagnostic aid for LNB. Previously, Molins et al. (2015) have developed the metabolic biosignature for the detection of early LB using similar LC-MS-based methodology for small molecule metabolites and detected 95 MFs that distinguished LB patients from healthy controls[23]. In that study, the majority of tentatively identified metabolites were lipophilic or lipid structures leading to the hypothesis that infection induces changes in lipid mediators and markers of the inflammatory response. However, 91 MFs selected based on the statistical analyses in our study were different from the 95 MFs selected by Molins et al. (2015). The only potentially shared MF was detected at *m/z* 464.19144 (**MF#3**) corresponding to the molecular formula of $C_{23}H_{29}NO_9$ with mass error of 0.297 ppm. The top identification of Molins et al. for their MF at *m/z* 464.1916 was the peptide alanylcysteinyl-α-aspartylarginine (Ala-Cys-Asp-Arg), with the matching molecular formula $C_{16}H_{29}N_7O_7S$. However, our MS/MS fragmentation data did not support this identification of tetrapeptide structure to our MF. Molins et al. also reported more than five alternate chemical structures for this MF, though these structures were not specified, preventing us from comparing them with our data. Additionally, while Molins et al. detected several lipids in their study, we did not identify any lipids. However, their presence among the unknown biomarkers cannot be ruled out.

Linear mixed models for repeated measures (see section "Statistical analysis") were used to assess the potential effects of known variability in between factors (sex, age, antibiotic treatment, hospital, use of APAP and symptom duration) on the most prominent MFs. While the models revealed no large numbers of MFs that are statistically significantly affected by sex (ten MFs) and antibiotic treatment (four MFs). The hospital had statistically significant effect on a total of 32 MFs. Specifically, two MFs in group iii and 30 MFs in group iv were affected by the hospital where the patients were treated during the study and most importantly where the blood samples were taken. For example, in group iv, **MF#20** showed the largest decrease of approximately 90% on amount (normalized average peak areas of MFs) of samples from Helsinki University Hospital compared to those from Turku University Hospital ($p < 0.001$) and **MF#30** showed an increase of almost 365% ($p < 0.001$). However, the treating hospital factor did not have any significant ($p < 0.05$) effect on the group i medicinal MFs. We can only speculate that the variation in metabolic profiles due to the recruiting hospital may reflect differences in sample management protocols or environmental conditions, such as the material and possible medium of the serum containers. For future studies to minimize site-specific factors, it is necessary to control these conditions more carefully, i.e., standardized sample collection methods, including similar containers and identical sample processing protocols should be used.

The use of APAP significantly impacted 33 MFs that were distinct from those in group i. Notably, all 21 MFs in group ii were so strongly affected by APAP usage that their significance in the pretreatment versus post-treatment timepoint comparison was entirely diminished ($p > 0.05$). In group iv, seven MFs were also strongly impacted by APAP; however, they remained statistically significant in the timepoint comparison ($p < 0.05$). These findings are crucial, as they underscore the substantial effect of APAP on metabolic profiles. This highlights the importance of carefully accounting for external factors, such as medication use, when interpreting metabolomic data to prevent confounding effects that could obscure biologically meaningful results. Additionally, we also

investigated the potential influence of LNB symptom duration, experienced by the patients, on the detected MFs. However, our statistical analysis revealed that symptom duration did not contribute to the observed metabolic profiles in this study, apart from one MF.

Despite the evident impact of known variables on certain MFs, no single MF exhibited a consistent influence across all variables. Furthermore, the persistence of statistical significance in these MFs even after adjusting for all the variables underscores their consistency as stable and prominent biomarkers. This resilience suggests that these MFs are indicators of metabolic alterations associated with LNB. For more detailed information of how individual prominent MFs were affected, see Table S3.

While our findings provide valuable insights into the metabolic signatures of LNB, we recognize that the study's design and scope are still in the early stages. It is important to emphasize that this research should be viewed as a preliminary step in the exploration of metabolomics as a complementary diagnostic approach for LNB.

Conclusions

The UHPLC-MS/MS method appears to be a promising tool for metabolite detection in LB/LNB diagnostics. Its high sensitivity and accuracy make it an asset for uncovering metabolic signatures indicative of LNB and related conditions. It is essential, however, to note that the clinical applicability of these biomarkers will require thorough validation via additional studies.

We also recognize the limitations of our study. The retrospective nature of the clinical serum samples, stored at -20 °C rather than the more optimal -80 °C, may have influenced the metabolomic profiles, leading to the loss of specific metabolites. The extended storage period prior to analysis could have impacted metabolite stability, as some metabolites are prone to degradation or transformation over time. This degradation may introduce variability, complicating data interpretation.

However, it is worth noting that the storage was conducted in freezer rooms and conventional freezers that lacked frost-free systems. Unlike frost-free freezers, which cycle through temperature fluctuations and could further compromise sample integrity, these storage conditions likely minimized additional variability. Despite these limitations, our findings underscore the feasibility and practicality of using routinely stored clinical samples for metabolomic analysis in the context of disease diagnosis, providing a valuable approach for real-world applications.

We also acknowledge that some of the MFs might not be specific for LNB but rather be general infection markers. To address these issues, future research should include freshly collected serum or plasma samples from individuals with various disseminated forms of LB and with other infections for comparison.

During the metabolomic in silico analysis, certain compounds posed a challenge for the software due to their inherent tendency to produce fragments very easily. These fragments can masquerade as separate compounds, complicating the identification process to accurately detect genuine compounds from their fragmented counterparts. Therefore, other data processing tools, such as supervised or unsupervised machine learning, could provide easier ways to interpret the data, and possibly to distinguish patients with acute infections from previously infected or healthy subjects.

Additionally, it is worth noting that the sample handling process, while kept minimal and monitored using IS, may still introduce variability. Variables such as above-mentioned storage conditions, transportation, and processing time can influence the integrity of the samples and potentially impact the results. Future studies should aim to optimize sample handling protocols to limit potential variations introduced during this stage of the process.

In summary, this study is a preliminary exploration of metabolomics as a diagnostic tool for LNB. For the time being, lumbar puncture-based CSF analyses remain critical and established diagnostic methods for LNB. The results presented here, however, suggest that UHPLC-MS/MS-based metabolomics can serve as a complementary tool in LNB diagnostics. Further research, including validation using larger and more diverse patient cohorts, is necessary to determine the feasibility of incorporating metabolomics into routine clinical practice alongside traditional laboratory assays.

Data availability

The UHPLC-MS/MS raw data, the processed data from Compound Discoverer™ analyses and detailed statistical data presented in this study are available on request from the corresponding author.

Received: 20 September 2024; Accepted: 25 February 2025
Published online: 11 March 2025

References

1. Stanek, G., Wormser, G. P., Gray, J. & Strle, F. Lyme borreliosis. *Lancet* **379**, 461–473 (2012).
2. Strle, F. & Stanek, G. Clinical manifestations and diagnosis of Lyme borreliosis. *Curr. Probl. Dermatol.* **37**, 51–110. https://doi.org/10.1159/000213070 (2009).
3. Nadelman, R. B. et al. The clinical spectrum of early Lyme borreliosis in patients with culture-confirmed erythema migrans. *Am. J. Med.* **100**, 502–508 (1996).
4. Nadelman, R. B. & Wormser, G. P. Erythema migrans and early Lyme disease. *Am. J. Med.* **98**, 15S–24S (1995).
5. Stanek, G. et al. Lyme borreliosis: Clinical case definitions for diagnosis and management in Europe. *Clin. Microbiol. Infect.* **17**, 69–79 (2011).
6. Branda, J. A. & Steere, A. C. Laboratory diagnosis of Lyme borreliosis. *Clin. Microbiol. Rev.* **34**. (2021).
7. Wormser, G. P. et al. The clinical assessment, treatment, and prevention of Lyme disease, human granulocytic anaplasmosis, and babesiosis: Clinical practice guidelines by the infectious diseases society of America. *Clin. Infect. Dis.* **43**, 1089–1134 (2006).

8. Mygland, Å. et al. EFNS guidelines on the diagnosis and management of European Lyme neuroborreliosis. *Eur. J. Neurol.* **17**. (2010).
9. Kullberg, B. J., Vrijmoeth, H. D., van de Schoor, F. & Hovius, J. W. Lyme borreliosis: Diagnosis and management. *BMJ* **m1041** https://doi.org/10.1136/bmj.m1041 (2020).
10. Kalish, R. A. et al. Persistence of Immunoglobulin M or Immunoglobulin G antibody responses to *Borrelia burgdorferi* 10–20 years after active Lyme disease. *Clin. Infect. Dis.* **33**, 780–785 (2001).
11. Feder, H. M., Gerber, M. A., Luger, S. W. & Ryan, R. W. Persistence of serum antibodies to *Borrelia burgdorferi* in patients treated for Lyme disease. *Clin. Infect. Dis.* **15**, 788–793 (1992).
12. Grążlewska, W. & Holec-Gąsior, L. Antibody cross-reactivity in serodiagnosis of Lyme disease. *Antibodies* **12**. (2023).
13. Wojciechowska-Koszko, I. et al. Cross-reactive results in serological tests for borreliosis in patients with active viral infections. *Pathogens* **11**. (2022).
14. Zhang, H. et al. Integrative system biology strategies for disease biomarker discovery. *Comb. Chem. High. Throughput Screen.* **15**, 286–298 (2012).
15. Engström, M. T. et al. Rapid qualitative and quantitative analyses of proanthocyanidin oligomers and polymers by UPLC-MS/MS. *J. Agric. Food Chem.* **62**, 3390–3399 (2014).
16. García-Cañaveras, J. C. et al. LC-MS untargeted metabolomic analysis of drug-induced hepatotoxicity in HepG2 cells. *Electrophoresis* **36**, 2294–2302 (2015).
17. Hu, Q. et al. The orbitrap: A new mass spectrometer. *J. Mass Spectrom.* **40**, 430–443 (2005).
18. Ntasi, G., Tsarbopoulos, A., Mikros, E. & Gikas, E. Targeted metabolomics: The LC-MS/MS based quantification of the metabolites involved in the methylation biochemical pathways. *Metabolites* **11**, 416 (2021).
19. Martín-Blázquez, A. et al. Untargeted LC-HRMS-based metabolomics to identify novel biomarkers of metastatic colorectal cancer. *Sci. Rep.* **9**, 20198 (2019).
20. Rodgers, M. A., Saghatelian, A. & Yang, P. L. Identification of an overabundant cholesterol precursor in hepatitis B virus replicating cells by untargeted lipid metabolite profiling. *J. Am. Chem. Soc.* **131**, 5030–5031 (2009).
21. Kerstholt, M. et al. Role of glutathione metabolism in host defense against *Borrelia burgdorferi* infection. *Proc. Natl. Acad. Sci.* **115**. (2018).
22. Molins, C. R. & e *t al.* Metabolic differentiation of early Lyme disease from Southern tick–associated rash illness (STARI). *Sci. Transl. Med.* **9**. (2017).
23. Molins, C. R. et al. Development of a metabolic biosignature for detection of early Lyme disease. *Clin. Infect. Dis.* **60**, 1767–1775. (2015).
24. Pegalajar-Jurado, A. et al. Identification of urine metabolites as biomarkers of early Lyme disease. *Sci. Rep.* **8**, 12204 (2018).
25. Fitzgerald, B. L. et al. Metabolic response in patients with post-treatment Lyme disease symptoms/syndrome. *Clin. Infect. Dis.* **73**. (2021).
26. Fitzgerald, B. L. et al. Host metabolic response in early Lyme disease. *J. Proteome Res.* **19**, 610–623 (2020).
27. Kortela, E. et al. Oral doxycycline compared to intravenous ceftriaxone in the treatment of Lyme neuroborreliosis: A multicenter, equivalence, randomized, open-label trial. *Clin. Infect. Dis.* **72**, 1323–1331 (2021).
28. Schymanski, E. L. et al. Identifying small molecules via high resolution mass spectrometry: Communicating confidence. *Environ. Sci. Technol.* **48**, 2097–2098 (2014).
29. Sumner, L. W. et al. Proposed minimum reporting standards for chemical analysis. *Metabolomics* **3**, 211–221 (2007).
30. Carini, M., Aldini, G. & Orioli, M. Maffei Facino, R. *In vitro* metabolism of a nitroderivative of acetylsalicylic acid (NCX4016) by rat liver: LC and LC–MS studies. *J. Pharm. Biomed. Anal.* **29**, 1061–1071 (2002).
31. Kerns, E. H., Rourick, R. A., Volk, K. J. & Lee, M. S. Buspirone metabolite structure profile using a standard liquid chromatographic-mass spectrometric protocol. *J. Chromatogr. B Biomed. Sci. Appl.* **698**, 133–145 (1997).
32. David, A. et al. Acetaminophen metabolism revisited using non-targeted analyses: Implications for human biomonitoring. *Environ. Int.* **149**, 106388 (2021).
33. Modick, H. et al. Rapid determination of N-acetyl-4-aminophenol (paracetamol) in urine by tandem mass spectrometry coupled with on-line clean-up by two dimensional turbulent flow/reversed phase liquid chromatography. *J. Chromatogr. B.* **925**, 33–39 (2013).
34. Rousar, T. Cysteine conjugates of acetaminophen and p-aminophenol are potent inducers of cellular impairment in human proximal tubular kidney HK-2 cells. *Arch. Toxicol.* **97**, 2943–2954 (2023).
35. Cooper, A. J. L. Enzymology of cysteine S-conjugate β-lyases. *Adv. Pharmacol.* **27**, 71–113. https://doi.org/10.1016/S1054-3589(08)61030-3 (1994).
36. Keski-Hynnilä, H. & e Comparison of electrospray, atmospheric pressure chemical ionization, and atmospheric pressure photoionization in the identification of apomorphine, dobutamine, and entacapone phase II metabolites in biological samples. *Anal. Chem.* **74**, 3449–3457 (2002).
37. Yamaguchi, J. I. et al. Identification of rat urinary and biliary metabolites of esonarimod, a novel antirheumatic drug, using liquid chromatography/electrospray ionization tandem mass spectrometry with post column addition of 2-(2-methoxyethoxy)ethanol, a signal-enhancing modifi. *Drug Metab. Dispos.* **29**, 806–812 (2001).
38. Mrochek, J. E., Katz, S., Christie, W. H. & Dinsmore, S. R. Acetaminophen metabolism in man, as determined by high-resolution liquid chromatography. *Clin. Chem.* **20**, 1086–1096 (1974).
39. Kępka, A., Pancewicz, S. A., Janas, R. M. & Świerzbińska, R. Serum carnitine concentration is decreased in patients with Lyme borreliosis. *Postepy. Hig. Med. Dosw.* **70**, 180–185 (2016).
40. Cristofano, A. et al. Serum levels of acyl-carnitines along the continuum from normal to Alzheimer's dementia. *PLos One.* **11**, e0155694 (2016).
41. Liu, P. et al. Altered arginine metabolism in Alzheimer's disease brains. *Neurobiol. Aging.* **35**, 1992–2003 (2014).
42. Luiking, Y. C., Poeze, M., Ramsay, G. & Deutz, N. E. P. The role of arginine in infection and sepsis. *J. Parenter. Enter. Nutr.* **29**, (2005).
43. Sylvestre, D. A., Slupsky, C. M., Aviv, R. I., Swardfager, W. & Taha, A. Y. Untargeted metabolomic analysis of plasma from relapsing-remitting multiple sclerosis patients reveals changes in metabolites associated with structural changes in brain. *Brain Res.* **1732**, 146589 (2020).
44. Chung, K. P. et al. Increased plasma acetylcarnitine in sepsis is associated with multiple organ dysfunction and mortality: A multicenter cohort study. *Crit. Care Med.* **47**, 210–218 (2019).
45. Karetnikova, E. S. et al. Is homoarginine a protective cardiovascular risk factor? *Arterioscler. Thromb. Vasc Biol.* **39**, 869–875 (2019).
46. Sobczak, A. et al. Tobacco smoking decreases plasma concentration of the emerging cardiovascular risk marker, L-homoarginine. *Circ. J.* **78**, 1254–1258 (2014).
47. Lin, C. W. & Fishman, W. H. L-homoarginine: An organ-specific, uncompetitive inhibitor of human liver and bone alkaline phosphohydrolases. *J. Biol. Chem.* **247**, 3082–3087 (1972).
48. Virmani, A. Role of carnitine esters in brain neuropathology. *Mol. Aspects Med.* **25**, 533–549 (2004).
49. Nasca, C. et al. Acetyl-L-carnitine deficiency in patients with major depressive disorder. *Proc. Natl. Acad. Sci.* **115**, 8627–8632 (2018).
50. Yan, X. et al. Mass spectral library of acylcarnitines derived from human urine. *Anal. Chem.* **92**, 6521–6528 (2020).

51. Chafai, A., Fromm, M. F., König, J. & Maas, R. The prognostic biomarker L-homoarginine is a substrate of the cationic amino acid transporters CAT1, CAT2A and CAT2B. *Sci. Rep.* **7**, 4767 (2017).
52. Martens-Lobenhoffer, J., Surdacki, A. & Bode-Böger, S. M. Fast and precise quantification of L-homoarginine in human plasma by HILIC-isotope dilution-MS–MS. *Chromatographia* **76**, 1755–1759 (2013).
53. Humen Metabolome Database. Glutamyltryptophan. https://hmdb.ca/metabolites/HMDB0028830
54. Bojkowski, C. J., Arendt, J., Shih, M. C. & Markey, S. P. Melatonin secretion in humans assessed by measuring its metabolite, 6-sulfatoxymelatonin. *Clin. Chem.* **33**, 1343–1348 (1987).
55. Human Metabolome Database. Pipericine. https://hmdb.ca/metabolites/HMDB0031678
56. Chemspider L-Leucyl-L-alanyl-L-aspartic acid (Leu-Ala-Asp). http://www.chemspider.com/Chemical-Structure.28639354.html?Rid=887e1d94-2e1e-4825-b5c8-007504665d12&page_num=0

Acknowledgements

Graphical abstract and integrated workflow figure were created with BioRender®. The authors thank all members of Natural Chemistry Research Group for their kind help.

Author contributions

IK: Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Funding acquisition Visualization. AP: Conceptualization, Resources, Writing - Review & Editing, Funding acquisition. TR: Methodology, Software, Formal analysis, Data Curation, Writing - Review & Editing, Visualization. SH: Methodology, Writing - Review & Editing, Supervision. EK: Resources, Writing - Review & Editing. MJK: Resources, Writing - Review & Editing. JO: Resources, Writing - Review & Editing. JH: Conceptualization, Resources, Writing - Review & Editing, Supervision, Project administration, Funding acquisition. MK: Conceptualization, Methodology, Validation, Resources, Writing - Review & Editing, Supervision, Project administration, Funding acquisition. All authors contributed to the intellectual content and to writing, reviewing and editing of the manuscript and approved the final manuscript.

Funding

The research was funded by the University of Turku, and by two grants from Sakari Alhopuro Foundation to AP (20230181 and 20200177) and one grant from the Turku University Foundation to IK (081451).

Declarations

Competing interests

The authors declare no competing interests.

Institutional review board information

All individuals participating in the original study provided informed consent, and ethical approval was granted by the National Committee on Medical Research Ethics in Finland. Permission to use human serum samples was granted by the Hospital district of Southwest Finland.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-92189-0.

Correspondence and requests for materials should be addressed to I.K.