



# **1H NMR-based metabolomics for infertility diagnosis in men with varicocele**

Filipe Tenorio Lira Neto<sup>1,2</sup> · Ronmilson Alves Marques<sup>3</sup> · Alexandre de Freitas Cavalcanti Filho<sup>2</sup> · Leslie Clifford Noronha Araujo<sup>4</sup> · Salvador Vilar Correia Lima<sup>5</sup> · Licarion Pinto<sup>6</sup> · Ricardo Oliveira Silva<sup>6</sup>

Received: 15 October 2019 / Accepted: 16 July 2020 / Published online: 26 July 2020  
© Springer Science+Business Media, LLC, part of Springer Nature 2020

## **Abstract**

**Purpose** “Omics” techniques have been used to understand and to identify biomarkers of male infertility. We report on the first metabolomics models created to diagnose varicocele and infertility among men with varicocele.

**Methods** We recruited 35 infertile men with varicocele (VI group), 21 fertile men with varicocele (VF group) and 24 fertile men without varicocele (C group). All men underwent standard semen analysis, scrotal duplex ultrasonography, and sexual hormone level measurement. Hydrogen-1 nuclear magnetic resonance (<sup>1</sup>H NMR) spectra of seminal plasma were used to create metabolomics models to discriminate between men with and without varicocele, and between fertile and infertile men with varicocele.

**Results** Using the statistical formalisms partial least square discriminants analysis and genetic algorithm-based linear discriminant analysis (GA-LDA), we created two models that discriminated the three groups from each other with accuracy of 92.17%. We also created metabolomics models using orthogonal partial least square discriminants analysis and GA-LDA that discriminated VF group from VI group, with an accuracy of 94.64% and 100% respectively. We identified 19 metabolites that were important in group segregation: caprate, 2-hydroxy-3-methylvalerate, leucine, valine, 3-hydroxybutyrate, lactate, alanine, 4-aminobutyrate, isoleucine, citrate, methanol, glucose, glycosides, glycerol-3-phosphocoline, n-acetyltirosine, glutamine, tyrosine, arginine, and uridine.

**Conclusions** <sup>1</sup>HNMR-based metabolomics of seminal plasma can be used to create metabolomics models to discriminate between men with varicocele from those without varicocele, and between fertile men with varicocele from those infertile with varicocele. Furthermore, the most important metabolites for group segregation are involved in the oxidative stress caused by varicocele.

**Keywords** Metabolome · Metabolomics · Fertility · Semen · NMR spectroscopy

✉ Filipe Tenorio Lira Neto  
ftenorio.uro@gmail.com

Ronmilson Alves Marques  
milson.ufpe@gmail.com

Alexandre de Freitas Cavalcanti Filho  
alexfcf18@gmail.com

Leslie Clifford Noronha Araujo  
leslie.clifford@gmail.com

Salvador Vilar Correia Lima  
salvadorvilarcorreia@me.com

Licarion Pinto  
licarion.pinto@ufpe.br

Ricardo Oliveira Silva  
ricardo.silva@ufpe.br

<sup>1</sup> Department of Surgery, Universidade Federal de Pernambuco, Rua Guilherme Salazar, n 150, apt 601E, Recife, PE ZIP 52061-275, Brazil

<sup>2</sup> Instituto de Medicina Integral Prof. Fernando Figueira, Rua dos Coelho, 300, Coelhos, Recife, PE ZIP 50070-902, Brazil

<sup>3</sup> Fundamental Chemistry Department, Universidade Federal de Pernambuco, Av. Jornalista Anibal Fernandes, s/n° Cidade Universitária, Recife, PE ZIP 50740-560, Brazil

<sup>4</sup> Faculdade de Medicina de Olinda, Rua Dr. Manoel de Almeida Belo, n 1333, Olinda, PE ZIP 53030-030, Brazil

<sup>5</sup> Department of Surgery, Universidade Federal de Pernambuco, Hospital das Clínicas - Campus UFPE, Av. Prof. Moraes Rego, s/n° - Bloco “A” - Térreo Cidade Universitária, Recife, PE ZIP 50670-420, Brazil

<sup>6</sup> Fundamental Chemistry Department, Universidade Federal de Pernambuco, Av. Jornalista Anibal Fernandes, s/n° Cidade Universitária, Recife, PE ZIP 50740-560, Brazil

## Introduction

Varicocele is defined as an abnormal dilation of internal spermatic veins due to incompetent venous valves, leading to venous reflux. This condition is found in 15% of all men, and its prevalence increases to 40% in men with primary infertility and to 80% in men with secondary infertility [1]. Although considered a common etiology of male infertility, the cause and effect relationship between varicocele and infertility is still under debate. Studies of infertile men with and without varicocele have revealed that the presence of varicocele has a negative impact on semen analysis (SA) parameters, whereas studies including fertile men showed no correlation between varicocele and SA parameters [2–4]. In fact, it is estimated that 80% of the men with varicocele are fertile, which highlights the complexity of this condition [5]. Despite the controversy, surgical treatment of clinical varicoceles in infertile men has consistently been shown to improve semen parameters [6], pregnancy rate [7], and testosterone levels [8]. In addition, SA parameters are not good predictors of male fertility [9, 10]; thus, counseling men with varicocele regarding their future fertility status based solely on SA data is challenging.

The “omics” methods use high-throughput techniques to study various biological systems in a step-by-step fashion. This integrated approach relies heavily on chemical analytical methods, producing a massive amount of data that needs to be interpreted using bioinformatics and computational analysis tools in order to deliver clinically useful information [11]. Recently, different “omics” techniques have been used to better understand male infertility conditions, helping the development of new diagnostic, therapeutic, and prognostics tools [12].

The set of metabolites in a biological sample, the metabolome, is the end product of gene transcription and is influenced by various physiological and pathological processes, having a direct relationship to phenotypes [13]. Metabonomics is defined as the study of the dynamic multiparametric metabolic response of living systems to endogen or exogen stimuli [14]. Metabonomics strategies analyze the changes of the endogenous metabolic profile caused by a certain stimulus, using spectral data obtained from biological samples and multivariate statistical tools [14–16]. Hydrogen-1 nuclear magnetic resonance ( $^1\text{H}$  NMR) spectroscopy is a powerful tool for chemical structural elucidation and chemical dynamic studies, and one of the most used techniques in metabonomics studies [17, 18]. In addition, metabonomics studies apply chemometrics and multivariate statistical tools to process and analyze data obtained by  $^1\text{H}$  NMR spectroscopy. This strategy has been successfully used in several studies in which changes in the metabolome of human biofluids were used in the diagnosis and prognosis of different conditions [15, 19].

The multivariate statistical tools employed in metabonomics studies are commonly divided into unsupervised and supervised methods. Both use information provided by an instrumental

analysis to assess differences among study groups. The supervised methods also use information that enables correspondence from samples to a class/group in order to build a predictive model. Principal components analysis (PCA) is the statistical formalism most commonly used for unsupervised studies because of its ability to reduce a large number of variables to a smaller number of uncorrelated variables, facilitating the task of evaluate differences between samples [20, 21]. Among supervised methods, discriminant analysis (DA) formalisms are the most commonly used in metabonomics studies, especially linear discriminant analysis (LDA) coupled with variable selection, the partial least square discriminant analysis (PLS-DA), and the orthogonal partial least square discriminant analysis (OPLS-DA). The greatest advantage of DA is the ability to identify variables that discriminate the classes even if the variation of these variables is lower than other variables. Therefore, in these cases where the metabolic profile do not present intense changes, DA formalisms are best suited for creating models capable of discriminating different groups of samples [22].

PLS-DA is a statistical formalism used to optimize separation between different groups of samples, by linking a data matrix  $X$ , containing the raw dataset, to a matrix  $I$  containing the group or class membership [23]. It creates multidimensional score plots depicting the segregation between different groups, making the visual interpretation of the data easier. Furthermore, PLS-DA also provides other statistics such as variable importance in projection (VIP), that can be used to identify the most important variables for group segregation and their relative concentration, and  $R^2$  and  $Q^2$  statistics, which are used to evaluate the predictive accuracy of the model [24].

The OPLS-DA is a variant of the PLS-DA that uses the information contained in the class matrix  $I$  to decompose the dataset matrix  $X$  into scores separated by orthogonal projections and loadings. OPLS-DA is often used in substitution of PLS-DA to separate predictive from non-predictive (orthogonal) variation related to groups. In this way, OPLS-DA builds more easily interpretable models compared with PLS-DA data since only two classes are used [25].

The LDA is a supervised statistical formalism that performs dimensional reduction for pattern-classification. LDA enables the determination of a decision surface or injunction in space, so that samples of one class are segregated on one side of the decision axis, and samples of the other class are grouped on the other [26]. The main limitation of LDA is the need for more samples than the number of variables studied, which requires a variable selection method in our case. The Genetic Algorithm–Linear Discriminant Analysis employs genetic algorithms to select variables with not only better discrimination power, but also more biochemically informative. This is made by building Discriminant Functions (DF) which are linear combinations of the selected variables. Normally, LDA results are presented using score and loading plots. Scores are the results obtained for each sample, while loadings

indicate the importance of each variable (metabolite) for discrimination. These plots provide simpler LDA discrimination models with clearer physiological meaning and higher diagnostic efficiency [27]. In the loading plots, the variables most important for class discrimination are presented. The loading values are normalized and can be used in conjunction with the correspondent signal (for example, NMR spectrum) to calculate the score of each sample that can be used in the discrimination. In the score plots, the similarities or differences between the classes are represented, providing class discrimination based on the characteristics analyzed and weighted by the loadings. It is important to note that the discrimination occurs by the analysis of the values from all of the variables expressed in the loadings. Therefore, the results of one or more variables may be different for samples from the same class. Fisher's linear discriminant score plot gives a visual interpretation of the data structure, allowing the identification of clusters of samples from the same class. The GA-LDA formalism does not provide information regarding the relative concentration of the variables identified as the most important for group segregation.

There are several reports using  $^1\text{H}$  NMR-based metabolomics to identify biomarkers useful in the diagnosis of some male infertility conditions [12]. Most of these studies have used seminal plasma, since this biofluid is composed in part by epididymal and testicular fluid, making it the most likely to carry metabolites from the testis and epididymis, organs most commonly involved in the pathogenesis of male infertility. However, up to this date, there are no published studies applying metabolomics to investigate male infertility due to varicocele. Men with varicocele have a high level of oxidative stress in the testes due to the release of reactive oxygen species (ROS) by leukocytes and abnormal spermatozoa [28]. There are studies that have demonstrated increased seminal concentration of ROS as well as end products of oxidative stress when men with varicocele were compared with fertile men [29]. Therefore, we can expect changes in the seminal metabolome of infertile men with varicocele. In addition, there are some proteomics studies that point towards a differential expression of seminal proteins in men with varicoceles when compared with controls without the condition [30–33]. Since proteome and metabolome are directly interconnected, these findings also suggest that the seminal metabolome may be altered in men with varicocele, and that such alterations can be used to classify these men into different phenotypes such as fertile versus infertile.

The aims of this pilot study were to study changes in the seminal metabolome caused by varicocele using  $^1\text{H}$  NMR spectra of seminal plasma, and, subsequently, to develop metabolomics models capable of segregating men regarding the presence of varicocele and their fertility status.

## Methods

### Patient and study groups

This study was approved by the Research Ethics Committee of Universidade Federal de Pernambuco and Instituto de Medicina Integral Prof. Fernando Figueira (approval n. 2.075.028). Men aged between 18 and 50 years old attending a male fertility outpatient clinic were evaluated by the same male fertility specialist. Ninety volunteers were enrolled and had semen samples collected; however, ten were excluded due to the bad quality of the sample spectra. The 80 remaining participants were initially stratified into two groups: Control group (C group)—24 fertile healthy volunteers without palpable varicocele, who had at least one child born in the last 12 months, had no past history of infertility or fertility treatments, and were seeking vasectomy; and the varicocele group (V group)—56 volunteers with palpable varicocele, independently of their fertility status, with no history of fertility treatments. The V group was then subdivided into two groups: Varicocele fertile group (VF group)—21 fertile volunteers with palpable varicocele who had at least one child born in the last 12 months, and with no history of infertility or fertility treatments; and varicocele infertile group (VI group)—35 infertile volunteers with palpable varicocele who failed to conceive after 12 months of regular and unprotected intercourse, with no history of other infertility causes or fertility treatments and no known cause of female infertility in the partner. Infertility was defined based on the Practice Committee of the American Society for Reproductive Medicine [34]. Participants were excluded if they presented any of the following: evidence of urogenital infection, urological diseases diagnosed by andrological examination, genetic defects, history of cryptorchidism, current or recent (< 12 months) use of testosterone or any other anabolic steroids, history of chemotherapy or radiotherapy, history of scrotal trauma, history of scrotal surgery. All the participants underwent full andrological anamnesis and physical exam. Puberty age was considered as the age when the participant first noticed an increase of penile or testicular size, or the appearance of pubic hair. Physical exam was performed in a warm and well illuminated room, with the participant in supine and standing positions. Testicular size was measured using a Prader orchidometer and varicocele grade was classified based on the Dubin and Amelar criteria [35]. The female partners were questioned about their age, past history of fertility and/or fertility treatments, past history of pelvic surgeries or pelvic pathologies, and about the regularity and intensity of their menstrual cycle. However, further female fertility evaluation was not available at our institution. Written informed consent was obtained from all participants before their inclusion in the study and anonymity was granted.

All participants underwent SA, scrotal duplex ultrasonography (USGD), and sexual hormones levels measurement. Semen analysis was performed according to the World Health Organization guidelines, seminal pH was measured using pH testing strips [36]. Scrotal duplex ultrasonography (GE Logiq S8 GE Healthcare, Wauwaosa, WI, USA) using a linear high-frequency probe (SL 15-8-MHz GE Healthcare, Wauwaosa, WI, USA) was used to evaluate testicular size, varicocele diameter, and venous reflux. Venous blood samples were drawn from each participant between 7 and 11 a.m., after an overnight fast. Total testosterone (reference range: 206 to 1200 ng/dL), estradiol (reference range: 11.6 to 41.2 pg/mL), follicle-stimulating hormone (FSH) (reference range 1.4 to 18.1 mUI/mL), luteinizing hormone (LH) (reference range 1.5 to 9.3 mUI/mL), and sexual hormone-binding protein (reference range 10 to 57 nmol/L) levels were measured in real-time by a solid-phase chemiluminescent immunoassay with the use of an automated analyzer (ADVIA Centaur XP, Siemens Healthcare Diagnostics, [www.siemens-healthineers.com](http://www.siemens-healthineers.com)). Albumin levels (reference range 3.4 to 4.8 g/dL) were measured using a colorimetric assay (Abbott Diagnostics, Abbott Park, IL, USA) with an autoanalyzer (Architect® c16000, Abbott Diagnostics, [www.corelaboratory.abbott](http://www.corelaboratory.abbott)). Free testosterone levels (reference range 49.9 to 199.9 pg/mL) were calculated using the validated formula of Vermeulen et al. [37].

### <sup>1</sup>H NMR analysis

<sup>1</sup>H NMR spectra were acquired at the Analytics Central of the Fundamental Chemistry Department at the Universidade Federal de Pernambuco. Following semen liquefaction, the samples were centrifuged at 3000g for 15 min to separate the spermatozoa from the supernatant seminal plasma. The seminal plasma samples were kept frozen at -20 °C until <sup>1</sup>H NMR analysis. After thawing, 400 µL of seminal plasma was added to 200 µL of D<sub>2</sub>O, and each sample was submitted to <sup>1</sup>H NMR analysis separately and randomly. <sup>1</sup>H NMR spectra were acquired using a VNMR400 spectrometer operating at 400 MHz. The Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence with presaturation of the water signal was used as follows: spectral window of 4.8 kHz, saturation delay of 2.0 s, acquisition time of 1.704 s, 90° RF pulse, temperature of 27 °C, 88 cycles, tau equal to 0.0004 s, bigtau equal to 0.07 s, and 128 scans. Each spectrum was acquired using total time equal to 10.0 min 4 seg. The line broadening used was 0.3 Hz, baseline and phase distortions were corrected manually, and the signal attributed to the citrate methylene group ( $\delta$  2.65 ppm) was used as internal chemical shift reference. Using the MestreNova 9.0 software, the region of the spectra between  $\delta$  8.01 and 0.69 ppm was divided into regions equals to 0.04 ppm. The region between  $\delta$  5.50 and 4.50 ppm was excluded because it contained a residual signal of water. A

matrix with 80 rows (cases) and 159 variables (bins of <sup>1</sup>H NMR spectra plus class variable) was built and submitted to multivariate analysis.

### Statistical analysis

To investigate the distribution of demographic and clinical data between groups, univariate tests were performed using GraphPad Prism 6 software (GraphPad Software, Inc., La Jolla, CA). The Shapiro-Wilk's test was used to analyze the normality of data distribution. One-way analysis of variance (ANOVA) coupled with Tukey's test when necessary was used to compare the following variables among the 3 groups: age, partner age, puberty age, testicular sizes by physical exam, BMI, sperm progressive motility, sperm morphology, semen volume, total testosterone, free testosterone, estradiol, LH, and testicular sizes by USGD. Kruskal-Wallis coupled with Dunn's test when necessary was used to compare the following variables among the 3 groups: sperm concentration, total sperm count, total progressive sperm count, FSH, and testicular vein sizes. Fisher's exact test was used to compare varicocele grade and the proportion of bilateral varicocele between VF and VI groups. A *P* value of <0.05 was set as the level of statistical significance.

### Multivariate statistical analysis

We built metabonomics models aiming to differentiate among the following: (i) control, varicocele fertile, and varicocele infertile groups; (ii) the control group from the varicocele group; and (iii) the varicocele fertile group from the varicocele infertile group. PLS-DA and GA-LDA were used to build the model using all three groups, while OPLS-DA and also GA-LDA were used to build models using two groups (i.e., C group versus V group; and VF group versus VI group).

All metabonomics models based on PLS-DA were performed using the web-based platform for metabonomics studies MetaboAnalyst 4.0. [38], and the GA-LDA-based models were performed using the MATLAB® 2010a computer environment (The MathWorks, Inc., Natick, Massachusetts, USA).

In the pre-processing step of the model using OPLS-DA to discriminate the VF group from the VI group, spectral data were normalized by sum, while each variable was normalized by autoscaling. For all other metabonomics models, each sample was normalized by mean and each variable was normalized by autoscaling. Latent variables were built using leave one out cross validation (LOOCV), and PLS-DA/OPLS-DA were validated using a permutation test with 2000 class permutations. The R<sup>2</sup> value was determined from LOOCV. PLS-DA and OPLS-DA models also provided a quantitative measure of the discriminating power of each spectral bin. For this purpose, we made use of the VIP score. The formalism used to build GA-LDA models consists on the division of the samples

into a training set containing 70% of the samples, and a test set composed of the remaining 30%, using the Kennard Stone algorithm [39]. This strategy makes it possible not only to evaluate the quality of the model but also if it could predict new samples with accuracy. For the GA-LDA models, the thresholds used to distinguish the classes in each metabonomic model were calculated by the mean distance between the mean scores of the samples from each class.

The efficiency of each metabonomic model was calculated using the confusion matrix, a matrix where the columns represent the actual groups and the rows represent the groups predicted by the model. The better the model correctly assigns samples to the groups, the higher is the accuracy of the model.

The chemical shifts identified by the models as the most important variables for group discrimination had their correspondent metabolites assigned using the using human metabolome database (HMDB) [40] and the work by Paiva et al. [41].

## Results

### Group demographics

The clinical characteristics of the three groups are shown in Table 1. Participants' and partners' ages, as well as puberty age, were similar among the three groups. The VI group participants showed smaller testes were compared with the VF and C groups. The grade of varicoceles did not differ between the VF and the VI groups, but the VI group had a higher proportion of men with bilateral varicoceles and bilaterally larger venous diameters on USGD. Regarding SA parameters, there were no differences between the C and VF groups, but the VI had lower sperm concentration, lower total sperm count, lower sperm motility, and lower total progressive motile sperm count when compared with the other two groups. The VI group showed higher FSH levels than the C and VF groups, and higher LH levels than the C group. Interestingly, the C group had lower levels of serum total testosterone than the VI group, but free testosterone levels were not different among the groups.

### Spectral data and multivariate analyses

An example of  $^1\text{H}$  NMR spectrum of seminal plasma with some attributions of signals is shown in Fig. 1. After obtaining  $^1\text{H}$  NMR spectra, we performed an exploratory analysis using the PCA formalism, but the model built was not able to segregate the groups. In sequence, PLS-DA formalism was used to obtain a metabonomics model that was able to discriminate the 3 groups (C group versus VF group versus VI group). Figure 2 shows a tridimensional PLS-DA score plot, which reveals that the samples are grouped in 3 different groups as

expected, and the VIP score showing the most important chemical shifts for discrimination.

The PLS-DA metabonomics model presented  $R^2$  value equal to 0.82 and discriminated the participants based on their biochemical status. The permutation test resulted in a  $P$  value equal to 0.023. Based on the chemical shifts of the variables indicated by the VIP score, we identified the following 13 metabolites that were important in group segregation: caprate, valine, 3-hydroxybutyrate, lactate, 4-aminobutyrate (GABA), isoleucine, citrate, glycosides, n-acetyltyrosine, glutamine, tyrosine, arginine, and uridine (Table 2). There were other 5 chemical shifts indicated by the VIP, but that we were unable to identify their correspondent chemical compounds (1.41, 2.53, 2.69, 5.53, and 6.77 ppm).

Using the OPLS-DA formalism, we built a metabonomics model able to discriminate healthy fertile participants (C group) from participants with varicocele independent of their fertility status (V group) with an accuracy of 91.25%. Figure 3 presents the score plot and permutation tests from the OPLS-DA model. Table 3 shows the confusion matrix from OPLS-DA metabonomics model, considering the 95% confidence limit.

We also used OPLS-DA to create a metabonomics model that discriminated, with high accuracy (94.64%), fertile men with varicocele from infertile men with varicocele (Fig. 4). Table 4 shows the confusion matrix from this OPLS-DA metabonomics model, considering the 95% confidence limit.

The models created using PLS-DA and OPLS-DA were validated by the LOOCV and permutation tests. In order to further validate our findings, we also used the GA-LDA formalism to create other metabonomics models based on a training sample set and validated using a test sample set.

The first model created using GA-LDA was able to segregate the 3 groups (C group versus VF group versus VI group) utilizing two linear discriminant functions (DF1 and DF2) (Fig. 5a). This model was created using 56 samples (training set) and validated using 24 samples (test set). The importance of each variable on the model is shown by the loading (bar) plot that represents the Fisher's loading of discriminant functions 1 and 2, respectively (Fig. 5 b and c). The direction of the bar points to the location of the sample, and the intensity ( $y$ -axis) is correlated to the importance of the variable. From the 20 variables selected by the model, we identified 3 metabolites that were also selected by the PLS-DA model (i.e., 3-hydroxybutyrate, lactate, and isoleucine), and another 6 new metabolites not previously selected; 2-hydroxy-3-methylvalerate, leucine, alanine, methanol, glucose, and glycerol-3-phosphocoline (GPC) (Table 2). There were 11 other chemical shifts that were selected by the genetic algorithms, but that we could not assign their correspondent chemical compounds (1.21, 2.09, 2.25, 2.45, 2.57, 3.13, 3.97, 4.41, 4.45, 7.01, and 7.09). The model was 92.17% accurate and showed high specificity (Tables 5 and 6)

**Table 1** Clinical and laboratorial characteristics

	C (n = 24)	VF (n = 21)	VI (n = 35)	P value
Clinical characteristics				
Age (years), mean ± SD	34.2 ± 4.3	33.7 ± 4.9	32.0 ± 6.1	0.22*
Partner age (years), mean ± SD	29.2 ± 3.8	32.3 ± 4.4	30.7 ± 5.1	0.14*
Puberty age (years), mean ± SD	14.2 ± 1.5	14.8 ± 1.8	13.8 ± 1.4	0.08*
Right testis size by physical exam (mL), mean ± SD	22.7 ± 5.0 <sup>a</sup>	19.0 ± 3.5 <sup>a</sup>	15.7 ± 5.4 <sup>b</sup>	< 0.05*
Left testis size by physical exam (mL), mean ± SD	22.0 ± 5.6 <sup>a</sup>	17.4 ± 5.0 <sup>b</sup>	12.5 ± 5.6 <sup>c</sup>	< 0.05*
BMI (kg/m <sup>2</sup> ), mean ± SD	28.0 ± 4.0	26.3 ± 4.2	26.2 ± 5.0	0.23*
Right varicocele grade %				
1		67%	48%	
2		33%	43%	
3	NA	0%	9%	0.53 <sup>†</sup>
Left varicocele grade %				
1		24%	17%	
2		38%	26%	
3	NA	38%	57%	0.38 <sup>†</sup>
Bilateral varicoceles %				
	NA	14%	60%	< 0.05 <sup>†</sup>
Semen analysis parameters				
Sperm concentration (10 <sup>6</sup> /mL), median ± IQR	38.9 ± 05.4 <sup>a</sup>	40.0 ± 56.9 <sup>a</sup>	5.0 ± 12.6 <sup>b</sup>	< 0.05 <sup>§</sup>
Total sperm count (10 <sup>6</sup> /mL), median ± IQR	133 ± 312 <sup>a</sup>	105 ± 136 <sup>a</sup>	14 ± 31 <sup>b</sup>	< 0.05 <sup>§</sup>
Progressive motility (%), mean ± SD	65 ± 20 <sup>a</sup>	55 ± 22 <sup>a</sup>	30 ± 15 <sup>b</sup>	< 0.05*
Total progressive motile sperm count (10 <sup>6</sup> /mL), median ± IQR	74 ± 241 <sup>a</sup>	55 ± 89 <sup>a</sup>	2 ± 10 <sup>b</sup>	< 0.05 <sup>§</sup>
Morphology (%), mean ± SD	10.0 ± 6.1	8.6 ± 3.5	7.8 ± 4.5	0.30*
Volume (mL), mean ± SD	4.0 ± 1.5	3.0 ± 1.1	3.2 ± 1.6	0.06*
pH, median ± IQR	8.0 ± 0.5	7.5 ± 1.0	7.3 ± 1.0	NA <sup>¶</sup>
Sexual hormones levels				
Testosterone (ng/dL), mean ± SD	345 ± 135 <sup>a</sup>	422 ± 158 <sup>a,b</sup>	446 ± 179 <sup>b</sup>	< 0.05*
Free testosterone (ng/dL), mean ± SD	7.5 ± 2.3	8.4 ± 2.6	8.8 ± 3.4	0.25*
Estradiol (pg/mL), mean ± SD	33.9 ± 12.0	39.5 ± 12.3	31.7 ± 12.3	0.16*
FSH (mUI/mL), median ± IQR	3.9 ± 2.4 <sup>a</sup>	3.6 ± 2.3 <sup>a</sup>	6.1 ± 4.9 <sup>b</sup>	< 0.05 <sup>§</sup>
LH (mUI/mL), mean ± SD	3.5 ± 1.3 <sup>a</sup>	3.9 ± 1.6 <sup>a,b</sup>	4.9 ± 2.3 <sup>b</sup>	< 0.05*
Ultrasonography parameters				
Right testis size by ultrasound (mL), mean ± SD	16.0 ± 3.9 <sup>a</sup>	14.7 ± 4.7 <sup>a,b</sup>	12.5 ± 4.5 <sup>b</sup>	< 0.05*
Left testis size by ultrasound (mL), mean ± SD	14.6 ± 4.3 <sup>a</sup>	11.8 ± 4.3 <sup>a,b</sup>	11.1 ± 4.3 <sup>b</sup>	< 0.05*
Right testicular vein size by color Doppler (cm), median ± IQR	1.9 ± 0.7 <sup>a</sup>	2.0 ± 0.9 <sup>a,b</sup>	2.5 ± 0.8 <sup>b</sup>	< 0.05 <sup>§</sup>
Left testicular vein size by color Doppler (cm), median ± IQR	2.0 ± 0.7 <sup>a</sup>	3.2 ± 1.0 <sup>b</sup>	3.8 ± 1.3 <sup>b</sup>	< 0.05 <sup>§</sup>

BMI body mass index, FSH follicle-stimulating hormone, LH luteinizing hormone, SD standard deviation, IQR interquartile range

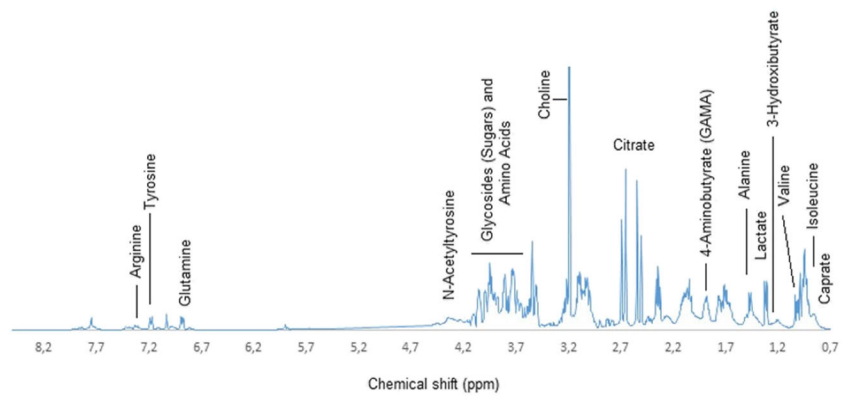
\*Analysis of variance; <sup>†</sup>Fischer's exact test; <sup>§</sup>Kruskal-Wallis test. <sup>a-c</sup> Means in a row without common superscript letters differ ( $P < 0.05$ )

<sup>¶</sup>Test not applicable due to the low accuracy of the pH measurement tool (pH testing strips)

In addition, another GA-LDA model was able to segregate the C group from the V group with high accuracy (Fig. 6 and Tables 7 and 8). The metabolites that had the highest influence on this model were 3-hydroxybutyrate, lactate, and arginine as can be seen in Fig. 6.

Furthermore, we built a different GA-LDA model that segregated the VF group from the VI group (Fig. 7 and Tables 9 and 10). Caprate, alanine, and arginine were the most important metabolites identified in this model.

**Fig. 1** Typical <sup>1</sup>H NMR spectrum (400 MHz, D<sub>2</sub>O) of seminal plasma. The areas under the peaks are associated with the concentration of metabolites weighted by the number of hydrogen nuclei in each chemical environment. Some peaks had their respective chemical compounds assigned in the spectrum



**Discussion**

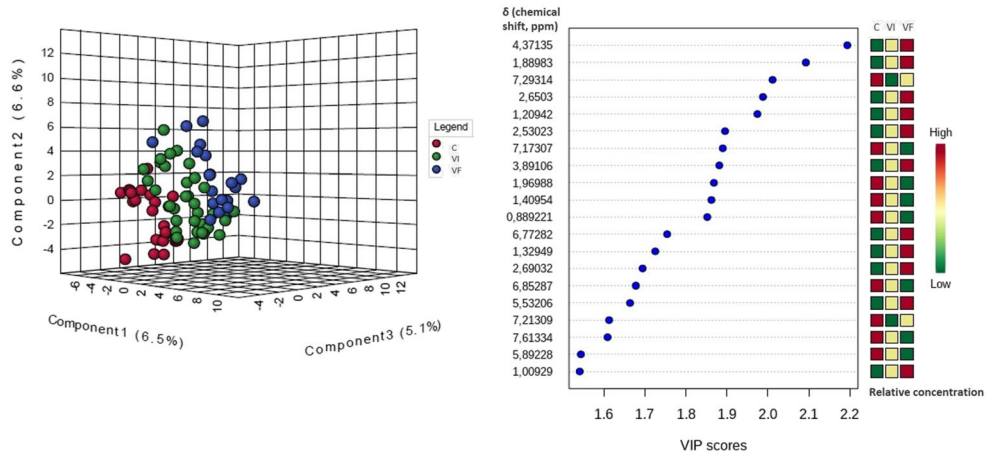
This study is the first to use a metabonomics approach to create models able to differentiate men with varicocele from healthy controls as well as to discriminate fertile men with varicocele from infertile ones. Using the statistical formalisms PLS-DA and GA-LDA, we created two models that distinguished each of the three groups from the others with high accuracy. In addition, using OPLS-DA and GA-LDA, we built models that segregated healthy fertile men from men with varicocele, independently of their fertility status, with an accuracy of 91.25% and 97.06% respectively. Also applying OPLS-DA and GA-LDA, we created two metabonomics models that discriminated group VF from group VI, with an accuracy of 94.64% and 100% respectively. This proof of concept study demonstrates that <sup>1</sup>H NMR-based metabonomics is capable of extracting and analyzing clinically useful information from the seminal plasma of men with varicocele.

This study demonstrated that men with varicocele may present different phenotypes. The VI group had clinical alterations due to the varicocele, such as decreased testicular volume, decreased sperm concentration, decreased total sperm count, decreased sperm motility, and decreased semen volume, when compared with healthy controls (Table 1). In

addition, the levels of testosterone, FSH, and LH were higher in the VI group. These findings might have little clinical importance, however, since all the levels are within the suggested reference values [42, 43]. The ability of these metabonomics models to discriminate men with varicocele from healthy controls with high accuracy corroborates the close relationship between metabolome and phenotypes.

Furthermore, these findings are in conformity with the results of studies using proteomics to analyze changes in the semen of adolescents and men with and without varicocele. When compared with a control group of adolescents without varicocele, the adolescents with varicocele had higher expression of proteins related to the inflammatory and immune responses, even when they did not show changes in SA parameters [33, 44]. The presence of alterations in the seminal metabolome and proteome of individuals with varicocele who do not have abnormal SA suggests that varicocele is a heterogeneous condition with implications that go beyond the impact on the classic SA parameters. The ability of the metabonomics models to segregate infertile men with varicocele from fertile men with varicocele reinforces the heterogeneity of the disease. The clinical importance of this is that, to date, there is no test able to discriminate accurately between these two groups [45]. This model can be very useful in clinical practice, helping in the management of men who have varicocele and want

**Fig. 2** Results of PLS-DA using 80 samples and three classes—control group (C), 24 samples (red dots); varicocele infertile group (VI), 35 samples (green dots); and varicocele fertile group (VF), 21 samples (blue dots). Score plot (left) and VIP score plot (right) showing the 20 most important variables for discrimination



**Table 2** Most important variables for group discrimination and their relative concentration

$\delta$ (chemical shift, ppm)	Metabolite	Control group	Varicocele infertile group	Varicocele fertile group
0.89	Caprate <sup>a</sup>	High	Middle	Low
0.92	2-Hydroxy-3-methylvalerate <sup>b</sup>	NA	NA	NA
0.97	Leucine <sup>b</sup>	NA	NA	NA
1.01	Valine <sup>a</sup>	Low	Middle	High
1.21	3-Hydroxybutyrate <sup>a,b</sup>	Low	Middle	High
1.33	Lactate <sup>a,b</sup>	Low	Middle	High
1.49	Alanine <sup>b</sup>	NA	NA	NA
1.89	4-Aminobutyrate (GABA) <sup>a</sup>	Low	Middle	High
1.97	Isoleucine <sup>a,b</sup>	High	Middle	Low
2.37	Caprate <sup>a</sup>	High	Middle	Low
2.65	Citrate <sup>a</sup>	Low	Middle	High
3.37	Methanol <sup>b</sup>	NA	NA	NA
3.73	Glucose <sup>b</sup>	NA	NA	NA
3.89	Glycosides <sup>a</sup>	Low	Middle	High
4.33	Glycerol-3-phosphocoline <sup>b</sup>	NA	NA	NA
4.37	N-Acetyltyrosine <sup>a</sup>	Low	Middle	High
5.89	Uridine <sup>a</sup>	High	Middle	Low
6.85	Glutamine <sup>a</sup>	High	Middle	Low
7.17	Tyrosine <sup>a</sup>	High	Middle	Low
7.21–7.29	Arginine <sup>a</sup>	High	Low	Middle
7.61	Glutamine <sup>a</sup>	High	Middle	Low
7.73	N-Acetyltyrosine <sup>b</sup>	Low	Middle	High

Assigned using Human Metabolome Database (HMDB) and Paiva et al. (2015), considering chemical shift ( $\delta$ ) ( $\pm$  0.02 ppm)

NA not assessed; <sup>a</sup> by PLS-DA model; <sup>b</sup> by GA-LDA model

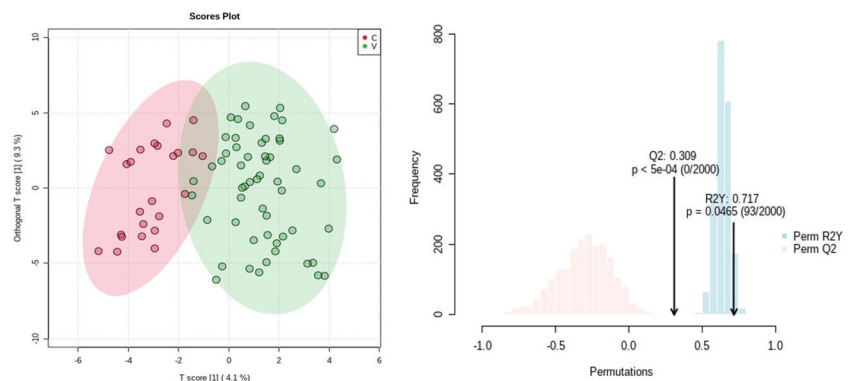
to know about their fertility potential even before they have begun the attempts to conceive. We highlight that, to date, there are no studies published in the literature using any “omics” techniques that compared these two groups of men with varicocele.

For GA-LDA models, the samples were separated into two sets: the training set, composed of 70% of the samples was used to build the model; and the test set, composed by the remaining 30% of the samples and

used to test the model. This strategy of validation confirmed the high accuracy of the GA-LDA models to distinguish among all 3 groups of men included in the study.

The first GA-LDA model was the more complex, aiming to use the metabolome changes to differentiate each group from the others. Therefore, this was the model with the lowest accuracy, both on the training set and the test set, 92.17% and 86.17% respectively (Table 5).

**Fig. 3** Results of the OPLS-DA model to discriminate the control group (C), 24 samples (red dots), from the varicocele group (V), 56 samples (green dots). Score plot (left) and permutation test (right)





**Table 3** Confusion matrix—OPLS-DA metabonomics model. Control group (24 samples) versus varicocele group (56 samples)

Model	Groups		P value
	Varicocele	Control	
Varicocele	53	4	0.04
Control	3	20	

Accuracy 91.25%, sensitivity 94.64%, specificity 83.33%, Positive predictive value 92.98%, and negative predictive value 86.96%

The second GA-LDA model segregated healthy fertile men from men with varicocele independently of their fertility status and resulted in only one false positive. This model had an accuracy of 97.06% and correctly classified all the patients with varicocele. These results indicate that this model is useful as a screening test to identify men with varicocele.

The most important metabonomics model created in this study was the one used to segregate fertile men with varicocele from infertile men with varicocele. In the validation step of this GA-LDA model, only one fertile man was misclassified as infertile, giving an accuracy of 91.67%.

This study identified a total of 19 metabolites important for group segregation. Valine, 3-hydroxybutyrate, lactate, GABA, citrate, glycosides, and n-acetyltyrosine were found in lower concentrations in the C group when compared with VF group, while their concentration was in the middle range in the VI group. On the other hand, caprate, isoleucine, uridine, glutamine, and tyrosine were expressed in the opposite way, with higher concentrations in the C group, lower concentrations in the VF group, and concentrations in the middle range in the VI group. Arginine was the only metabolite found in lower concentration in the VI group and higher concentrations in the C group. In addition, 2-hydroxy-3-methylvalerate, leucine, alanine, methanol, glucose, and glycerol-3-phosphocholine were also important, but their relative concentrations were not determined. It is important to highlight that similar metabolites were found independently of the multivariate statistical tool used; however, the influence of each metabolite was different, varying with the models. These

**Table 4** Confusion matrix—OPLS-DA metabonomics model. Varicocele fertile group (21 samples) versus varicocele infertile group (35 samples)

Model	Groups		P value
	Varicocele fertile	Varicocele infertile	
Varicocele infertile	1	33	0.01
Varicocele fertile	20	2	

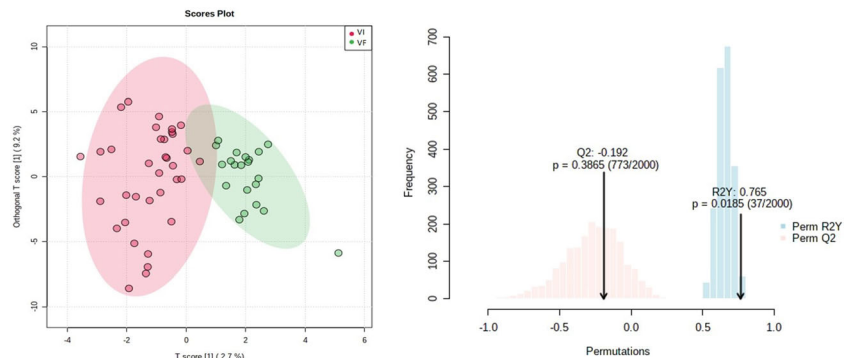
Accuracy 94.64%, sensitivity 95.24%, specificity 95.24%, Positive predictive value 97.06%, and negative predictive value 90.91%

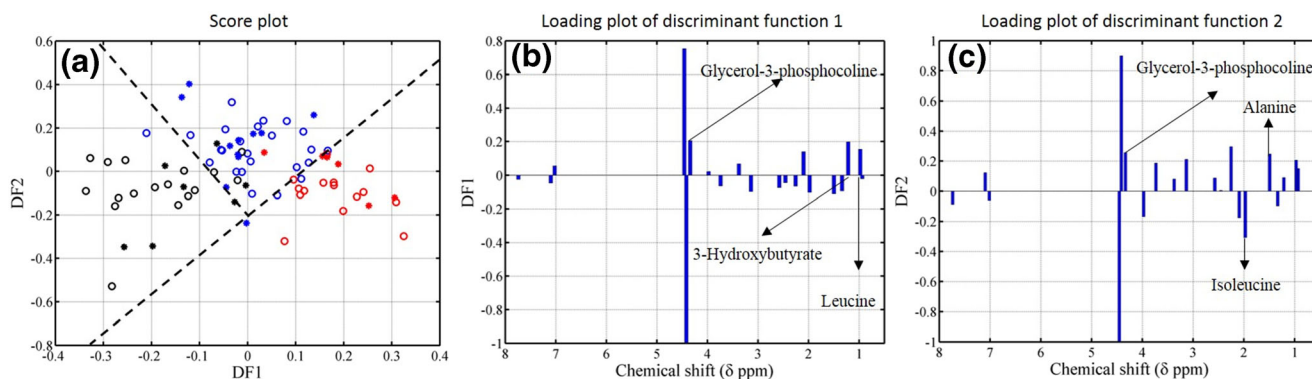
alterations suggest that varicocele causes changes in the metabolism of amino acids, sugars, and its derivatives, as a homeostatic response.

Valine, leucine, and isoleucine are essential amino acids; changes in their seminal concentrations have been shown in infertile men [46–48]. In this study, the C group had lower levels of valine and higher levels of isoleucine when compared with the VF group, and in the VI, the concentration of these amino acids fell in the middle range. These findings suggest a disruption of the amino acid metabolism in men with varicocele, but the pathways linking their metabolism and male infertility remain unclear. 2-Hydroxy-3-methylvalerate is an organic acid generated by isoleucine metabolism and was selected by the GA-LDA formalism as one of the discriminant metabolites, reinforcing the evidence of an altered amino acid metabolism in men with varicocele.

Alanine is a non-essential aminoacidic produced by alanine aminotransferase and has a pivotal role in energy metabolism through the glucose-alanine cycle. Seminal levels of alanine and seminal alanine aminotransferase activity are decreased in infertile men, which could be an indication of impaired sperm energy metabolism [49–52]. Arginine is found in high concentrations in human semen and is a precursor for the synthesis of nitric oxide, a messenger molecule that is essential for normal spermatogenesis and motility [28, 53]. Herein, we observed that infertile men with varicocele had lower levels of arginine, suggesting that the homeostatic response involves an increase of nitric oxide production, consuming arginine. Tyrosine is important in protein synthesis and is a precursor

**Fig. 4** Results of OPLS-DA model to discriminate the varicocele infertile group (VI), 35 samples (red dots), from the varicocele fertile group (VF), 21 samples (green dots). Score plot (left) and permutation test (right)





**Fig. 5** **a** Score plot of GA-LDA to discriminate varicocele infertile (blue), varicocele fertile (red), control (black) groups. Training samples (balls) and test samples (dots). The importance of each variable selected for the

model are shown by the loading of discriminant functions 1 and 2 in **b** and **c** respectively

for synthesis of dopamine and noradrenaline, which have an effect in sperm motility [54, 55]. On the other hand, *n*-acetyltyrosine is derived from tyrosine and is mainly found as a human urinary metabolite. There are reports indicating that seminal tyrosine levels are altered in infertile men [46, 50]. Despite some evidence towards tyrosine having an antioxidant function, a direct link between tyrosine and male fertility is still missing [56]. We identified lower seminal levels of tyrosine in men with varicocele than in healthy volunteers. Moreover, when we compared the VF group to the VI group, varicocele fertile men had lower concentrations of tyrosine. Zhao et al. and Mumcu et al. have demonstrated that seminal glutamine level is an indication of male infertility, but with levels varying among different categories of infertile men [46, 47]. In our study, healthy fertile men had higher levels of glutamine when compared with fertile men with varicocele, while the levels in infertile men with varicocele fell in the intermediate range between the other two groups.

The second group of metabolites identified is associated to carbohydrate metabolism. Sugars and its derivatives can be found in high concentrations in human seminal plasma, where they play several functions, such as energy source and cell

signaling. These metabolites were important variables for group discrimination. Glucose is an energy source for mature sperm, and the lack of this substrate can impair sperm motility, hyperactivation, and acrosome reaction [57]. Glucose was one of the most important variables for group discrimination in the GA-LDA model, a finding that highlights its crucial role in male fertility. Surprisingly, no other study has identified glucose as a discriminant metabolite for male infertility. Citrate is the first product of the Krebs cycle and is stored in large amounts by the glandular epithelial cells in the peripheral region of the human prostate. Citrate has been advocated as a biomarker of prostate function and prostate cancer [58]. Citrate concentration is the main responsible for seminal pH, but the association between seminal citrate levels and male fertility is controversial [59, 60]. In addition to other studies using  $^1\text{H}$  NMR spectroscopy of human seminal plasma, we also found that citrate was a pivotal factor for group segregation [46–48]. Lactate is the final product of the glycolytic pathway, an important step for adenosine triphosphate (ATP) production in mature sperm cells. Moreover, lactate is the only energy source for developing germ cells, since these cells are unable to metabolize glucose and rely on Sertoli cell-produced lactate to produce ATP [57, 61]. We found that

**Table 5** Confusion matrix—GA-LDA metabolomics model to discriminate the 3 groups. Training set: Control group 17 samples, Varicocele fertile group 15 samples, and Varicocele infertile group 24 samples

Model	Groups		
	Control (C)	Varicocele fertile (VF)	Varicocele infertile (VI)
Control	14	0	1
Varicocele fertile	0	15	2
Varicocele infertile	3	0	21

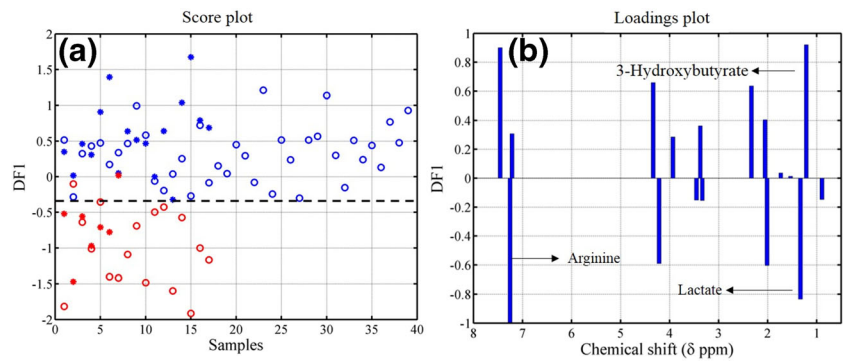
Accuracy 92.17%, sensitivity (82.35% for C, 100.00% for VF and 87.50% for VI), specificity (97.44% for C, 95.12% for VF and 90.62% for VI)

**Table 6** Confusion matrix—GA-LDA metabolomics model to discriminate the 3 groups. Test set: Control group 7 samples, Varicocele fertile group 6 samples, and Varicocele infertile group 11 samples.

Model	Groups		
	Control (C)	Varicocele fertile (VF)	Varicocele infertile (VI)
Control	5	0	0
Varicocele fertile	0	5	1
Varicocele infertile	2	1	10

Accuracy 86.17%, sensitivity (71.43% for C, 83.33% for VF and 90.91% for VI), specificity (100.00% for C, 94.44% for VF and 76.92% for VI)

**Fig. 6** **a** Score plot of GA-LDA to discriminate Varicocele group (blue) from control group (red); training samples (balls) and test samples (dots). The importance of each variable selected is shown by the loading of discriminant function 1 in **b**



lactate was one of the most important metabolites for group segregation, confirming the findings of several other studies which had used <sup>1</sup>H NMR spectroscopy to study seminal plasma [46–48]. However, in these studies, lactate levels were decreased in oligospermic men with idiopathic infertility, whereas we found decreased levels of lactate in healthy fertile men when compared with men with varicocele. Thus, it seems that glucose metabolism is altered in infertile men, but in different ways depending on the cause of infertility.

Although glucose is the main energy source for testicular germ cells and sperm, ketone bodies may play a relevant role in sperm motility [62]. As a ketone body, 3-hydroxybutyrate may affect sperm motility. In addition, it also has a signaling function, being able to inhibit histone deacetylases enzymes [63]. The inhibition of these enzymes has been shown to improve glucose metabolism and to promote resistance to oxidative stress and could be important for spermatogenesis and sperm function [64]. Since varicocele results in oxidative stress, the high seminal levels of 3-hydroxybutyrate observed in varicocele groups indicate that this is the homeostatic response of a live system [28].

Gamma-aminobutyric acid (GABA) is the most important inhibitory neurotransmitter in humans. However, there is a paucity of data regarding the functions that GABA may exert on spermatogenesis and sperm. A few studies have demonstrated the presence of GABA and its receptors in human semen and sperm [65–67]. Furthermore, it has been shown that GABA is able to induce sperm hyperactivation and acrosome reaction [66, 68]. GABA and glutamine are two possible

products of glutamate metabolism. We found that seminal levels of these two metabolites varied symmetrically, indicating that the formation of GABA is preferential in men with varicocele. Sperm count is possibly also important in this context, since in the VI group, where we found low sperm concentration, we also observed an intermediate GABA seminal relative concentration. Caprate is a medium-chain fatty acid composed of 10 carbons found frequently in dairy products and may function as a membrane stabilizer [69]. In patients with varicocele, oxidative stress causes damages to sperm membrane. This requires consumption of caprate to restore the cell membrane, resulting in low caprate concentrations in this group.

Uridine is one of the five standard nucleosides which make up nucleic acids, but the role of seminal uridine is still unclear. Uridine may be a precursor of metabolites required for capacitation, but could also play a role in lipid and carbohydrate metabolism [70, 71]. Zhang et al. demonstrated that seminal uridine is an important discriminant metabolite in men with asthenozoospermia [70]. Our work also revealed that it is a discriminant variable, with a possible higher consumption of uridine in men with varicocele. Methanol is a toxic alcohol found in small amounts in healthy individuals, derived mainly from the ingestion of ripe fruits and by the metabolization of pectin, as a structural component of cell walls of plants [72]. The GA-LDA model using the 3 groups selected methanol as an important variable for group discrimination; however, we were unable to find any study about the possible functions of methanol in human semen.

**Table 7** Confusion matrix—GA-LDA metabolomics model. Training set: control group (17 samples) versus varicocele group (39 samples)

Model	Groups	
	Control (C)	Varicocele (V)
Control	16	0
Varicocele	1	39

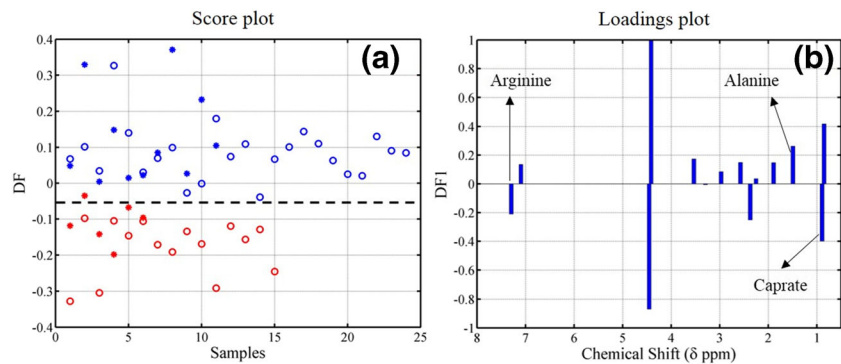
Accuracy 97.06%, sensitivity (94.12% for C and 100.00% for V), specificity (100.00% for C and 94.12% for V)

**Table 8** Confusion matrix – GA-LDA metabolomics model. Test set: control group (7 samples) versus varicocele group (17 samples)

Model	Groups	
	Control (C)	Varicocele (V)
Control	6	0
Varicocele	1	17

Accuracy 92.86%, sensitivity (85.71% for C and 100.00% for V), specificity (100.00% for C and 85.71% for V)

**Fig. 7 a** Score plot of GA-LDA to discriminate the varicocele infertile group (blue) from the varicocele fertile group (red); training samples (balls) and test samples (dots). The importance of each variable selected is shown by the loading of discriminant function 1 in **b**



Glycerol-3-phosphocoline is a cell membrane component and plays a role in cholesterol transport and metabolism. Some studies using  $^1\text{H}$  NMR spectroscopy have shown abnormal seminal levels of GPC in infertile men with altered SA parameters [46, 70, 73, 74]. GPC was an important discriminant variable in the GA-LDA model using the 3 groups, suggesting an altered phospholipid metabolism in men with varicocele.

Other researchers have demonstrated that it is possible to use the metabolomics approach to create models capable of diagnosing certain phenotypes related to male infertility with high accuracy. A study using  $^1\text{H}$  NMR spectroscopy of seminal plasma was able to generate models with accuracy of 92.4% to discriminate fertile men from infertile, and 92.2% to differentiate infertile men with normal SA from infertile men with oligospermia [50]. Another similar study constructed OPLS-DA models capable of segregating men with idiopathic infertility (with normal SA) from infertile men with oligospermia with accuracy of 89% [49]. Also using  $^1\text{H}$  NMR of seminal plasma and OPLS-DA, another group created a model that discriminated fertile men from men with asthenozoospermia with an accuracy of 85.7% [70]. In addition, using  $^1\text{H}$  NMR spectra of seminal plasma, a different group was able to build PCA and PLS-DA models to segregate normozoospermic from oligoasthenoteratozoospermic men [46].

The main limitation of the present study was the small number of samples used in this pilot study. Therefore, future

studies with larger sample sizes are required to confirm our findings. Furthermore, the evaluation of infertile men's partners was made by interview only, and no additional tests were performed to identify possible undiagnosed causes of female infertility. This means that we could not definitely exclude female fertility factors, and they may have been a confusion factor in our analyses. However, the young age of the partners and the absence of symptoms or previous history of conditions that could cause infertility decrease the likelihood of the presence of female factor infertility.

One of the strengths of our study was the use of the infertility definition suggested by the American Society for Human Reproduction, which is currently considered the best definition to be applied in research [34]. Moreover, all participants were evaluated by the same specialist in male infertility, and all USGD were performed by the same genitourinary radiology specialist, as well all laboratory tests were done in a central laboratory, aspects that reduce the chance of bias. It should be emphasized that there was no change in the clinical management of the participants and that all samples collected and clinical exams performed are already part of the initial evaluation of infertile men, factors that facilitate the incorporation of  $^1\text{H}$  NMR spectroscopy in the clinical context. The metabolomics models presented here can be very useful in clinical practice, helping in the management of men who have varicocele and want to know about their fertility potential even before they have begun the attempts to conceive.

**Table 9** Confusion matrix—GA-LDA metabolomics model. Training set: varicocele infertile group (24 samples) versus varicocele fertile group (15 samples)

Model	Groups	
	Varicocele infertile (VI)	Varicocele fertile (VF)
Varicocele infertile	24	0
Varicocele fertile	0	15

Accuracy 100%, sensitivity (100.00% for VI and 100% for VF), specificity (100% for VI and 100.00% for VF)

**Table 10** Confusion matrix—GA-LDA metabolomics model. Test set varicocele infertile group (11 samples) versus varicocele fertile group (6 samples).

Model	Groups	
	Varicocele infertile (VI)	Varicocele fertile (VF)
Varicocele infertile	11	1
Varicocele fertile	0	5

Accuracy 91.67%, sensitivity (100.00% for VI and 83.33% for VF), specificity (83.33% for VI and 100.0% for VF)

## Conclusion

The present study demonstrated the  $^1\text{H}$  NMR spectroscopy of seminal plasma can be used in conjunction with multivariate statistical tools to create metabonomics models useful in discriminating between men with and without varicocele, and between fertile and infertile men with varicocele. In addition, the most important metabolites for group segregation are involved in the oxidative stress caused by varicocele and its response, highlighting the importance of these mechanisms in the pathogenesis of varicocele. Further studies are needed to confirm these findings.

**Authors' contributions** FTLN, RAM, LCNA, LP, and ROS conceived and implemented the study. FTLN, RAM, and AFCF collected and performed analysis. FTLN, RAM, SVCL, LP, and ROS interpreted the data and drafted the manuscript. FTLN, RAM, SVCL, LP, and ROS critically revised the manuscript. All authors read and approved the final manuscript.

**Funding information** This research was supported by the Pernambuco Science and Technology Foundation - FACEPE (process number APQ-0765-1.06/14) and by National Council for Scientific and Technological Development (CNPq, process 306616/2013-0).

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflicts of interest.

**Ethics approval** This study was approved by the Research Ethics Committee of Universidade Federal de Pernambuco and Instituto de Medicina Integral Prof. Fernando Figueira (approval n. 2.075.028).

## References

- Gorelick JJ, Goldstein M. Loss of fertility in men with varicocele. *Fertil Steril*. 1993;59(3):613–6.
- WHO. The influence of varicocele on parameters of fertility in a large group of men presenting to infertility clinics. *World Health Organization. Fertil Steril*. 1992;57(6):1289–93.
- Pasqualotto FF, Lucon AM, de Goes PM, Sobreiro BP, Hallak J, Pasqualotto EB, et al. Semen profile, testicular volume, and hormonal levels in infertile patients with varicoceles compared with fertile men with and without varicoceles. *Fertil Steril*. 2005;83(1):74–7. <https://doi.org/10.1016/j.fertnstert.2004.06.047>.
- Zargooshi J. Sperm count and sperm motility in incidental high-grade varicocele. *Fertil Steril*. 2007;88(5):1470–3. <https://doi.org/10.1016/j.fertnstert.2007.01.016>.
- Sigman M. There is more than meets the eye with varicoceles: current and emerging concepts in pathophysiology, management, and study design. *Fertil Steril*. 2011;96(6):1281–2. <https://doi.org/10.1016/j.fertnstert.2011.10.026>.
- Goldstein M, Gilbert BR, Dicker AP, Dwosh J, Gnecco C. Microsurgical inguinal varicocelectomy with delivery of the testis: an artery and lymphatic sparing technique. *J Urol*. 1992;148(6):1808–11.
- Marmar JL, Agarwal A, Prabakaran S, Agarwal R, Short RA, Benoff S, et al. Reassessing the value of varicocelectomy as a treatment for male subfertility with a new meta-analysis. *Fertil Steril*. 2007;88(3):639–48. <https://doi.org/10.1016/j.fertnstert.2006.12.008>.
- Li F, Yue H, Yamaguchi K, Okada K, Matsushita K, Ando M, et al. Effect of surgical repair on testosterone production in infertile men with varicocele: a meta-analysis. *International journal of urology : official journal of the Japanese Urological Association*. 2012;19(2):149–54. <https://doi.org/10.1111/j.1442-2042.2011.02890.x>.
- ASRM. Diagnostic evaluation of the infertile male: a committee opinion. *Fertil Steril*. 2015;103(3):e18–25. <https://doi.org/10.1016/j.fertnstert.2014.12.103>.
- van der Steeg JW, Steures P, Eijkemans MJ, Habbema JDF, Hompes PG, Kremer JA, et al. Role of semen analysis in subfertile couples. *Fertil Steril*. 2011;95(3):1013–9. <https://doi.org/10.1016/j.fertnstert.2010.02.024>.
- Karczewski KJ, Snyder MP. Integrative omics for health and disease. *Nat Rev Genet*. 2018;19(5):299–310. <https://doi.org/10.1038/nrg.2018.4>.
- Kovac JR, Pastuszak AW, Lamb DJ. The use of genomics, proteomics, and metabolomics in identifying biomarkers of male infertility. *Fertil Steril*. 2013;99(4):998–1007. <https://doi.org/10.1016/j.fertnstert.2013.01.111>.
- Mayr M. Metabolomics: ready for the prime time? *Circ Cardiovasc Genet*. 2008;1(1):58–65. <https://doi.org/10.1161/CIRCGENETICS.108.808329>.
- 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(11):1181–9. <https://doi.org/10.1080/004982599238047>.
- Godoy MM, Lopes EP, Silva RO, Hallwass F, Koury LC, Moura IM, et al. Hepatitis C virus infection diagnosis using metabonomics. *J Viral Hepat*. 2010;17(12):854–8. <https://doi.org/10.1111/j.1365-2893.2009.01252.x>.
- Dunn WB, Ellis DI. Metabolomics: current analytical platforms and methodologies. *TrAC Trends Anal Chem*. 2005;24(4):285–94. <https://doi.org/10.1016/j.trac.2004.11.021>.
- de Azevedo WM, de Oliveira Luna AJ, Silva EF, Silva RO. The effect of ultrasonic waves in conducting polymer solution. *Ultrason Sonochem*. 2006;13(5):433–7. <https://doi.org/10.1016/j.ultsonch.2005.10.002>.
- Silva RO, Neves Filho RAW, Azevedo R, Srivastava RM, Gallardo H. Complete  $^1\text{H}$  and  $^{13}\text{C}$  NMR signal assignments and chemical shift calculations of four 1,2,4-oxadiazole-based light-emitting liquid crystals. *Struct Chem*. 2010;21(3):485–94. <https://doi.org/10.1007/s11224-009-9576-z>.
- Pontes TA, Barbosa AD, Silva RD, Melo-Junior MR, Silva RO. Osteopenia-osteoporosis discrimination in postmenopausal women by  $^1\text{H}$  NMR-based metabonomics. *PLoS One*. 2019;14(5):e0217348. <https://doi.org/10.1371/journal.pone.0217348>.
- Robertson DG. Metabonomics in toxicology: a review. *J Toxicol Sci*. 2005;85(2):809–22. <https://doi.org/10.1093/toxsci/kfi102>.
- Worley B, Powers R. Multivariate analysis in metabolomics. *Curr Metab*. 2013;1(1):92–107. <https://doi.org/10.2174/2213235x11301010092>.
- Trygg J, Lundstedt T. Chemometrics Techniques for Metabonomics. In: Lindon JC, Nicholson JK, Holmes E, editors. *The Handbook of metabonomics and metabolomics*. Amsterdam: Elsevier BV; 2007.
- Gromski PS, Muhamadali H, Ellis DI, Xu Y, Correa E, Turner ML, et al. A tutorial review: metabolomics and partial least squares-discriminant analysis—a marriage of convenience or a shotgun wedding. *Anal Chim Acta*. 2015;879:10–23. <https://doi.org/10.1016/j.aca.2015.02.012>.
- Szymańska E, Saccenti E, Smilde AK, Westerhuis JA. Double-check: validation of diagnostic statistics for PLS-DA models in

- metabolomics studies. *Metabolomics*. 2012;8(Suppl 1):3–16. <https://doi.org/10.1007/s11306-011-0330-3>.
25. Bylesjö M, Rantalainen M, Cloarec O, Nicholson JK, Holmes E, Trygg J. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J Chemom*. 2006;20(8–10):341–51. <https://doi.org/10.1002/cem.1006>.
  26. Brereton RG, Lloyd GR. Partial least squares discriminant analysis: taking the magic away. *J Chemom*. 2014;28(4):213–25. <https://doi.org/10.1002/cem.2609>.
  27. Jarvis RM, Goodacre R. Genetic algorithm optimization for pre-processing and variable selection of spectroscopic data. *Bioinformatics (Oxford, England)*. 2005;21(7):860–8. <https://doi.org/10.1093/bioinformatics/bti102>.
  28. Agarwal A, Hamada A, Esteves SC. Insight into oxidative stress in varicocele-associated male infertility: part 1. *Nat Rev Urol*. 2012;9(12):678–90. <https://doi.org/10.1038/nrurol.2012.197>.
  29. Agarwal A, Prabakaran S, Allamaneni SS. Relationship between oxidative stress, varicocele and infertility: a meta-analysis. *Reprod BioMed Online*. 2006;12(5):630–3.
  30. Agarwal A, Sharma R, Durairajanayagam D, Ayaz A, Cui Z, Willard B, et al. Major protein alterations in spermatozoa from infertile men with unilateral varicocele. *Reprod Biol Endocrinol : RB&E*. 2015;13:8. <https://doi.org/10.1186/s12958-015-0007-2>.
  31. Camargo M, Intasqui P, Bertolla RP. Proteomic profile of seminal plasma in adolescents and adults with treated and untreated varicocele. *Asian J Androl*. 2016;18(2):194–201. <https://doi.org/10.4103/1008-682X.168788>.
  32. Del Giudice PT, da Silva BF, Lo Turco EG, Fraietta R, Spaine DM, Santos LF, et al. Changes in the seminal plasma proteome of adolescents before and after varicocelectomy. *Fertil Steril*. 2013;100(3):667–72. <https://doi.org/10.1016/j.fertnstert.2013.04.036>.
  33. Zylbersztein DS, Andreoni C, Del Giudice PT, Spaine DM, Borsari L, Souza GH, et al. Proteomic analysis of seminal plasma in adolescents with and without varicocele. *Fertil Steril*. 2013;99(1):92–8. <https://doi.org/10.1016/j.fertnstert.2012.08.048>.
  34. ASRM. Definitions of infertility and recurrent pregnancy loss: a committee opinion. 2013(1556–5653 (Electronic)).
  35. Dubin L, Amelar RD. Varicocelectomy: 986 cases in a twelve-year study. *Urology*. 1977;10(5):446–9.
  36. WHO. Laboratory Manual for the Examination and Processing of Human Semen. 5th ed; 2010.
  37. Vermeulen A, Verdonck L, Kaufman JM. A critical evaluation of simple methods for the estimation of free testosterone in serum. *J Clin Endocrinol Metab*. 1999;84(10):3666–72. <https://doi.org/10.1210/jcem.84.10.6079>.
  38. Chong J, Soufan O, Li C, Caraus I, Li S, Bourque G, et al. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res*. 2018;46(W1):W486–W94. <https://doi.org/10.1093/nar/gky310>.
  39. Kennard RW, Stone LA. Computer aided design of experiments. *Technometrics*. 1969;11(1):137–48. <https://doi.org/10.2307/1266770>.
  40. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vazquez-Fresno R, et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res*. 2018;46(D1):D608–d17. <https://doi.org/10.1093/nar/gkx1089>.
  41. Paiva C, Amaral A, Rodriguez M, Canyellas N, Correig X, Balleca JL, et al. Identification of endogenous metabolites in human sperm cells using proton nuclear magnetic resonance ((1) H-NMR) spectroscopy and gas chromatography-mass spectrometry (GC-MS). *Andrology*. 2015;3(3):496–505. <https://doi.org/10.1111/andr.12027>.
  42. Milewicz A, Krzyzanowska-Swiniarska B, Miazgowski T, Jedrzejuk D, Arkowska A, Mieszczanowicz U, et al. The reference values of sex hormones and SHBG serum levels in subjects over 65 years old - the PolSenior study. *Endokrynol Pol*. 2013;64(2):82–93.
  43. Trost LW, Mulhall JP. Challenges in testosterone measurement, data interpretation, and methodological appraisal of interventional trials. *J Sex Med*. 2016;13(7):1029–46. <https://doi.org/10.1016/j.jsxm.2016.04.068>.
  44. Del Giudice PT, Belardin LB, Camargo M, Zylbersztein DS, Carvalho VM, Cardozo KH, et al. Determination of testicular function in adolescents with varicocele - a proteomics approach. *Andrology*. 2016;4(3):447–55. <https://doi.org/10.1111/andr.12174>.
  45. Chen SS. Significant predictive factors for subfertility in patients with subclinical varicocele. *Andrologia*. 2017;49(10). <https://doi.org/10.1111/and.12781>.
  46. Mumcu A, Karaer A, Dogan B, Tuncay G. Metabolomics analysis of seminal plasma in patients with idiopathic Oligoasthenoatozoospermia using high-resolution NMR spectroscopy. *Andrology*. 2020;8(2):450–6. <https://doi.org/10.1111/andr.12707>.
  47. Zhao K, Zhang J, Xu Z, Xu Y, Xu A, Chen W, et al. Metabolomic profiling of human spermatozoa in idiopathic asthenozoospermia patients using gas chromatography-mass spectrometry. *Biomed Res Int*. 2018;2018:8327506–8. <https://doi.org/10.1155/2018/8327506>.
  48. Qiao S, Wu W, Chen M, Tang Q, Xia Y, Jia W, et al. Seminal plasma metabolomics approach for the diagnosis of unexplained male infertility. *PLoS One*. 2017;12(8):e0181115. <https://doi.org/10.1371/journal.pone.0181115>.
  49. Jayaraman V, Ghosh S, Sengupta A, Srivastava S, Sonawat HM, Narayan PK. Identification of biochemical differences between different forms of male infertility by nuclear magnetic resonance (NMR) spectroscopy. *J Assist Reprod Genet*. 2014;31(9):1195–204. <https://doi.org/10.1007/s10815-014-0282-4>.
  50. Gupta A, Mahdi AA, Ahmad MK, Shukla KK, Jaiswar SP, Shankwar SN. 1H NMR spectroscopic studies on human seminal plasma: a probative discriminant function analysis classification model. *J Pharm Biomed Anal*. 2011;54(1):106–13. <https://doi.org/10.1016/j.jpba.2010.07.021>.
  51. Gershbein LL, Thielen DR. Enzymatic and electrolytic profiles of human semen. *Prostate*. 1988;12(3):263–9. <https://doi.org/10.1002/pros.2990120309>.
  52. Chyra-Jach D, Kaletka Z, Dobrakowski M, Machoń-Grecka A, Kasperczyk S, Bellanti F, et al. Levels of macro- and trace elements and select cytokines in the semen of infertile men. *Biol Trace Elem Res*. 2020. <https://doi.org/10.1007/s12011-019-02022-9>.
  53. Zini A, O'Bryan MK, Schlegel PN. Nitric oxide synthase activity in human seminal plasma. *Urology*. 2001;58(1):85–9. [https://doi.org/10.1016/s0090-4295\(01\)01001-9](https://doi.org/10.1016/s0090-4295(01)01001-9).
  54. Hou Y, Yin Y, Wu G. Dietary essentiality of “nutritionally non-essential amino acids” for animals and humans. *Exp Biol Med (Maywood, NJ)*. 2015;240(8):997–1007. <https://doi.org/10.1177/1535370215587913>.
  55. Lavieri JC, Pierini AA. L-dopa and oligozoospermia. *Andrologia*. 1978;10(1):74–9. <https://doi.org/10.1111/j.1439-0272.1978.tb01320.x>.
  56. van Overveld FW, Haenen GR, Rhemrev J, Vermeiden JP, Bast A. Tyrosine as important contributor to the antioxidant capacity of seminal plasma. *Chem Biol Interact*. 2000;127(2):151–61. [https://doi.org/10.1016/s0009-2797\(00\)00179-4](https://doi.org/10.1016/s0009-2797(00)00179-4).
  57. Dias TR, Alves MG, Silva BM, Oliveira PF. Sperm glucose transport and metabolism in diabetic individuals. *Mol Cell Endocrinol*. 2014;396(1–2):37–45. <https://doi.org/10.1016/j.mce.2014.08.005>.
  58. Gregorio EP, Alexandrino AP, Schuquel ITA, da Costa WF, Rodrigues MAF. Seminal citrate is superior to PSA for detecting clinically significant prostate cancer. *Int Braz J Urol*. 2019;45(6):1113–21. <https://doi.org/10.1590/s1677-5538.Ibju.2018.0730>.

59. Said L, Galeraud-Denis I, Carreau S, Saad A. Relationship between semen quality and seminal plasma components: alpha-glucosidase, fructose and citrate in infertile men compared with a normospermic population of Tunisian men. *Andrologia*. 2009;41(3):150–6. <https://doi.org/10.1111/j.1439-0272.2008.00906.x>.
60. Zopfgen A, Priem F, Sudhoff F, Jung K, Lenk S, Loening SA, et al. Relationship between semen quality and the seminal plasma components carnitine, alpha-glucosidase, fructose, citrate and granulocyte elastase in infertile men compared with a normal population. *Hum Reprod*. 2000;15(4):840–5. <https://doi.org/10.1093/humrep/15.4.840>.
61. Rato L, Alves MG, Socorro S, Duarte AI, Cavaco JE, Oliveira PF. Metabolic regulation is important for spermatogenesis. *Nat Rev Urol*. 2012;9(6):330–8. <https://doi.org/10.1038/nrurol.2012.77>.
62. Tanaka H, Takahashi T, Iguchi N, Kitamura K, Miyagawa Y, Tsujimura A, et al. Ketone bodies could support the motility but not the acrosome reaction of mouse sperm. *Int J Androl*. 2004;27(3):172–7. <https://doi.org/10.1111/j.1365-2605.2004.00471.x>.
63. Gregoretti IV, Lee YM, Goodson HV. Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis. *J Mol Biol*. 2004;338(1):17–31. <https://doi.org/10.1016/j.jmb.2004.02.006>.
64. Newman JC, Verdin E.  $\beta$ -hydroxybutyrate: much more than a metabolite. *Diabetes Res Clin Pract*. 2014;106(2):173–81. <https://doi.org/10.1016/j.diabres.2014.08.009>.
65. Leader A, Minuk GY, Mortimer D. Seminal plasma gamma-aminobutyric acid (GABA) levels in normospermic men. *Clin Invest Med*. 1992;15(4):346–8.
66. Calogero AE, Hall J, Fishel S, Green S, Hunter A, D'Agata R. Effects of gamma-aminobutyric acid on human sperm motility and hyperactivation. *Mol Hum Reprod*. 1996;2(10):733–8. <https://doi.org/10.1093/molehr/2.10.733>.
67. Ritta MN, Calamera JC, Bas DE. Occurrence of GABA and GABA receptors in human spermatozoa. *Mol Hum Reprod*. 1998;4(8):769–73. <https://doi.org/10.1093/molehr/4.8.769>.
68. Wistrom CA, Meizel S. Evidence suggesting involvement of a unique human sperm steroid receptor/Cl<sup>-</sup> channel complex in the progesterone-initiated acrosome reaction. *Dev Biol*. 1993;159(2):679–90. <https://doi.org/10.1006/dbio.1993.1274>.
69. Dahlgren D, Sjoblom M, Hedeland M, Lennernas H. The in vivo effect of transcellular permeation enhancers on the intestinal permeability of two peptide drugs enalaprilat and hexarelin. *Pharmaceutics*. 2020;12(2). <https://doi.org/10.3390/pharmaceutics12020099>.
70. Zhang X, Diao R, Zhu X, Li Z, Cai Z. Metabolic characterization of asthenozoospermia using nontargeted seminal plasma metabolomics. *Clin Chim Acta*. 2015;450:254–61. <https://doi.org/10.1016/j.cca.2015.09.001>.
71. Maher AD, Patki P, Lindon JC, Want EJ, Holmes E, Craggs M, et al. Seminal oligouridinoses: low uridine secretion as a biomarker for infertility in spinal neurotrauma. *Clin Chem*. 2008;54(12):2063–6. <https://doi.org/10.1373/clinchem.2008.112219>.
72. Turner C, Spanel P, Smith D. A longitudinal study of methanol in the exhaled breath of 30 healthy volunteers using selected ion flow tube mass spectrometry, SIFT-MS. *Physiol Meas*. 2006;27(7):637–48. <https://doi.org/10.1088/0967-3334/27/7/007>.
73. Bonechi C, Collodel G, Donati A, Martini S, Moretti E, Rossi C. Discrimination of human semen specimens by NMR data, sperm parameters, and statistical analysis. *Syst Biol Reprod Med*. 2015;61(6):353–9. <https://doi.org/10.3109/19396368.2015.1054003>.
74. Hamamah S, Seguin F, Bujan L, Barthelemy C, Mieusset R, Lansac J. Quantification by magnetic resonance spectroscopy of metabolites in seminal plasma able to differentiate different forms of azoospermia. *Hum Reprod*. 1998;13(1):132–5. <https://doi.org/10.1093/humrep/13.1.132>.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.