www.jpnim.com Open Access eISSN: 2281-0692
Journal of Pediatric and Neonatal Individualized Medicine 2025;14(2):e140202
doi: 10.7363/140202
Received: 2025 Jan 24; revised: 2025 Apr 23; accepted: 2025 Apr 29; published online: 2025 Sept 03

Original article

Urinary GC-MS and ^{1}H-NMR metabolomics of Sardinian cystic fibrosis patients reveal unique mutation-class dependent signatures: preliminary results

Martina Spada[1]*, Antonio Noto[1]*, Cristina Piras[1], Giulia Diana[1], Giulia Paci[1], Valentina Vincis[2], Karolina Krystyna Kopeć[3], Michele Mussap[4], Vassilios Fanos[2], Luigi Atzori[1], Maurizio Zanda[5]****

[1]Department of Biomedical Sciences, University of Cagliari, Cittadella Universitaria, Monserrato, Cagliari, Italy

[2]Department of Surgical Sciences, University of Cagliari, and Neonatal Intensive Care Unit, AOU Cagliari, Cagliari, Italy

[3]Department of Mechanical, Chemical and Materials Engineering, University of Cagliari, Cagliari, Italy

[4]IRCCS San Camillo Hospital, Venice, Italy

[5]Cystic Fibrosis Center, Azienda Ospedaliera Brotzu, Cagliari, Italy

*These Authors contributed equally to this work and share first authorship.

**These Authors contributed equally to this work and share last authorship.

Abstract

Cystic fibrosis (CF) is an autosomal recessive disorder caused by mutations in the CF transmembrane conductance regulator (*CFTR*) gene. The CFTR protein acts as an ion channel, and its deficiency results in an increased density and viscosity of secretion. CF shows high phenotypic variability because of the intervention of genetic and environmental factors. In this context, metabolomics is a useful tool to identify clinical biomarkers. Gas chromatography mass spectrometry (GC-MS) and ^{1}H-nuclear magnetic resonance spectroscopy (^{1}H-NMR) were used to study urine samples of 35 patients affected by CF with different genotypes (F508del/F508del, T338I/T338I, and F508del/T338I). The multivariate statistical analysis allowed the separation of the samples based on the metabolomics profile. A good separation between F508del/F508del vs. T338I/T338I genotypes, and F508del/F508del vs. F508del/T338I genotypes were observed. Moreover, the comparison between the two groups T338I/T338I vs. F508del/T338I did not highlight significant differences. The variables of importance responsible for the separation were sugars, organic acids, amino acids, and polyols. Metabolomic analysis has proven to be a useful tool to discriminate among the different subclasses of CF, mirroring the complexity of the pathological

condition. The present study represents a starting point for better understanding physiopathological changes and identifying new clinical biomarkers.

Keywords

Cystic fibrosis, gas chromatography mass spectrometry, ^{1}H-nuclear magnetic resonance spectroscopy, multivariate statistical analysis, clinical biomarkers.

Corresponding author

Cristina Piras, Department of Biomedical Sciences, University of Cagliari, Cittadella Universitaria, Monserrato, Cagliari, Italy; email: cristina.piras@unica.it.

How to cite

Spada M, Noto A, Piras C, Diana G, Paci G, Vincis V, Kopeć KK, Mussap M, Fanos V, Atzori L, Zanda M. Urinary GC-MS and ^{1}H-NMR metabolomics of Sardinian cystic fibrosis patients reveal unique mutation-class dependent signatures: preliminary results. J Pediatr Neonat Individual Med. 2025;14(2):e140202. doi: 10.7363/140202.

Introduction

Cystic fibrosis (CF) is a genetic disease inherited with an autosomal recessive mechanism caused by the mutation of the CF transmembrane conductance regulator (*CFTR*) gene, which codes for the homonymous protein, whose primary function is to act as a chloride channel [1]. The dysfunction of this protein is responsible for the increased density and viscosity of the mucous secretions of the secretory epithelia. CF is a disease that may present a chronic and evolutionary course with multi-organ involvement and poor prognosis [2]. Moreover, the phenotype of the disease can go from severe to mild or even asymptomatic forms. The classic manifestation is represented by chronic progressive bronchopneumopathy, exocrine pancreatic insufficiency, and elevated levels of electrolytes (Cl-, Na+) in sweat; in addition, a possible involvement can affect the liver, biliary tract, intestine, paranasal sinuses, endocrine pancreas, and vas deferens [3-7]. In Sardinia, the T338I and the F508del are the most frequent mutations, and the latter, in homozygous, is responsible for severe clinical pictures involving lung and pancreas functions [8]. On the other hand, patients with the T338I mutation, both homozygous and heterozygous (even with severe mutation F508del), present a mild clinical picture, with absent lung involvement and normal pancreatic function [1, 9, 10]. A recent approach, named metabolomics, has been applied to investigate the phenotype of CF at the molecular level [11, 12]. This method may identify specific metabolites and their relative variations in CF patients, providing further insights into the pathophysiology. In this scenario, there has been growing interest in applying metabolomics to characterize the metabolic phenotype of CF [13-16]. To date, metabolomic studies on the understanding of the basic pathophysiology of patients with CF are rather limited. Wetmore et al. in 2010 conducted an untargeted metabolomics analysis of cultured primary human airway epithelial cells from three separate cohorts of patients with CF compared to a group of disease-free subjects to understand the epithelial dysfunction caused by CF [15]. Impairment of more than 100 metabolites in CF patients was associated with decreased purine biosynthesis, increased tryptophan catabolism, decreased glutathione biosynthesis, and low levels of glucose metabolism mainly due to increased cellular sensitivity to oxidative stress. Joseloff et al. in 2014 conducted a metabolomic study of the serum profile of children with CF compared to subjects with non-CF lung disease [16]. The authors identify 92 altered metabolites in CF patients, mostly involved in lipid metabolism, oxidants, and markers consistent with abnormalities in bile acid processing. Furthermore, some identified metabolites were of bacterial origin, indicating intestinal dysbiosis in children with CF compared with non-CF. Microbial dysbiosis is a hallmark of CF and refers to the imbalance in the composition, diversity, and function of the microbial communities, particularly in the lungs and gastrointestinal tract. This dysbiosis may play a crucial role in disease progression, inflammation, and metabolic alterations in CF patients [17]. The present study evaluated patients with CF carrying the two most frequent mutations found in the Sardinian population. Specifically, the urinary metabolome of patients with T338I homozygous, F508del homozygous, and F508del/T338I heterozygous genotypes was analyzed, with the aim of comparing the metabolic phenotypes associated with these three different genotypes. This approach may help identify unique metabolic characteristics and specific differences that could serve as potential biomarkers for diagnosis and disease stratification.

Materials and methods

Patients

The study involved the ARNAS "G. Brotzu" Sardinian Regional Center for CF treatment, the Unit of Neonatal Intensive Care of the AOU of the Department of Surgery of the University of Cagliari, and the Unit of Clinical Metabolomics of the Department of Biomedical Sciences of the same University. The cohort included 35 patients (22 males, 13 females, aged between 1 and 54 years; mean age 21 ± 14) divided into three groups, based on genotype. The three populations consisted of F508del homozygous, T338I homozygous, and F508del/T338I heterozygous patients (**Fig. 1** and **Tab. 1**).

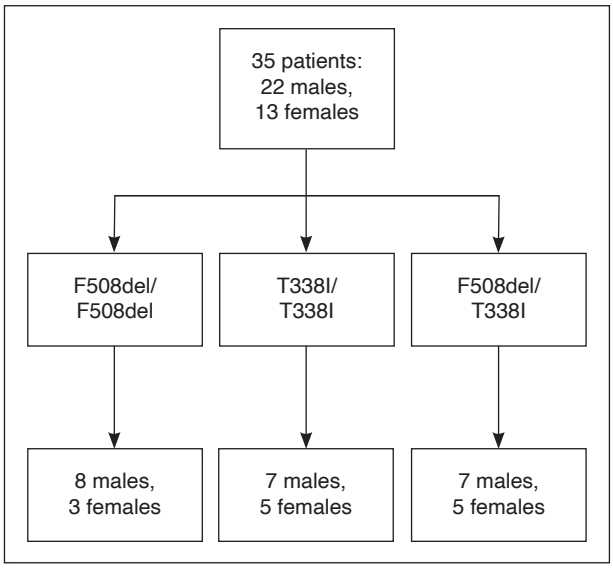

Figure 1. Patient flowchart of the study.

Table 1. Demographic and clinical features of cystic fibrosis (CF) patients.

Genotype	Gender	Age (years)	Weight (kg)	Height (meters)	Pancreatic insufficiency	Pulmonary involvement	FEV1 (%)
F508del/F508del	F	13	34.7	1.45	yes	yes	60
	F	18	59	1.7	yes	yes	64
	F	1	3.88	0.56	yes	no	-
	M	31	55	1.63	yes	yes	40
	M	23	57	1.67	yes	yes	100
	M	16	44.8	1.6	yes	yes	103
	M	25	65	1.72	yes	yes	67
	M	45	72	1.7	yes	yes	42
	M	36	76	1.78	yes	yes	92
	M	16	62	1.67	yes	yes	116
	M	32	70	1.75	yes	yes	44
T338I/T338I	F	29	43	1.63	no	no	71
	M	4	9.9	0.81	no	no	-
	M	22	83	1.71	no	no	101
	F	27	43	1.54	no	no	107
	M	29	63	1.78	no	no	118
	M	47	80	1.8	no	no	99
	M	16	52	1.71	no	no	100
	F	17	48.8	1.54	no	no	110
	F	5	17.5	1.06	no	no	-
	F	7	16.6	1.05	no	no	74
	M	14	42	1.51	no	no	105
	M	54	77	1.6	no	no	118
F508del/T338I	F	9	28.4	1.3	no	no	81
	F	3	9.5	0.79	no	no	-
	F	10	35	1.38	no	no	93
	M	25	47.5	1.58	no	no	80
	F	19	54.5	1.65	no	no	109
	M	11	41	1.39	no	no	88
	M	4	15.2	0.99	no	no	-
	M	32	81.5	1.77	no	no	89
	M	30	62	1.78	no	no	86
	F	16	51.7	1.52	no	no	77
	M	46	82	1.8	no	no	85
	M	8	31	1.27	no	no	109

FEV1: forced expiratory volume in the 1st second.

Pancreatic insufficiency is very frequent in CF patients, and it's always present in severe forms like F508del homozygous; fecal elastase is the parameter used to evaluate pancreatic function; values lower than 200 mcg/g indicate pancreatic insufficiency. On the contrary, pancreatic insufficiency is not present in mild forms, particularly in those with T338I mutation.

Pulmonary involvement is almost always present in severe forms (F508del homozygous), while its modest or even absent in mild one; it's characterized by cough, recurrent infections, chronic *Pseudomonas aeruginosa* or *Staphylococcus aureus* infection, reduced respiratory function; FEV1 (forced expiratory volume in the 1st second) is the main parameter of respiratory function and it's the parameter used in CF patients to monitor the progress of the disease.

Inclusion criteria were patients with CF, of both genders, of any age group, with the F508del/F508del, T338I/T338I, and F508del/T338I genotypes. Exclusion criteria were patients not affected by CF or affected by CF but with different genotypes than those mentioned above.

For each participant, a first-morning urine sample was collected at the day hospital, after informed consent, and stored in a freezer at -20°C before being transferred to the Clinical Metabolomics Laboratory of the Department of Biomedical Sciences of the University of Cagliari. About 1 mL of each urine sample was transferred to a sterile test tube containing a 1% concentration of sodium azide solution (acting as a bacteriostatic) and stored in a freezer at -80°C. The samples were subsequently prepared for gas chromatography mass spectrometry (GC-MS) and ^{1}H-nuclear magnetic resonance (^{1}H-NMR) spectroscopy analysis according to our internal protocol [18].

GC-MS analysis

150 µL of urine was transferred to a 2 mL Eppendorf tube with 200 µL of an aqueous urease solution (1 mg/mL) and subjected to ultrasound for 30 min. Then 800 µL of methanol was added to denature the enzyme. After centrifugation, 750 µL of the supernatant was taken, transferred into glass test tubes, and evaporated to dryness in an Eppendorf vacuum centrifuge. The dried samples were then derivatized with 50 µL of a solution of methoxamine in pyridine (10 mg/mL; Sigma-Aldrich). After 1 hour at 70°C, 50 µL of N-methyl-N-(trimethylsilyl) trifluoroacetamide (Sigma-Aldrich) was added, and the mixture was left to react at room temperature for 1 hour. Then, samples were diluted with 600 µL of anhydrous hexane containing undecane as an internal standard. One µL of each sample was injected splitless into a 7890A gas chromatograph coupled with a 5975C mass spectrometer (Agilent Technologies, Santa Clara, CA, USA) equipped with a 30 m × 0.25 mm ID, fused silica capillary column, with a 0.25 µM TG-5MS stationary phase (Thermo Fisher Scientific, Waltham, MA, USA). The injector temperature was 250°C, the gas flow through the column was 1 mL/min, and the transfer line temperature was 280°C. The column's initial temperature was kept at 70°C for 3 min, then increased to 250°C at 12°C/min and held for 4 min. Finally, the temperature was increased to 300°C at 50°C/min and kept for 1 min. Identification of metabolites was performed using the standard NIST 08 and GMD mass spectra libraries and, when available, by comparison with authentic standards [19]. Data processing was performed using MassHunter Software (Agilent Technologies). The total area of chromatograms (= 100) was used to normalize the measurement of the metabolites from each urine sample.

^{1}H-NMR analysis

Before analysis, samples were centrifuged for 10 min at 4°C at 12,000× g to remove solid particles. Then, 630 µL of the supernatant were mixed with 70 µL of potassium phosphate buffer in D2O (1.5 M, pH 7.4) containing sodium 3-trimethylsilyl-propionate-2,2,3,3,-d4 (TSP) as an internal standard (98 atom% D, Sigma-Aldrich, Milan, Italy). Finally, 650 µL were transferred to 5 mm NMR glass tubes for ^{1}H-NMR analysis.

^{1}H-NMR analysis was carried out using a Varian UNITY INOVA 500 spectrometer operating at 499.839 MHz for proton and equipped with a 5 mm double resonance probe (Agilent Technologies, Santa Clara, CA, USA). One-dimensional proton NMR spectra were obtained by using a 1D nuclear overhauser enhancement spectroscopy (NOESY) standard pulse sequence to suppress water signals with a relaxation delay of 3 s. For each sample, 256 free induction decays (FIDs) were collected into 64K data points with a spectral width of 6,000 Hz spectral with a 90° pulse, an acquisition time of 2 s, and a mixing time of 100 ms. The FIDs were weighted by an exponential function with a 0.5 Hz line-broadening factor prior to Fourier transformation. NMR spectra were phased and baseline corrected using an Advanced Chemistry Development (ACD) lab (Toronto, ON, Canada) Processor Academic Edition (Advanced Chemistry Development, 1

December 2010) and chemical shifts referenced internally to trisodium phosphate (TSP) at $\delta = 0.0$ ppm. The spectral region comprising the signal of residual water and urea (4.7-6.5 ppm) was removed. The final spectral regions were between 0.5-4.7 ppm and 6.5-9.5 ppm. The metabolites were identified and quantified using the Chenomx NMR Suite 7.1 (Chenomx Inc., Edmonton, AB, Canada), an integrated set of tools for identifying and quantifying metabolites in NMR spectra [20]. Chenomx NMR Suite is equipped with reference libraries containing numerous pH-sensitive compound models that are identical to the spectra of pure compounds obtained under similar experimental conditions. GraphPad Prism software (version 7.01, GraphPad Software, Inc., San Diego, CA, USA) was used to perform the univariate statistical analysis. Statistical significance was assessed by using the Mann-Whitney U test; $p \leq 0.05$ was considered statistically significant.

Multivariate statistical analysis

The multivariate statistical analysis was used to get the most information from complex spectrometric data. Different procedures were used: the principal component analysis (PCA), the partial least-square discriminant analysis (PLS-DA), and the orthogonal-projection on latent structures discriminant analysis (OPLS-DA). PCA is a data simplification technique that can summarize groups of multivariable data into two main components, PC1 and PC2. PC1 is a linear combination of the original starting variables and describes the significant variance in the data set; PC2, on the other hand, is a parameter that describes the degree of variance in the data set and is orthogonal to PC1. Graphically, the output from the PCA analysis consists of scores plot, indicating any grouping in the data sets in terms of metabolomic similarity, and loadings plot, indicating which variables are important for the patterns obtained in the scores plot. The PLS-DA uses a Y-matrix that contains the information of the class to which the sample belongs, and therefore, it is widely used for the classification of the sample. The PLS-DA is generally associated with the OPLS-DA, which maximizes the covariance between the measured data of the X-variable (peaks area of GC-MS chromatograms) and the response of the Y-variable (class assignment) within the groups. The goodness of the model was evaluated using a 7-fold cross-validation and "permutation test" (500 times). The permutation test was calculated by randomizing the Y-matrix while the X-matrix was kept constant. The permutation plot then displays the correlation coefficient between the original Y-variable and the permuted Y-variable on the X-axis versus the cumulative R^2 and Q^2 on the Y-axis and draws the regression line. The intercept is a measure of the overfit, Q^2 intercept value less than 0.05 is indicative of a valid model. The most significant variables were extracted from each model's plot of loadings (metabolites) to which univariate statistical analysis was applied. GraphPad Prism software (version 7.01, GraphPad Software, Inc., San Diego, CA, USA) was used to perform the univariate statistical analysis of the data. Statistical significance was assessed by using the Mann-Whitney U test; $p \leq 0.05$ was considered statistically significant. The Benjamini-Hochberg adjustment was subsequently applied to p-values, to acquire the level of significance for multiple testing. The relative concentrations of the metabolites in the different groups were compared using box-and-whisker plots.

Correlation-based network analysis

Two different software were used to create the network analysis: Dave (Data Analysis and Visualization Engine) and Cytoscape (version 3.9.1). Using Dave, an exploratory data analysis to gain insights into the metabolomic dataset was conducted. This step involved statistical analysis and data visualization to identify differentially expressed metabolites between groups. Dave's interactive visualizations can help to identify patterns and potential biomarkers associated with the disease. Next, we used the results from the data analysis to identify specific metabolite pathways that are significantly altered in patients with CF. The identified metabolite pathway data were imported into Cytoscape to construct the network representation. Each metabolite was represented as a node, and the biochemical interactions or relationships between metabolites were represented as edges. Different colours, shapes, and labels were used to denote different types of metabolites and their properties. Cytoscape offers various layout algorithms to arrange the nodes and edges in a visually appealing and informative way. After selecting an appropriate layout, the appearance of the pathway diagram was fine-tuned to highlight the most relevant information. This included differentiating upregulated and downregulated metabolites, highlighting key pathways or nodes, and providing annotations for essential metabolites.

Results

Thirty-five patients underwent metabolomics analysis, of which 11 belonged to the F508del/F508del population, 12 to the T338I/T338I population, and 12 to the F508del/T338I population.

The identified metabolites generated two data matrixes corresponding to the metabolites detected with GC/MS and ^{1}H-NMR. The results of the multivariate analysis applied to the GC-MS data of the samples are shown below. A first analysis of the spectral data was carried out through the application of PCAs, which did not show clusters or groupings nor the presence of strong outliers (data not shown). Subsequently, a supervised analysis was applied to verify a possible separation of the samples into three distinct classes according to the metabolomics profile (**Fig. 2**). The PLS-DA model was not statistically significant (R^2X = 0.336, R^2Y = 0.406, Q^2 = -0.0365), probably due to overlapping samples T338I/T338I (blue color) and F508del/T338I (empty circle), indicating a similar urinary profile between the two groups. The same result was highlighted by a PLS-DA model obtained with the ^{1}H-NMR approach. For this reason, a pairwise evaluation was carried out.

At first, the metabolomics profiles of patients with the F508del/F508del genotype and patients with the T338I/T338I genotype were compared (**Fig. 3**). The OPLS-DA model shows good separation between the F508del/F508del genotype and the T338I/T338I genotype, indicating a different metabolomic profile between the two groups of samples. Indeed, the OPLS-DA model (**Fig. 3**) was established with one predictive and one orthogonal component and showed good values of R^2X, R^2Y, and Q^2 (0.551, 0.882, 0.441, respectively). The validity of the OPLS-DA model was evaluated through a permutation test using 500 times. A value of Q^2 intercept = -0.332 indicates the statistical validity of the OPLS-DA model. The univariate statistical analysis was performed comparing the relative concentrations of metabolites obtained with GC-MS and ^{1}H-NMR analysis. The significantly altered metabolites between the two groups (p-value < 0.05) were: glucose, 4-hydroxyphenylacetic acid, citric acid, scyllo-inositol, 3-hydroxyisovaleric acid, maltose, pyruvic acid, isobutyric acid, tyrosine, 3-hydroxy-3-methylglutaric acid, and methylhistidine, as shown in **Fig. 4**. The relative concentration (± standard deviation) and respective p-value for each metabolite are reported in **Tab. 2**. Biochemical network mapping from comparing the urine metabolome of F508del/F508del and T338I/T338I patients showed a significant metabolic shift (**Fig. 5**). Based on biochemical and structural similarities and the number of interconnected metabolites, the resulting relationships indicated several metabolites strongly associated with each other.

Next, the metabolomics profile of patients with the F508del/F508del genotype and patients with the F508del/T338I genotype was

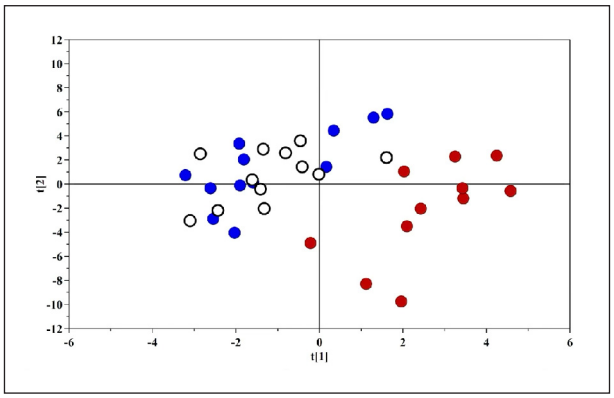

Figure 2. Partial least-square discriminant analysis (PLS-DA) score plot model obtained with multivariate statistical analysis conducted on urine samples acquired by gas chromatography mass spectrometry (GC-MS). The blue circles represent T338I/T338I samples, the red ones F508del/F508del samples and the empty circles the F508del/T338I samples. The model was not statistically significant.

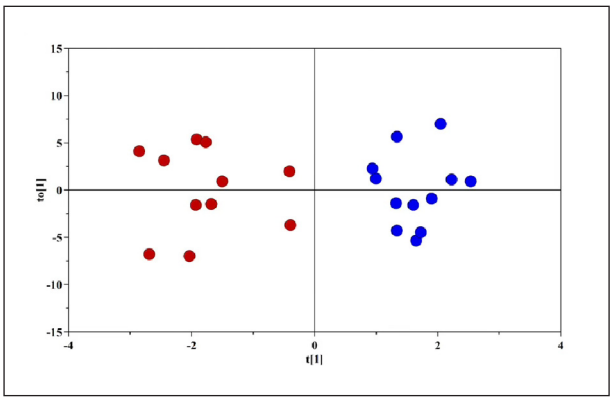

Figure 3. Orthogonal-projection on latent structures discriminant analysis (OPLS-DA) score plot model obtained with multivariate statistical analysis conducted on urine samples acquired by gas chromatography mass spectrometry (GC-MS): F508del/F508del (red color), T338I/T338I (blue color). Q^2 intercept value less than 0.05 is indicative of a valid model.

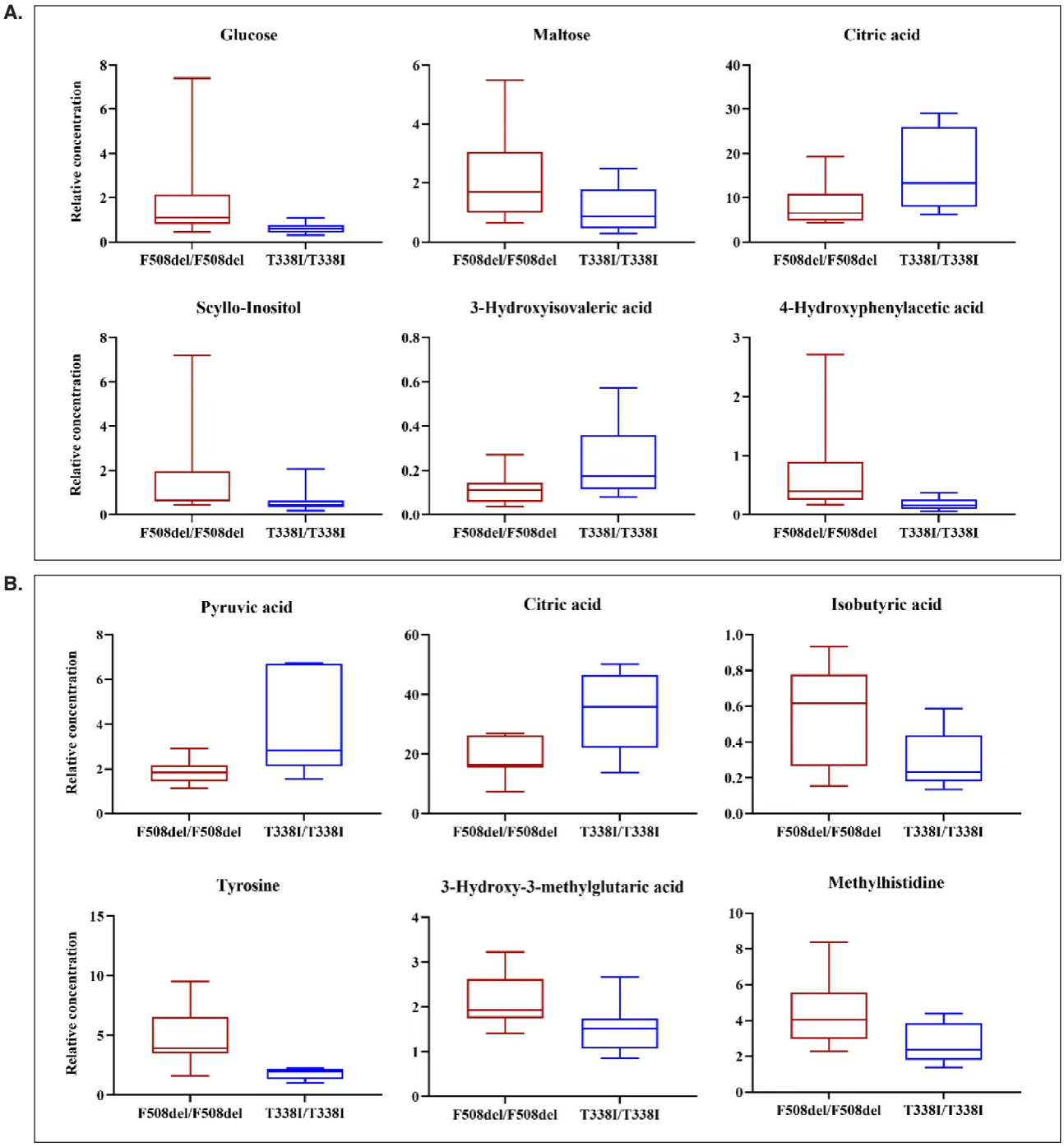

Figure 4. Altered metabolites between the F508del/F508del and T338I/T338I groups. The Mann-Whitney U test was used to verify the corresponding significance. The relative concentrations of the metabolites identified by **A**) gas chromatography mass spectrometry (GC-MS) and **B**) ^{1}H-nuclear magnetic resonance (^{1}H-NMR) analysis were shown using box-and-whisker plots.

compared (**Fig. 6**). An OPLS-DA model was built with one predictive and one orthogonal component. The analysis showed suitable statistical parameters (R^2X = 0.548, R^2Y = 0.683, Q^2 = 0.400). A permutation test using 500 times was performed to assess the validity of the statistical mode. A value of Q^2 intercept = -0.414 indicates the statistical validity of the OPLS-DA model. A Mann-Whitney U test revealed significant concentration differences of glucose, citric acid, 3-hydroxyisovaleric acid, 3-(3-hydroxyphenyl)-3-hydroxypropionic acid (HPHPA), pyruvic acid, 4-carboxyglutamic acid, tyrosine, 3-hydroxy-3-methylglutaric acid and methylhistidine between the two groups (**Fig. 7**). The relative concentration (± standard deviation) and respective p-value for each metabolite are reported in **Tab. 2**. A significant

Table 2. Relative concentration (± standard deviation) and respective p-value for discriminant metabolites among the F508del/F508del, T338I/T338I and F508del/T338I groups.

Metabolites	Mean ± SD F508del/F508del	Mean ± SD T338I/T338I	Mean ± SD F508del/T338I	p-value [a] F508del/F508del vs. T338I/T338I	p-value [a] F508del/F508del vs. F508del/T338I
GC-MS					
Glucose	1.901 ± 2.086	0.6232 ± 0.243	0.751 ± 0.399	**0.016**	**0.050**
Maltose	2.136 ± 1.444	1.107 ± 0.7109	-	**0.033**	-
Citric acid	8.488 ± 4.361	16.06 ± 8.932	20.91 ± 10.11	**0.041**	**0.016**
Scyllo-inositol	1.615 ± 2.013	0.633 ± 0.543	-	**0.025**	-
3-hydroxyisovaleric acid	0.117 ± 0.069	0.239 ± 0.155	0.243 ± 0.131	**0.050**	**0.033**
4-hydroxyphenylacetic acid	0.699 ± 0.738	0.188 ± 0.104	-	**0.008**	-
HPHPA	0.197 ± 0.209	-	0.8415 ± 0.882	-	**0.040**
^{1}H-NMR					
Pyruvic acid	1.862 ± 0.5122	3.954 ± 2.314	4.692 ± 3.104	**0.016**	**0.050**
Citric acid	20.64 ± 6.444	32.98 ± 13.5	38.62 ± 17.08	**0.033**	**0.042**
Isobutyric acid	0.5535 ± 0.268	0.297 ± 0.161	-	**0.050**	-
Tyrosine	4.683 ± 2.505	1.732 ± 0.496	2.426 ± 1.585	**0.008**	**0.032**
3-hydroxy-3-methylglutaric acid	2.152 ± 0.606	1.526 ± 0.6113	1.203 ± 0.4125	**0.025**	**0.008**
Methylhistidine	4.371 ± 1.954	2.645 ± 1.084	2.482 ± 0.9605	**0.041**	**0.025**
4-carboxyglutamic acid	22.57 ± 7.5	-	14.44 ± 4.825	-	**0.016**

^{1}H-NMR: ^{1}H-nuclear magnetic resonance; GC-MS: gas chromatography mass spectrometry; HPHPA: 3-(3-hydroxyphenyl)-3-hydroxy propionic acid; SD: standard deviation.
[a] p-value with Benjamini-Hochberg correction.
The dash indicates that the metabolite was not significantly altered in that group.

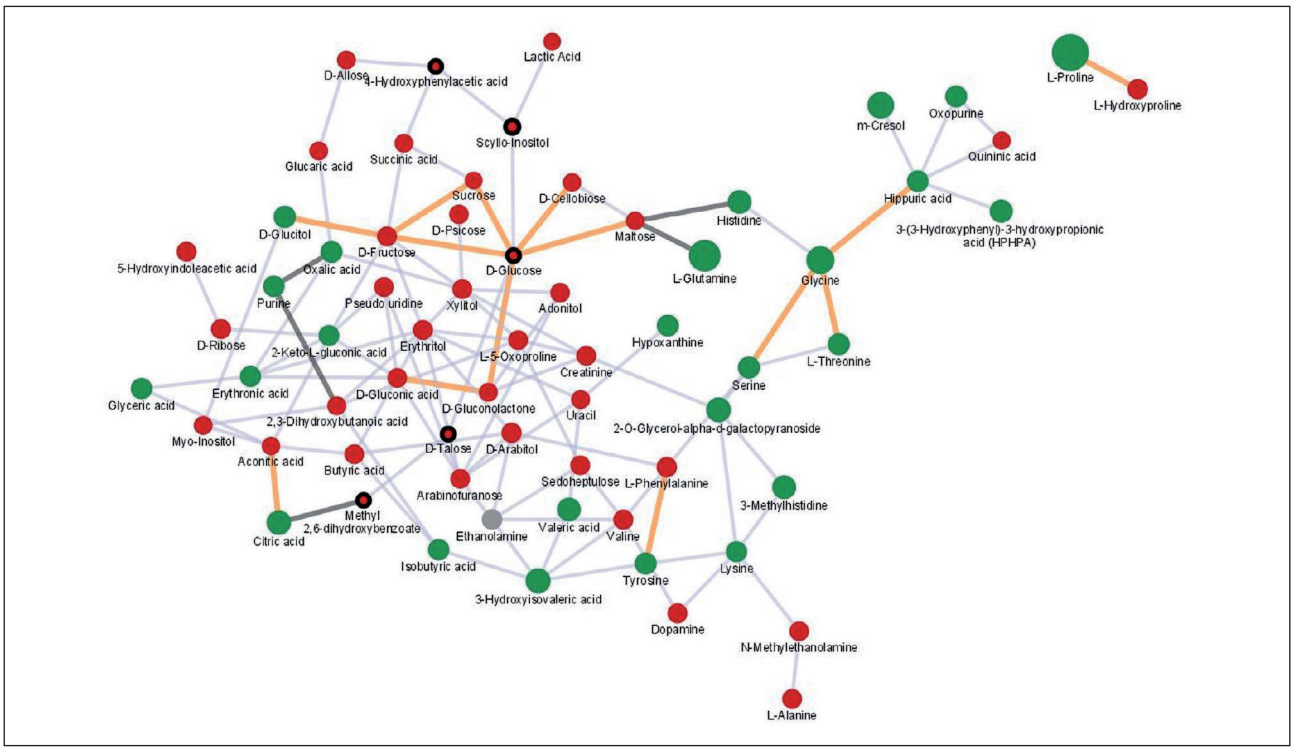

Figure 5. Biochemical similarity network displaying changes in urinary metabolites between homozygosity for F508del and homozygosity for T338I cohorts.
Nodes represent metabolites and display the direction of the fold change in homozygosity for T338I versus homozygosity for F508del. An orange line indicates a biochemical relationship. A pink line represents a positive correlation, while a gray line indicates a negative correlation. Metabolites shown in red are decreased, and those in green are increased. Thick black borders identify metabolites significantly altered.

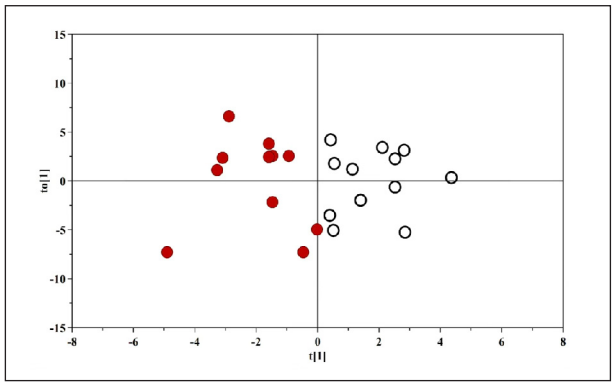

Figure 6. Orthogonal-projection on latent structures discriminant analysis (OPLS-DA) score plot model obtained with multivariate statistical analysis conducted on urine samples acquired by gas chromatography mass spectrometry (GC-MS): F508del/F508del (red color), F508del/T338I (empty circle). Q^2 intercept value less than 0.05 is indicative of a valid model.

metabolic shift between the urinary metabolome of F508del/F508del and F508del/T338I patients was highlighted by the biochemical network mapping. Several metabolites were strongly associated with each other, as shown in **Fig. 8**, which reported the relationships based on biochemical and structural similarities and the interconnections among metabolites.

Finally, the metabolomic profile of patients with the T338I/T338I genotype and patients with the F508del/T338I genotype was compared (**Fig. 9**). In this case, the OPLS-DA model was not statistically significant ($R^2X = 0.326$, $R^2Y = 0.580$, $Q^2 = 0.0736$). The statistical model indicated a substantial similarity in the metabolomics profile of the two groups.

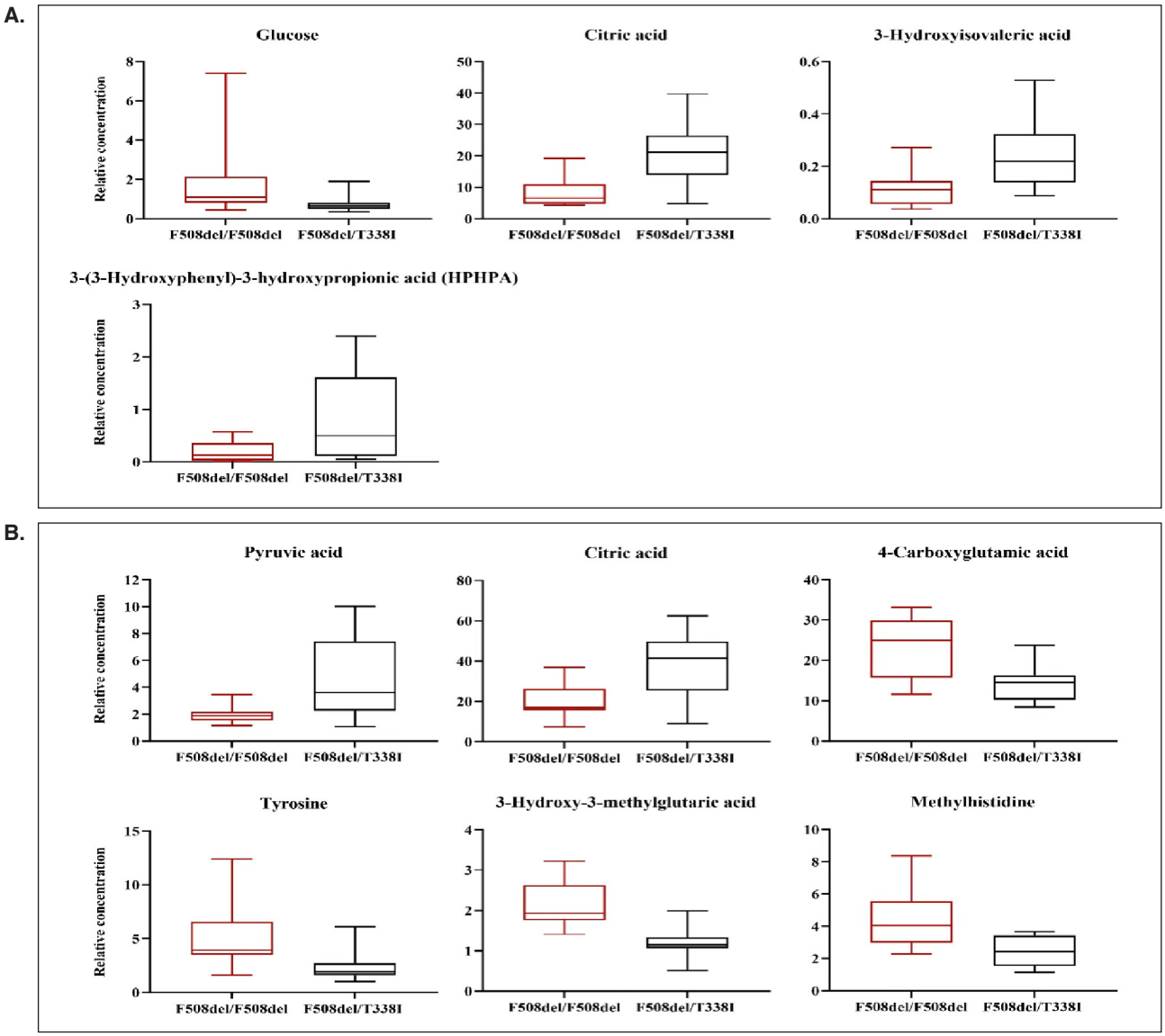

Figure 7. Altered metabolites between the F508del/F508del and F508del/T338I groups. The Mann-Whitney U test was used to verify the corresponding significance. The relative concentrations of the metabolites identified by **A**) gas chromatography mass spectrometry (GC-MS) and **B**) ^{1}H-nuclear magnetic resonance (^{1}H-NMR) analysis were showed using box-and-whisker plots.

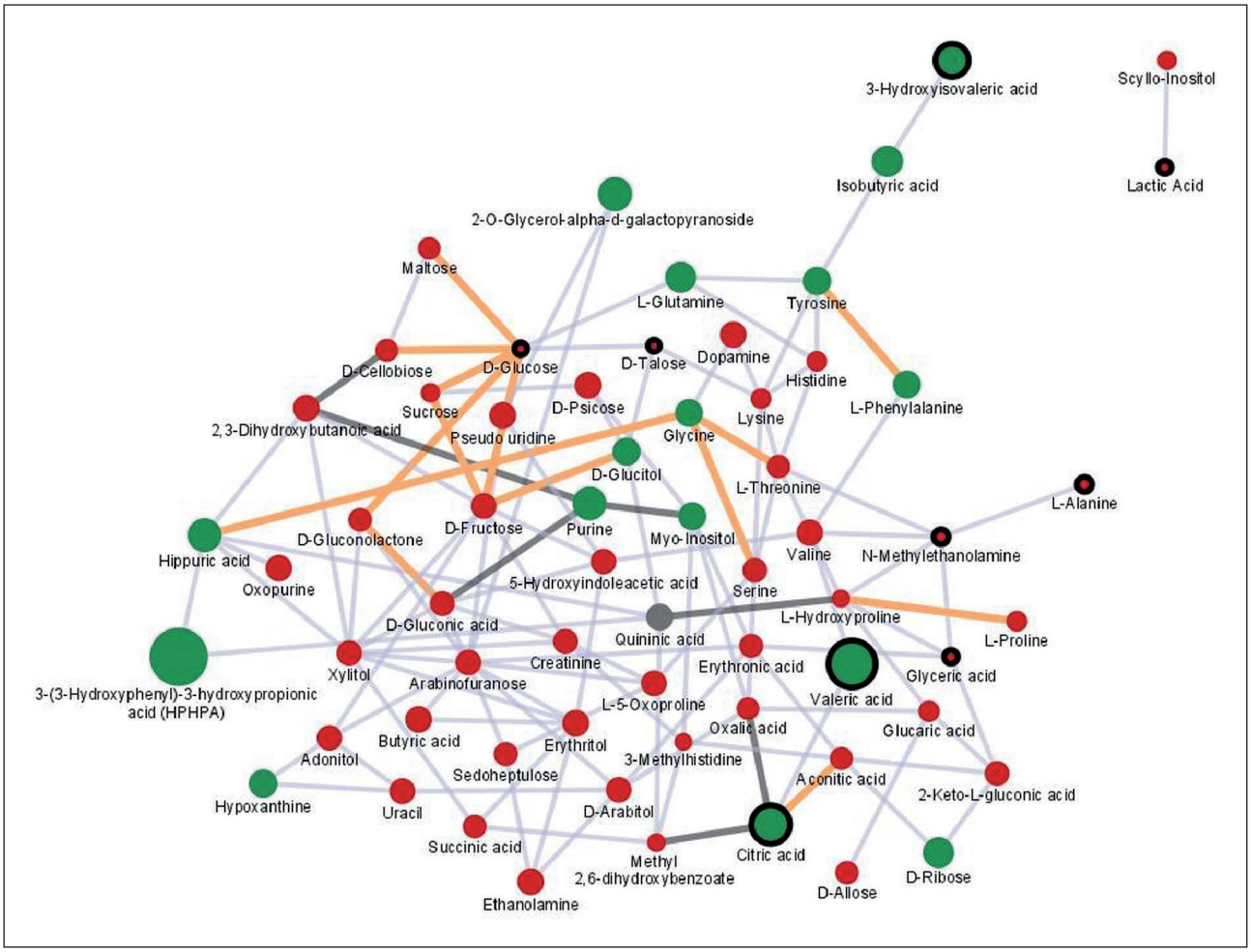

Figure 8. Biochemical similarity network displaying changes in urinary metabolites between homozygosity for F508del and heterozygosity for F508del/T338I cohorts.

Nodes represent metabolites and display the direction of the fold change in heterozygosity for F508del/T338I versus homozygosity for F508del. An orange line indicates a biochemical relationship. A pink line represents a positive correlation, while a gray line indicates a negative correlation. Metabolites shown in red are decreased, and those in green are increased. Thick black borders identify metabolites significantly altered.

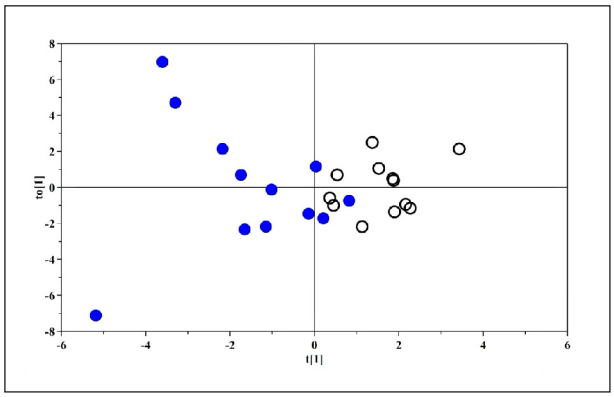

Figure 9. Orthogonal-projection on latent structures discriminant analysis (OPLS-DA) score plot model obtained with multivariate statistical analysis conducted on urine samples acquired by gas chromatography mass spectrometry (GC-MS) did not show a significant separation between the T338I/T338I (blue color) and F508del/T338I (empty circle) samples.

Discussion

The *CFTR* gene encodes a 1,480 amino acid long transmembrane protein with a symmetrical structure, it is a unique ATP-binding cassette (ABC) transporter that functions as a chloride (Cl-) and bicarbonate (HCO_3-) channel in epithelial cells. The properties of CFTR refer to how effectively it allows chloride and other anions to pass through the channel once it is open. These properties are crucial for normal physiological function and are directly linked to the pathology [21]. CF is caused by ~2,000 mutations in the *CFTR* gene with a wide range of disease severity, which can be reflected in complex intra- and inter-patient alterations. The phenotypic variability, observed not only in patients with different genotypes but also in patients with the same genotype, suggests

the presence of other factors capable of influencing the severity of the disease [9, 10]. For this reason, several studies have led to the discovery of gene modifiers (such as genes involved in the immune response or inflammation) and to the identification of epigenetic factors (such as tobacco, pollution, socioeconomic status, and adherence to therapies) which seem to have the ability to modulate the expression of the disease [22, 23].

Metabolomics, an emerging tool of systems medicine, allows the investigation of the molecular complexity of pathology by identifying the metabolic phenotypes' final product of the interaction between genetic and environmental factors. Therefore, it represents a valuable tool for 1) studying complex pathologies with variable expressivity, such as CF, and 2) bridging the gap between genotype and phenotype. Its integration into clinical practice holds promise for earlier diagnosis, more accurate disease monitoring, and personalized therapeutic strategies, ultimately improving patient outcomes.

The present study, conducted on patients with CF, showed a different metabolomic profile between the population with the F508del/F508del genotype and the populations with the T338I/T338I and F508del/T338I genotypes. The metabolomics profiles are consistent with the known clinical pictures typical of these populations under examination. Patients with the F508del/F508del genotype present a generally severe clinical picture, mainly characterized by progressive chronic bronchopneumopathy and exocrine pancreatic insufficiency. In contrast, patients with the T338I/T338I and F508del/T338I genotypes present a mild clinical picture, with no pulmonary involvement and normal pancreatic function [1, 2]. It is intriguing to note that the comparison between T338I/T338I and F508del/T338I shows a similar metabolic phenotypic profile, probably indicative of a possible adaptation thanks to the compensation given by the T338I mutation that affects the transmembrane domain 6 of the CFTR protein with the substitution of threonine with isoleucine in position 3. This mutation contributes to a partial alteration of the selectivity and permeability of the channel to ions, causing a reduced chloride conductance and altered permeability of bicarbonate, which causes a mild to moderate phenotype. In contrast, the F508del mutation causes a severe phenotype as it is essential for correct folding, trafficking, and channel gating.

The metabolomics profiles, highlighted by comparing patients with the F508del/F508del genotype to those with the T338I/T338I genotype and those with the F508del/T338I genotype, resulted in statistically significant altered metabolites. The significantly altered urinary metabolites identified a characteristic metabolomics fingerprint in the populations under investigation. This will be interpreted considering the information in the literature and by studying the network similarity among the urinary metabolites. In addition, representing the metabolism as a network may uncover unrelated but connected diseases' fingerprints. The untargeted network similarity approach allowed to evaluate the relationship among the detected metabolites for biological interpretation. Such relationships are formalized as networks, where the nodes correspond to the metabolites or features, and edges connect nodes if the corresponding metabolites are related. Interestingly, the comparison between the two networks corresponding to the two models (F508del/F508del vs. T338I/T338I and F508del/F508del vs. F508del/T338I) showed some diversity of correlations between the metabolites identified, reflecting the different metabolic hierarchies involved in the pathology.

As expected from the clinical phenotypes, glucose is a hub when comparing F508del/F508del vs. T338I/T338I and F508del/F508del vs. F508del/T338I genotypes. Glucose is related to other sugars, such as maltose, which strongly decreases in the mild form of CF (T338I). In contrast, in the severe form, these sugar metabolites increase in urine, due to the pancreatic involvement typical of patients with severe genotypes [24]. Pancreatic endocrine dysfunction, with consequently reduced insulin secretion, is responsible for the first phase of glucose intolerance and, subsequently, CF-related diabetes, both conditions that could justify the increase in urinary glucose [25, 26]. Another constituent is glutamine, which increased in both networks in patients affected with the T338I mutation. Glutamine is a precursor for gluconeogenesis, the process of glucose production from other non-carbohydrate constituents, which is a central metabolic pathway in the liver that allows the maintenance of blood glucose levels in fasting and starvation conditions following depletion of glycogen stores. Therefore, its increase may be interpreted as a homeostatic mechanism for glucose levels. In contrast, homozygous F508del/F508del displayed a reduced presence of glutamine, probably linked to its use for gluconeogenesis [27].

Another network-shared metabolite is scyllo-inositol, which appeared to be strongly increased in patients with the F508del/F508del genotype. It is a derivative of cyclohexane with six hydroxyl groups, and for this reason, it is also known as sugar alcohol. It plays an essential role in various cellular functions, including cell signalling, membrane integrity, and regulation of ion channels. While inositol has been studied in several disease conditions (such as intrauterine growth restriction) [28], its role in CF is still under investigation. The relationship between inositol, inositol derivatives, and enzymes involved in its metabolism concern its crucial role in cell signalling in various cellular processes, including the metabolism of glucose and lipids. Therefore, it could be intended as a marker of reduced pancreatic activity in the severe form of CF [29].

Moreover, our study indicated a reduced urinary level of pyruvic acid and citric acid in the homozygote F508del population. Pyruvic acid, an aliphatic monocarboxylate and glycolytic end-product, enters mitochondria via the inner membrane monocarboxylate transporter and is central to cellular energy production. Moreover, it incorporates antioxidant properties due to its α-keto-carboxylate structure, enabling it to neutralize peroxides and peroxynitrite directly [30]. Citric acid is a tricarboxylic acid, an intermediate metabolite of the Krebs cycle. Although the citric acid cycle plays a primary role in catabolism, anabolism, and energy production, it also supplies NADH and FADH2, important antioxidant agents. The reduced pyruvic acid and citric acid levels in the severe population (F508del homozygote) may indicate the presence of abnormal oxidative stress in CF [31]. As such, the role of oxidative stress in the progression of lung injury in severe homozygote F508del patients has been widely recognized and well described in the literature. It has been recently shown that, in patients with CF, there is an essential deficit of antioxidant molecules and an increase in oxidative stress [32]. The continuous imbalance between oxidant and antioxidant species led to chronic inflammation, which contributes to persistent cellular damage and prevents proper airway remodelling.

HPHPA resulted statistically increased in the heterozygote F508del/T338I population compared with the homozygote F508del population. It is a frequently detected organic acid in human urine, which has recently been reported as an abnormal phenylalanine metabolite resulting from bacterial metabolism in the gastrointestinal tract. HPHPA appears to derive from the action of intestinal bacteria of *Clostridia* species such as *C. difficile* [33-35]. *C. difficile* infection is often associated with antibiotics, the latter being frequently taken by patients with CF, making them patients at high risk of intestinal dysbiosis. The results confirm the known alteration of the intestinal microbiota of patients with CF [36].

The 4-hydroxyphenylacetic acid resulted statistically increased in the homozygote F508del population only in the comparison with the homozygote T338I population, while its precursor tyrosine resulted statistically increased in the homozygote F508del population both in the comparison with the homozygote T338I population and in the comparison with the heterozygote F508del/T338I population. Tyrosine is commonly used to screen small intestine diseases and bacterial overgrowth syndromes [37]. In particular, higher levels of 4-hydroxyphenylacetic acid have been associated with an overgrowth of small intestinal bacteria of *Clostridia* species, including *C. difficile*, *C. stricklandii*, *C. lituseburense*, *C. subterminale*, *C. putrefaciens* and *C. propionicum,* that is considered a dysbiotic profile [38]. Recent metabolomics studies conducted on fecal samples from pediatric and adult subjects have already demonstrated that the composition of the intestinal microbiota of CF subjects differs significantly from that of healthy controls [39-40].

A sign of nutrition deficiency in CF patients can be stated by measuring leucine and its cometabolites levels, such as 3-hydroxy-3-methylglutaric acid and 3-hydroxyisovaleric acid, which are the primary metabolites of the L-leucine pathway. The presence of these two metabolites with significant but opposite trends in the urine of the homozygote F508del population can be a sign of nutritional problems and muscle mass loss, commonly present in 25-30% of children with CF [41]. Notably, while life expectancy for severe patients has gradually improved during the last decades, nutrition is still a critical component for managing CF, and nutritional status is directly associated with both pulmonary involvement and survival. Therefore, an increased concentration of 3-hydroxy-3-methylglutaric acid and a reduced concentration of 3-hydroxyisovaleric acid in the severe population could be intended as indicators of the need for energy supplements [42].

Finally, urine methylhistidine, either 1-methylhistidine or 3-methylhistidine, is a potential bio-

marker for evaluating preclinical myopathy and muscle protein turnover. In severe CF patients, peripheral muscle dysfunction is an important systemic consequence of CF with significant clinical implications, such as exercise intolerance and reduced quality of life. Therefore, increased methylhistidine could be considered an indicator of muscle dysfunction [43].

Limitations and future direction

Although the findings of this study provide valuable insights and raise important hypotheses regarding metabolites alterations in CF, the small sample size of 35 subjects represents a significant limitation. This limited cohort reduces the statistical power of the analysis, making it difficult to draw definitive conclusions. Future research should aim to validate these findings in larger, more diverse cohorts, ideally through multicenter studies. Expanding the sample size will not only enhance the robustness of statistical analyses but also allow for subgroup comparisons and more detailed stratification based on clinical variables such as medication use.

Conclusions

Overall, this study creates a basis for future metabolic indicators of CF evolution. Many of the identified metabolites and pathways have potential as biomarkers to predict current and future disease and assess the impact of targeted therapies. While promising metabolites for describing and predicting CF have also been identified in multiple studies, further validation and exploration are needed. This study was the first in this field that combined metabolomics and network analyses from different CF genotypes. Our results may be important for a better understanding of the role played by the metabolism in the progression of this disease.

Institutional Review Board statement

The study was conducted in accordance with the Declaration of Helsinki.

Declaration of interest

The Authors declare no conflict of interest. This research received no external funding.

References

1. López-Valdez JA, Aguilar-Alonso LA, Gándara-Quezada V, Ruiz-Rico GE, Ávila-Soledad JM, Reyes AA, Pedroza-Jiménez FD. Cystic fibrosis: current concepts. Bol Med Hosp Infant Mex. 2021;78:584-96.
2. Scotet V, L'Hostis C, Férec C. The Changing Epidemiology of Cystic Fibrosis: Incidence, Survival and Impact of the CFTR Gene Discovery. Genes (Basel). 2020;11:E589.
3. Sosnay PR, Siklosi KR, Van Goor F, Kaniecki K, Yu H, Sharma N, Ramalho AS, Amaral MD, Dorfman R, Zielenski J, Masica DL, Karchin R, Millen L, Thomas PJ, Patrinos GP, Corey M, Lewis MH, Rommens JM, Castellani C, Penland CM, Cutting GR. Defining the disease liability of variants in the cystic fibrosis transmembrane conductance regulator gene. Nat Gen. 2013;45:1160-7.
4. Cantin AM, Hartl D, Konstan MW, Chmiel JF. Inflammation in cystic fibrosis lung disease: Pathogenesis and therapy. J Cyst Fibros. 2015;14:419-30.
5. Singh VK, Schwarzenberg SJ. Pancreatic insufficiency in Cystic Fibrosis. J Cyst Fibros. 2017;16:S70-8.
6. Koch C, Høiby N. Pathogenesis of cystic fibrosis. Lancet. 1993;34:1065-9.
7. O'Sullivan BP, Freedman SD. Cystic fibrosis. Lancet. 2009;373:891-1904.
8. Viviani L, Padoan R, Giglio L, Bossi A. The Italian registry for cystic fibrosis: what has changed in the last decade. Epidemiol Prev. 2003;27:91-6.
9. Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, Grzelczak Z, Zielenski J, Lok S, Plavsic N, Chou JL. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science. 1989;245:1066-73.
10. Férec C, Scotet V. Genetics of cystic fibrosis: Basics. Arch Pediatr. 2020;27(Suppl 1):eS4-7.
11. Nicholson JK, Lindon JC, Holmes E. Metabonomics: understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica. 1999;29:1181-9.
12. Fanos V, Antonucci R, Barberini L, Noto A, Atzori L. Clinical application of metabolomics in neonatology. J Matern Fetal Neonatal Med. 2012;25:104-9.
13. Chandler JD, Esther CR. Metabolomics of airways disease in cystic fibrosis. Curr Opin Pharmacol. 2022;65:102238.
14. Liessi N, Pedemonte N, Armirotti A, Braccia C. Proteomics and Metabolomics for Cystic Fibrosis Research. Int J Mol Sci. 2020;21:E5439.
15. Wetmore DR, Joseloff E, Pilewski J, Lee DP, Lawton KA, Mitchell MW, Milburn MV, Ryals JA, Guo L. Metabolomic profiling reveals biochemical pathways and biomarkers associated with pathogenesis in cystic fibrosis cells. J Biol Chem. 2010;285:30516-22.

16. Joseloff E, Sha W, Bell SC, Wetmore DR, Lawton KA, Milburn MV, Ryals JA, Guo L, Muhlebach MS. Serum metabolomics indicate altered cellular energy metabolism in children with cystic fibrosis. Pediatr Pulmonol. 2014;49:463-72.
17. Caley LR, White H, de Goffau MC, Floto RA, Parkhill J, Marsland B, Peckham DG. Cystic Fibrosis-Related Gut Dysbiosis: A Systematic Review. Dig Dis Sci. 2023;68: 1797-814.
18. Piras C, Mussap M, Noto A, De Giacomo A, Cristofori F, Spada M, Fanos V, Atzori L, Francavilla R. Alterations of the Intestinal Permeability are Reflected by Changes in the Urine Metabolome of Young Autistic Children: Preliminary Results. Metabolites. 2022;12:104.
19. Caria P, Tronci L, Dettori T, Murgia F, Santoru ML, Griffin JL, Vanni R, Atzori L. Metabolomic Alterations in Thyrospheres and Adherent Parental Cells in Papillary Thyroid Carcinoma Cell Lines: A Pilot Study. Int J Mol Sci. 2018;19:2948.
20. Lipfert M, Rout MK, Berjanskii M, Wishart DS. Automated Tools for the Analysis of 1D-NMR and 2D-NMR Spectra. Methods Mol Biol. 2019;2037:429-49.
21. Veit G, Avramescu RG, Chiang AN, Houck SA, Cai Z, Peters KW, Hong JS, Pollard HB, Guggino WB, Balch WE, Skach WR, Cutting GR, Frizzell RA, Sheppard DN, Cyr DM, Sorscher EJ, Brodsky JL, Lukacs GL. From CFTR biology toward combinatorial pharmacotherapy: expanded classification of cystic fibrosis mutations. Mol Biol Cell. 2016;27(3):424-33.
22. Drumm ML, Konstan MW, Schluchter MD, Handler A, Pace R, Zou F, Zariwala M, Fargo D, Xu A, Dunn JM, Darrah RJ, Dorfman R, Sandford AJ, Corey M, Zielenski J, Durie P, Goddard K, Yankaskas JR, Wright FA, Knowles MR, Gene Modifier Study Group Genetic modifiers of lung disease in cystic fibrosis. N Engl J Med. 2005;353:1443-53.
23. Corvol H, Blackman SM, Boëlle PY, Gallins PJ, Pace RG, Stonebraker JR, Accurso FJ, Clement A, Collaco JM, Dang H, Dang AT, Franca A, Gong J, Guillot L, Keenan K, Li W, Lin F, Patrone MV, Raraigh KS, Sun L, Zhou YH, O'Neal WK, Sontag MK, Levy H, Durie PR, Rommens JM, Drumm ML, Wright FA, Strug LJ, Cutting GR, Knowles MR. Genome-wide association meta-analysis identifies five modifier loci of lung disease severity in cystic fibrosis. Nat Commun. 2015;6:8382.
24. Szczesniak R, Rice JL, Brokamp C, Ryan P, Pestian T, Ni Y, Andrinopoulou ER, Keogh RH, Gecili E, Huang R, Clancy JP, Collaco JM. Influences of environmental exposures on individuals living with cystic fibrosis. Expert Rev Respir Med. 2020;14:737-48.
25. Baker, EH, Clark N, Brennan AL, Fisher DA, Gyi KM, Hodson ME, Philips BJ, Baines DL, Wood DM. Hyperglycemia and cystic fibrosis alter respiratory fluid glucose concentrations estimated by breath condensate analysis J Appl Physiol. 1985;102:1969-75.
26. Gur M, Zuckerman-Levin N, Masarweh K, Hanna M, Laghi L, Marazzato M, Levanon S, Hakim F, Bar-Yoseph R, Wilschanski M, Bentur L. The effect of probiotic administration on metabolomics and glucose metabolism in CF patients. Pediatr Pulmonol. 2022;57:2335-43.
27. Cruzat V, Macedo Rogero M, Noel Keane K, Curi R, Newsholme P. Glutamine: Metabolism and Immune Function, Supplementation and Clinical Translation. Nutrients. 2018;10:1564.
28. Dessì A, Atzori L, Noto A, Visser GH, Gazzolo D, Zanardo V, Barberini L, Puddu M, Ottonello G, Atzei A, De Magistris A, Lussu M, Murgia F, Fanos V. Metabolomics in newborns with intrauterine growth retardation (IUGR): urine reveals markers of metabolic syndrome. J Matern Fetal Neonatal Med. 2011;24(Suppl 2):35-9.
29. Caputo M, Bona E, Leone I, Samà MT, Nuzzo A, Ferrero A, Aimaretti G, Marzullo P, Prodam F. Inositols and metabolic disorders: From farm to bedside. J Tradit Complement Med. 2020;10:252-9.
30. Mallet RT, Sun J, Knott EM, Sharma AB, Olivencia-Yurvati AH. Metabolic cardioprotection by pyruvate: recent progress. Exp Biol Med (Maywood). 2005;230:435-43.
31. Moliteo E, Sciacca M, Palmeri A, Papale M, Manti S, Parisi GF, Leonardi S. Cystic Fibrosis and Oxidative Stress: The Role of CFTR. Molecules. 2022;27:5324.
32. Causer A, Shute JK, Cummings MH, Shepherd A, Gruet M, Costello J, Bailey S, Lindley M, Pearson C, Connett G, Allenby MI, Carroll MP, Daniels T, Saynor ZL. Circulating biomarkers of antioxidant status and oxidative stress in people with cystic fibrosis: A systematic review and meta-analysis. Redox Biol. 2020;32:101436.
33. Kumps A, Duez P, Mardens Y. Metabolic, nutritional, iatrogenic, and artifactual sources of urinary organic acids: a comprehensive table. Clin Chem. 2002;48:708-17.
34. Shaw W. Increased urinary excretion of a 3-(3-hydroxyphenyl)-3-hydroxypropionic acid (HPHPA), an abnormal phenylalanine metabolite of Clostridia spp. in the gastrointestinal tract, in urine samples from patients with autism and schizophrenia. Nutr Neurosci. 2010;13:135-43.
35. Keşli R, Gökçen C, Buluğ U, Terzi Y. Investigation of the relation between anaerobic bacteria genus clostridium and late-onset autism etiology in children. J Immunoassay Immunochem. 2014;35:101-9.
36. Burke DG, Fouhy F, Harrison MJ, Rea MC, Cotter PD, O'Sullivan O, Stanton C, Hill C, Shanahan F, Plant BJ, Ross RP. The altered gut microbiota in adults with cystic fibrosis. BMC Microbiol. 2017;17:58.
37. Chalmers RA, Valman HB, Liberman MM. Measurement of 4-hydroxyphenylacetic aciduria as a screening test for small-bowel disease. Clin Chem. 1979;25:1791-4.
38. Finegold SM, Molitoris D, Song Y, Liu C, Vaisanen ML, Bolte E, McTeague M, Sandle, R, Wexler H, Marlowe EM, Collins MD, Lawson PA, Summanen P, Baysallar M, Tomzynski TJ, Read E, Johnson E, Rolfe R, Nasir P, Shah H, Haake DA, Manning P, Kaul A. Gastrointestinal microflora studies in late-onset autism. Clin Infect Dis. 2002;35:S6-16.

39. Wang Y, Leong LEX, Keating RL, Kanno T, Abell GCJ, Mobegi FM, Choo JM, Wesselingh SL, Mason AJ, Burr LD, Rogers GB. Opportunistic bacteria confer the ability to ferment prebiotic starch in the adult cystic fibrosis gut. Gut Microbes. 2019;10: 367-81.
40. Kaakoush NO, Pickford R, Jaffe A, Ooi CY. Is there a role for stool metabolomics in cystic fibrosis? Pediatr Int. 2016;58: 808-11.
41. Dodge JA, Turck D. Cystic fibrosis: nutritional consequences and management. Best Pract Res Clin Gastroenterol. 2006;20:531-46.
42. Engelen MP, Com G, Wolfe RR, Deutz NE. Dietary essential amino acids are highly anabolic in pediatric patients with cystic fibrosis. J Cyst Fibros. 2013;12:445-53.
43. Gruet M, Troosters T, Verges S. Peripheral muscle abnormalities in cystic fibrosis: Etiology, clinical implications and response to therapeutic interventions. J Cyst Fibros. 2017;16:538-52.