



OPEN ACCESS

EDITED BY

Guihua Liu,
The Sixth Affiliated Hospital, Sun
Yat-Sen University, China

REVIEWED BY

Hiroshi Okada,
Dokko Medical University Saitama
Medical Center, Japan
Paula Intasqui,
Federal University of São Paulo, Brazil

*CORRESPONDENCE

Yi Guo
gy_guoyi@126.com
Xianting Jiao
jiaoxianting@xinhuamed.com.cn
Xiaoming Teng
tengxiaoming@hotmail.com

[†]These authors have contributed
equally to this work

SPECIALTY SECTION

This article was submitted to
Reproduction,
a section of the journal
Frontiers in Endocrinology

RECEIVED 30 September 2022

ACCEPTED 28 November 2022

PUBLISHED 04 January 2023

CITATION

Guo Y, Li J, Hao F, Yang Y, Yang H,
Chang Q, Kong P, Liu W, Jiao X and
Teng X (2023) A new perspective on
semen quality of aged male: The
characteristics of metabolomics
and proteomics.
Front. Endocrinol. 13:1058250.
doi: 10.3389/fendo.2022.1058250

COPYRIGHT

© 2023 Guo, Li, Hao, Yang, Yang,
Chang, Kong, Liu, Jiao and Teng. This is
an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use,
distribution or reproduction is
permitted which does not comply with
these terms.

A new perspective on semen quality of aged male: The characteristics of metabolomics and proteomics

Yi Guo^{1*†}, Jinli Li^{1†}, Fengdan Hao^{2†}, Yang Yang¹, Hao Yang¹,
Qirong Chang¹, Pengcheng Kong¹, Wenqiang Liu¹,
Xianting Jiao^{3*} and Xiaoming Teng^{1*}

¹Department of Assisted Reproduction, Shanghai First Maternity and Infant Hospital, School of Medicine, Tongji University, Shanghai, China, ²Department of Pediatrics, Tongji Hospital, School of Medicine, Tongji University, Shanghai, China, ³Department of Pediatric Cardiology, Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China

Background: Semen quality is negatively correlated with male age and is mainly quantified by a routine semen analysis, which is descriptive and inconclusive. Sperm proteins or semen metabolites are used as the intermediate or end-products, reflecting changes in semen quality, and hold much promise as a new biomarker to predict fertility in advanced-aged males.

Objectives: In this study, we sought to assess whether the semen metabolome and proteome of aged males can affect semen quality and serve as biomarkers for predicting semen quality.

Materials and methods: We retrospectively analyzed 12825 males that underwent semen routine analysis to understand the age-dependent changes in sperm quality. To identify the difference between aged and young adults, metabolomics (n=60) analyses of semen and proteomics (n=12) analyses of sperm were conducted. Finally, integrated machine learning of metabolomics was conducted to screen biomarkers to identify aging semen.

Results: We discovered that male age was positively correlated with sperm concentration as well as DNA fragmentation index (DFI), and negatively with progressive motile sperm count, total sperm count, sperm volume and progressive sperm motility. The differential metabolites were significantly enriched in various metabolic pathways, and four of these differential metabolites (Pipamperone, 2,2-Bis(hydroxymethyl)-2,2',2''-nitrotriethanol, Arg-Pro and Triethyl phosphate) were utilized to establish a biomarker panel to identify aging semen. Proteomic analysis showed that differential proteins were significantly enriched in protein digestion and absorption and some energy-related pathways. An integrated analysis of the metabolome and proteome identified differential energy metabolism and oxidative stress-related proteins, which could explain the decreased motility and the increased DFI of aging sperm

Discussion and conclusion: We provide compelling evidence that the changes in semen metabolome and sperm proteome are related to the decline of semen quality in aged males. Moreover, a biomarker panel based on four metabolites was established to identify aging semen.

KEYWORDS

advanced male age, metabolome, proteome, sperm quality, biomarker

1 Introduction

It is well-established that advanced maternal age is related to lower fertility and birth defects (1). Growing evidence substantiates that advanced male age exerts deleterious effects on semen quality, assisted reproductive outcomes, and offspring health (2). In this regard, advanced paternal age has been related to a higher risk of diseases in offspring, including achondroplasia, autism, schizophrenia, bipolar disorders, cancers and impaired cognitive abilities (3). Given that the exact mechanisms remain largely unknown, exploring the biological characteristics of aging semen and finding explanations for the decline of fertility and offspring health of advanced males is essential.

Routine semen analysis remains the main indicator for evaluating semen quality, including semen volume, pH, sperm count, motility, vitality, and morphology (4). In recent years, researchers have made vast strides with the discovery of more advanced semen tests to evaluate sperm function, including oxidative stress and DNA fragmentation index (DFI), for further evaluation of sperm function. Although semen analyses have identified several different parameters with different predictive values for sperm function, these parameters are often descriptive and can not provide a more etiological explanation. Accordingly, the application of routine semen analysis to diagnose male fertility has been controversial for many years (5, 6). In recent years, omics, including genomic, proteomic and metabolomic techniques, have been used to analyze molecular-level information of samples to diagnose or treat male infertility (7). Indeed, in-depth analyses of sperm molecular composition and data mining have contributed to our understanding of sperm function. A previous study identified 103 sperm proteins that play a key role in the fertilization, including sperm penetration, sperm acrosome reaction, and sperm-oocyte hybridization (8). Meanwhile, a comparative proteome analysis between asthenozoospermia and normozoospermia yielded eight differentially expressed proteins (DEPs), mainly related to energy and metabolism (9). More interestingly, one sperm protein CATSPER1, the sperm-

specific calcium channel, may be considered as a valuable tool to predict ART success (10). These DEPs are potentially more effective and reliable indicators to evaluate male fertility with respect to sperm function and fertilization. On the other hand, metabolites are thought to be the end products of cellular regulatory processes and biological systems following genetic or environmental changes (11). Compared with genes and proteins, endogenous metabolites are the closest to the phenotype and more easily detected (11). The seminal metabolome has recently become a research hotspot and is considered as a promising tool to uncover biomarkers of semen quality and fertilizing capacity (12). It has been found that 63 metabolites are potential biomarkers of male infertility, of which 17 are correlated with semen parameters (13). Reduced levels of metabolites in seminal plasma, including amino acids, lactate, citrate, creatinine, α -ketoglutaric acid, spermine and putrescine, have also been used to distinguish oligoasthenoteratozoospermia from normozoospermic controls (14). In addition, the metabolic profile of seminal plasma was used to predict spermatogenesis in nonobstructive azoospermic (NOA) patients in assisted reproduction (15). Taken together, the above findings suggest that applying proteomics and metabolomics in andrology can help overcome the limitations of conventional semen analysis.

It is widely acknowledged that aging affects a myriad of genetic, biochemical, and metabolic processes (16). However, no study has hitherto reported the effect of aging on semen proteomics and metabolomics. It remains unclear whether proteomics and metabolomics analyses can explain the decline of semen quality in aged males. Because semen is mainly composed of secretions from accessory glands, it contains a large number of metabolites, and is easier to be detected for being biomarkers. In addition, in the process of sperm-oocyte binding, sperm proteins are more involved in the fertilization and suitable for exploring sperm function. Herein, we sought to establish the profiles of semen metabolome and sperm proteome to identify age-related changes in metabolites and proteins between aged and young men. Meanwhile, we established a biomarker panel to predict aging semen.

2 Material and methods

2.1 Study design and patients

We retrospectively analyzed the data of 12828 sperm samples collected from males that underwent routine semen examinations and DFI tests from January 2020 to December 2021 at the Department of Assisted Reproduction of Shanghai First Maternity and Infant Hospital. All participants were categorized into five groups according to their age as follows: group A ($\leq 30y$), group B (31-35y), group C (36-40y), group D (41-45y) and group E ($\geq 46y$). Patients were excluded for the following reasons: history of sexually transmitted diseases, chromosome abnormalities, mumps orchitis, varicocele, testicular surgery, epididymal surgery, urethral surgery and vasectomy.

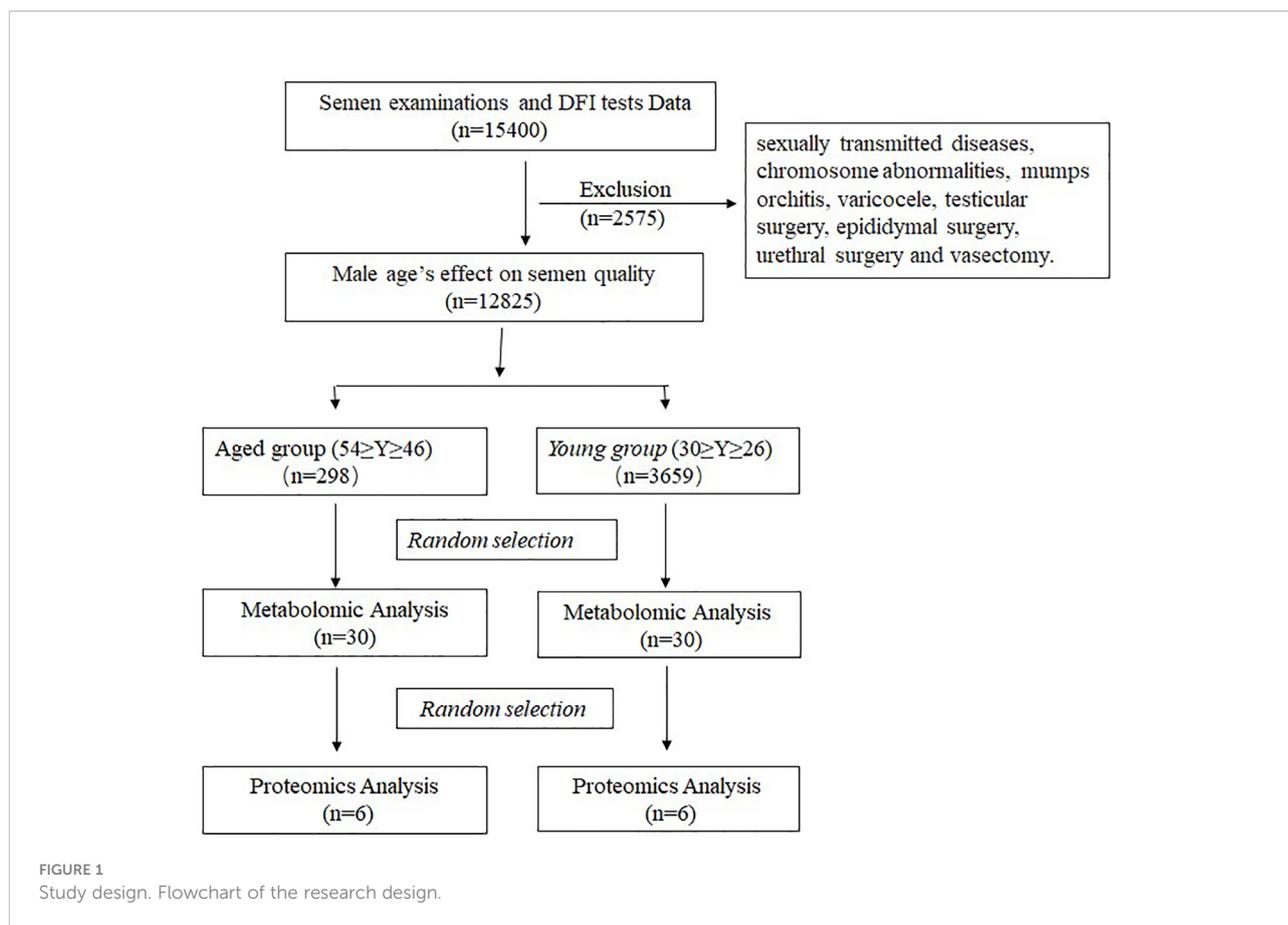
60 participants were enrolled in the metabolomics analysis, of which 12 participants were also enrolled in the proteomics analysis. Participants aged 45-54 were included in the aged semen (AS) group, while participants aged 26-30 comprised the young semen (YS) group. The exclusion criteria were as follow: 1) metabolic diseases (diabetes and cardiovascular diseases); 2) long-term exposure to toxins, tobacco, alcohol or drugs; 3) other known causes male infertility (genetic disease and

infection). The Shanghai First Maternity and Infant Hospital's medical ethics committee approved the study, and all participants provided written informed consent. The study design is illustrated in **Figure 1**.

2.2 Semen collection, analysis and purification

Semen samples were obtained by masturbation after 2-7 days of abstinence. After 30 minutes of liquefaction in an incubator (37°C), the following parameters, including semen volume, pH, sperm concentration, motility, total sperm count and progressive motile sperm count, were evaluated according to the World Health Organization guidelines. Semen volume was measured by weighing, and sperm motility and concentration were analyzed by computer-assisted sperm analysis (CASA) (Hamilton Thorne, USA). At the same time, part of the semen samples were used for DFI testing and the detection of sperm DFI was performed according to the kit's instructions (CellPro Biotech, China), which was based on the sperm chromatin structure assay (SCSA) method.

Sperm were isolated by the density gradient method. In brief, 2-3 mL of liquefied semen samples was carefully overlaid onto a



two-layer (40–80%) Percoll gradient (1:1:1, v/v/v). After centrifugation at 300×g for 15 mins, the supernatant was discarded and the sperm pellet was washed twice with PBS. Then the purified sperm was examined under a microscope (Olympus, Japan) to ensure no lymphocytes, epithelial cells and bacteria.

2.2 Metabolomic analysis

For metabolite extraction, 1 mL of metabolite extraction buffer consisting of methanol/acetonitrile/H₂O (2:2:1, v/v/v) was added to the semen sample and vortexed. The mixture was collected into a 1.5 mL centrifuge tube and centrifuged at 14000×g for 10 min to remove proteins. The supernatant was dried up and re-dissolved in a solvent containing acetonitrile/water (1:1, v/v), and then separated using liquid chromatography, UHPLC (1290 Infinity LC, Agilent Technologies) and subjected to quadrupole time-of-flight mass spectrometry (AB Sciex TripleTOF 6600). Briefly, a 2.1 mm × 100 mm ACQUITY UPLC BEH 1.7 μm column (Waters, Ireland) was used to separate the sample solution. Meanwhile, the mobile phase in both ESI positive and negative modes contained A=25 mM ammonium acetate and 25 mM ammonium hydroxide in water and B= acetonitrile. The gradient was as follows: 0–0.5 min, 95% B; 0.5–7 min, 95%–65% B; 7–8 min, 65%–40% B; 8–9 min, 40% B; 9–9.1 min, 40%–95% B; 9.1–12 min, 95% B. A random injection sequence was employed to avoid signal fluctuations during sample analysis. The quality control (QC) samples were inserted to ensure a reliable data analysis. In MS-only acquisition, the m/z range was set to 60–1000 Da, and the accumulation time for the TOF MS scan was set to 0.20 s/spectra. In the auto MS/MS acquisition, the m/z range was set to 25–1000 Da, and the accumulation time for the product ion scan was set to 0.05 s/spectra. The information-dependent acquisition (IDA) mode was used to acquire the product ion scan. The following parameters were set: the collision energy (CE) was set at 35 V ± 15 eV; declustering potential (DP) was set to ±60 V (± ESI); isotopes within 4 Da were excluded and 10 candidate ions were monitored per cycle. Metabolomic Analysis was conducted by Shanghai Applied Protein Technology Co.

Only variables with more than 50% nonzero measurements in one or more groups were retained during the extraction of ion features in the metabolome. Metabolite analysis was carried out by comparing their m/z accuracy (<25 ppm) and MS/MS spectra to an in-house database based on well-established standards. After sum normalization, an R package was used to perform orthogonal partial least-squares discriminant analysis (OPLS-DA). Pearson's correlation analysis was used to determine the relationship between two variables. The KEGG database was used for pathway enrichment analysis (<http://geneontology.org/>).

2.3 Proteomics analysis

Sperm samples (six replicates for each group) were isolated by the density gradient method and washed twice with PBS for the proteomics study. 2×10⁶ sperm for each sample were lysed by 200 μL SDT buffer (3×volume of sperm pellet) containing 4% SDS, 150 mM Tris-HCl pH 8.0 and 100 mM DTT. The lysates were further sonicated and boiled for 10 min. After centrifugation at 14000g for 30 min at 4°C, the supernatant was quantified with the BCA Protein Assay Kit (Bio-Rad, USA). 15 μg protein from each sample was individually separated by one-dimensional (1D) Sodium dodecyl sulfate (SDS) Polyacrylamide gel electrophoresis (PAGE) and then individually digested using in-gel trypsin for sample preparation. An equal aliquot of each sample was pooled, and the mixture was used for generating the DDA library. For the pool sample adhesive strip, the high-abundance strip (45KD–66.2KD) was cut as the high-abundance sample, and the remaining part regarded as the low-abundance sample, underwent in-gel trypsin method. The peptides from each sample were desalted using C18 Cartridges (Empore™ SPE Cartridges, Sigma). The peptides were estimated by UV light spectral density at 280 nm using an extinction coefficient of 1.1 of 0.1% (g/l) solution. The peptides of low-abundance sample were fractionated into ten fractions using the Thermo Scientific Pierce High pH Reversed-Phase Peptide Fractionation Kit. Each fraction was concentrated by vacuum centrifugation and reconstituted in 15 μl of 0.1% (v/v) formic acid. Collected peptides were desalted on C18 Cartridges (Empore™ SPE Cartridges C18 (standard density), bed I.D. 7 mm, volume 3 ml, Sigma) and reconstituted in 40 μl of 0.1% (v/v) formic acid. The iRT-Kits (Biognosys) were added to correct the relative retention time differences between runs with volume proportion 1:3 for iRT standard peptides versus sample peptides.

To create the query database for data-independent acquisition (DIA) mass spectra analysis, a spectral library was built using data-dependent acquisitions (DDA). The Q Exactive 480 mass spectrometer connected to a Thermo Scientific Easy nLC 1200 chromatography system was used to perform MS on all fractions for DDA library generation. First, the high-abundance sample and all fractions (from the low-abundance sample) were loaded onto a Thermo Scientific EASY-Spray™ C18 Trap column with the following specifics and dimensions: P/N 164946, 3 μm, 75 μm×2 cm. It was then separated using a Thermo Scientific EASY-Spray™ C18 LC Analytical Column with the following parameters: ES802, 2 μm, 75 μm×25 cm. The full-scan MS detection was conducted in the positive ion mode, with the scan range set to 300–1800 m/z. The resolution was 60000 at 200 m/z for MS1 scans and 15,000 for sequential MS2 scans.

LC-MS/MS set to the DIA mode was used to assess the peptides from each sample by Shanghai Applied Protein Technology. For DIA mass spectrometry, 2 μg peptides from each sample were mixed with iRT peptide before injection.

Chromatographic separation was performed with the Easy nLC-1200, and the Q-Exactive 480 mass spectrometer was used for automatic switching between MS and MS/MS acquisition. In each DIA cycle, one full MS-SIM scan was performed, and 30 DIA scans were performed utilizing a mass range of 350–1800 m/z with the following settings: SIM full scan resolution was set to 120,000 at 200 m/z; AGC 3e6; a maximum IT 50ms; profile mode; DIA scans were performed at a resolution of 15,000; AGC target 3e6; Max IT auto; and the normalized collision energy used was 30 eV. QC samples were injected with DIA mode at the start of MS to monitor performance, and QC was performed every five injections throughout the experiment. Quantitative Proteomics Analysis was conducted by Shanghai Applied Protein Technology Co.

For DDA library data, the FASTA sequence database (<http://www.uniprot.org>) was searched with SpectronautTM 14.4.200727.47784 (Biognosys) software. The iRT peptides sequence was added, and the parameters were set as follows: enzyme: trypsin, max missed cleavages: 2, fixed modification: carbamidomethyl(C), dynamic modification: oxidation(M) and acetyl (Protein N-term). All reported data were based on 99% confidence for protein identification as determined by false discovery rate ($FDR = \frac{N(\text{decoy}) \times 2}{N(\text{decoy}) + N(\text{target})} \leq 1\%$). The spectral library was constructed by importing the original raw files and DDA searching results into Spectronaut Pulsar XTM 12.0.20491.4 (Biognosys). DIA data was analyzed with SpectronautTM 14.4.200727.47784, searching the above constructed spectral library. The main software parameters were as follows: retention time prediction type: dynamic iRT, and interference on MS2 level correction and cross-run normalization were enabled. All results were filtered based on a Q cutoff value of 0.01 (equivalent to $FDR < 1\%$).

The protein differences were calculated using Fisher's exact test, and a P -value < 0.05 was statistically significant. Subsequently, we conducted gene ontology (GO) analysis on the DEPs using Blast2GO (<https://www.blast2go.com/>) and KEGG Orthology analysis using the KEGG database. Protein-protein interaction (PPI) networks were constructed using the STRING database (<http://string-db.org/>) to determine the functional relationship between DEPs.

2.4 Integrated machine learning

Each metabolite's comprehensive weight value was calculated, and ROC analysis was used to assess the effect of each metabolite on the area under the receiver operating characteristic curve (AUC) values. The candidate biomarker panel comprised the top-ranked metabolites with the highest AUC value. To validate candidate biomarkers, integrated analyses such as Logistic Regression (LR), Random Forest (RF), and Support Vector Machine (SVM) were used.

Subsequently, Logical Regression was used to construct the classification model, and the entire dataset was randomly divided into training and test sets in a 7:3 ratio. Notably, the training set was utilized for hyper-parameter optimization and model training, while the test set was used to validate of the model.

2.5 Statistical analysis

Multivariable linear regression models with adjustment were used to investigate the relationship between male age and sperm parameters. Statistical analysis was conducted using Stata 12.0 and significant difference were indicated as follows: * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$.

2.6 Large-scale data

The raw metabolomics data were uploaded to MetaboLights at EMBL-EBI under accession number MTBLS5289 and the raw proteomics data were uploaded to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) under accession number PXD035466

3 Results

3.1 Male age's effect on semen quality

Overall, a total of 12 825 semen samples were included in the present study. We calculated and recorded the semen pH, sperm concentration, volume, total sperm count, progressive motile sperm count, DFI and progressive sperm motility. The details of the semen parameters across age categories are shown in **Table 1**, and the trends are shown in **Figure 2**. After adjustment for abstinence time and semen pH, we evaluated the trends for the semen parameters with respect to age using multiple linear regression analyses. As shown in **Table 2**, the average semen volume decreased from $4.12 \text{ mL} \pm 1.62$ in group A to $3.24 \text{ mL} \pm 1.66$ in group E, with an average decline of 0.04 mL per year. The total sperm count was reduced from $94.35 \text{ million} \pm 75.59$ to $86.14 \text{ million} \pm 74.08$, with an average annual decline of 0.56 million. Moreover, the progressive motility was reduced from $34.07\% \pm 16.96$ to $23.03\% \pm 16.51$, with an average annual decline of 0.46%, and the progressive motile sperm count decreased from $35.27 \text{ million} \pm 38.18$ to $22.32 \text{ million} \pm 27.21$, with an average annual decline of 0.65 million. Finally, the sperm concentration increased from $23.75 \text{ million/mL} \pm 17.51$ to $30.16 \text{ million/mL} \pm 25.57$, with an average annual increment of 0.18 million/mL and sperm DFI increased from $10.24\% \pm 8.75$ to $21.82\% \pm 15.05$, with an average annual increment of 0.49%.

TABLE 1 The descriptive statistics of semen parameters and DFI by male age category.

Age group	A (≤30)	B (31-35)	C (36-40)	D (41-45)	E (≥46)
Number (%)	4638(36.16)	5020(39.14)	2094(16.33)	744(5.80)	329(2.57)
Abstinence time(days)	4.21 ± 2.87	4.29 ± 2.88	4.49 ± 3.24	4.71 ± 4.40	4.88 ± 3.29
pH	7.22 ± 0.18	7.22 ± 0.18	7.23 ± 0.19	7.23 ± 0.18	7.22 ± 0.20
Volume(ml)	4.12 ± 1.62	3.96 ± 1.55	3.79 ± 1.64	3.52 ± 1.50	3.24 ± 1.66
Sperm concentration(10 ⁶ /ml)	23.75 ± 17.51	23.79 ± 17.96	25.55 ± 19.28	25.67 ± 19.98	30.16 ± 25.57
Total sperm count (10 ⁶)	94.35 ± 75.59	90.84 ± 73.78	92.19 ± 79.98	86.63 ± 77.74	86.14 ± 74.08
Progressive motility (%)	34.07 ± 16.96	32.15 ± 16.83	30.06 ± 17.03	26.95 ± 16.98	23.03 ± 16.51
Progressive motile sperm count (10 ⁶)	35.27 ± 38.18	32.03 ± 35.63	30.02 ± 36.00	25.12 ± 28.76	22.32 ± 27.21
DFI (%)	10.24 ± 8.75	11.94 ± 10.12	14.62 ± 11.43	17.79 ± 12.88	21.82 ± 15.05

Data are expressed as mean ± standard deviation. DFI DNA fragmentation index.

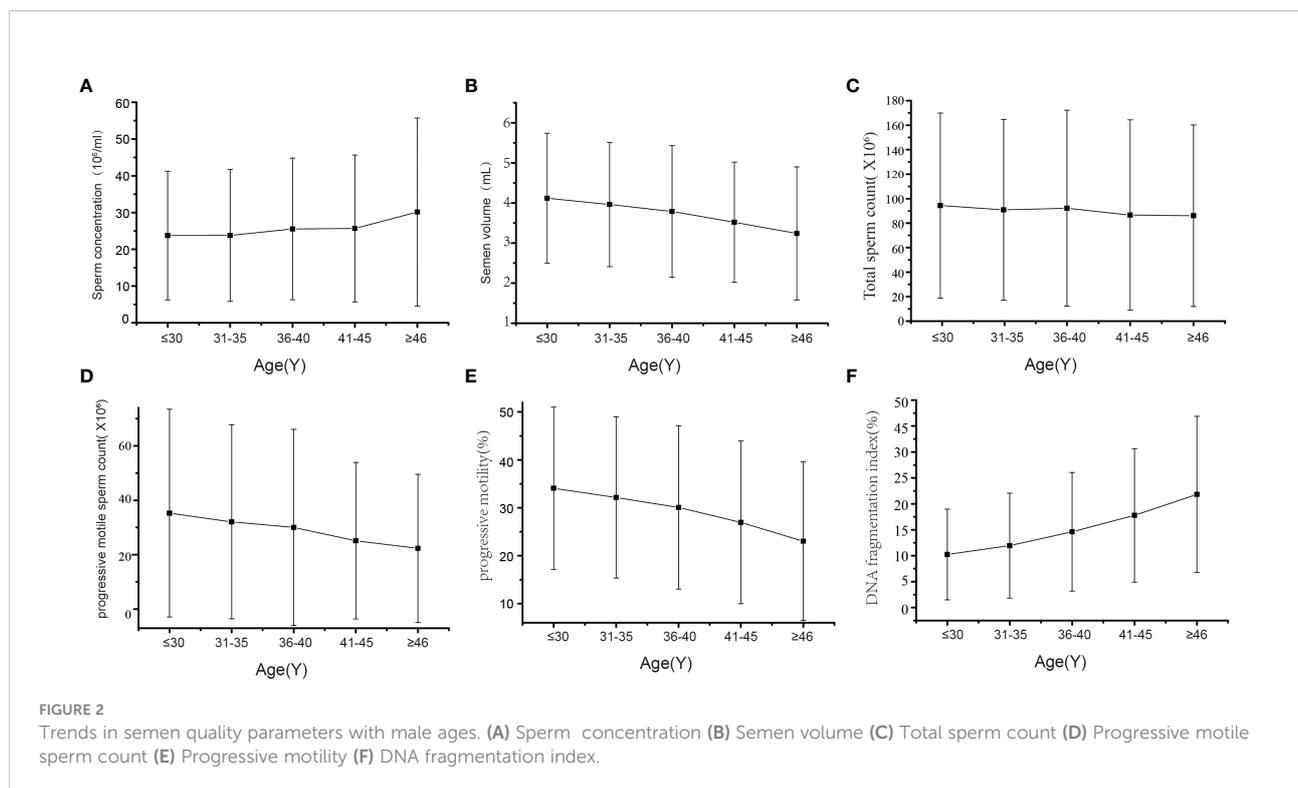


TABLE 2 Linear regression analyses of age trend in semen parameters.

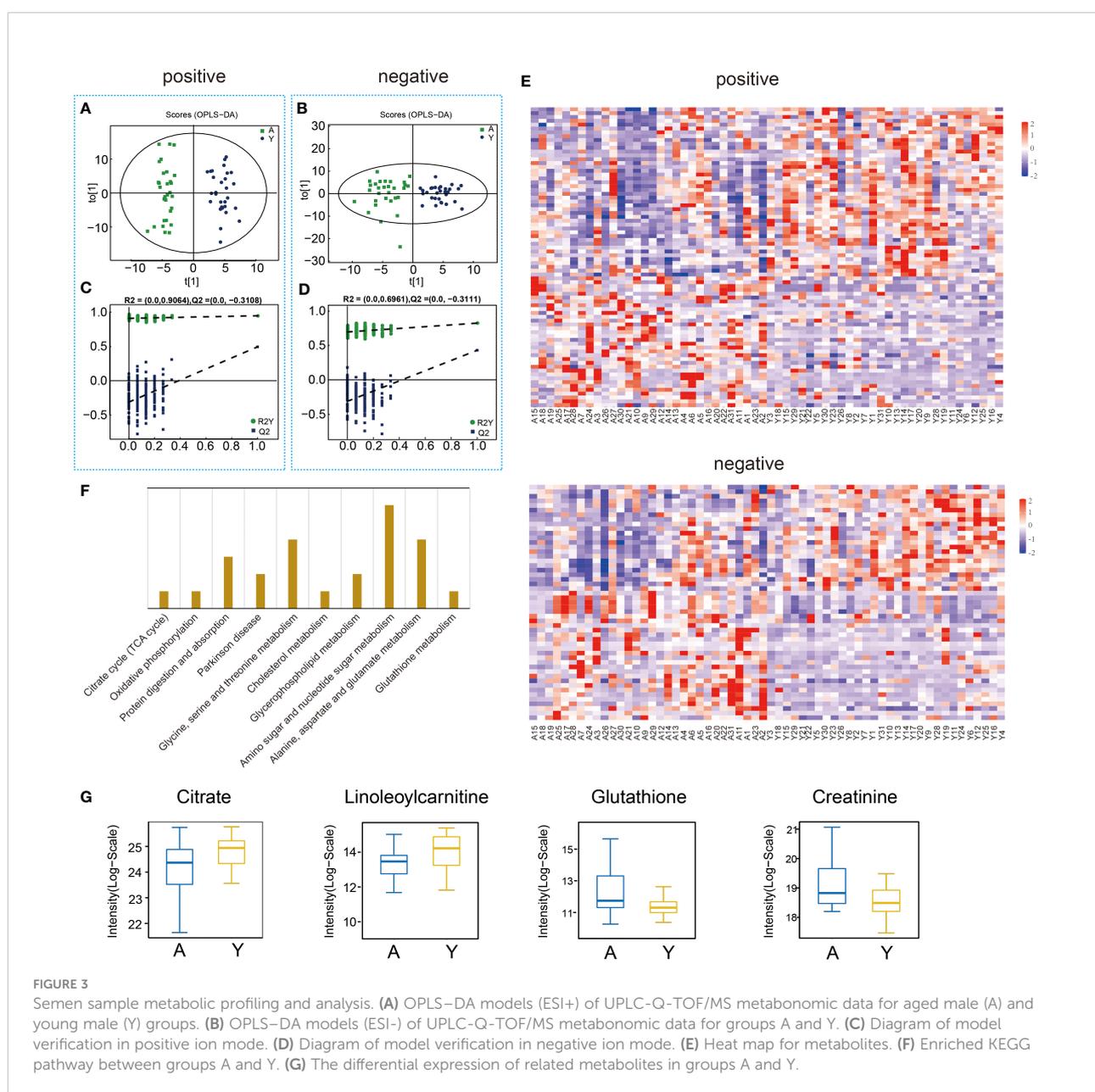
Semen parameter	B (95% CI)	SE	t	P value
Volume (ml)	-0.04 (-0.05, -0.04)	0.00	-15.44	<0.001
Sperm concentration (10 ⁶ /ml)	0.18 (0.11, 0.24)	0.03	5.70	<0.001
Total sperm number (10 ⁶)	-0.56 (-0.81, -0.32)	0.12	-4.55	<0.001
Progressive motility (%)	-0.46 (-0.52, -0.41)	0.03	-16.17	<0.001
Progressive motile sperm count (10 ⁶)	-0.65 (-0.77, -0.53)	0.06	-10.64	<0.001
DFI(%)	0.49 (0.46, 0.53)	0.17	28.78	<0.001

Noted: Adjustment for abstinence time and pH. CI, Confidence Interval; SE, Standard Error; DFI, DNA Fragmentation Index.

3.2 Metabolomics analysis of semen

Thirty semen samples per group for aged and young participants were detected using ultra-high-performance liquid chromatography-quadrupole time-of-flight mass spectrometry (UPLC-Q-TOF/MS). 1331 positive-ion-mode and 870 negative-ion-mode metabolites were identified. The supervised OPLS-DA analysis revealed that the aged and young groups were completely divided into positive and negative ionization modes, which were validated by typical 7-fold cross-validation. Accordingly, high values were obtained for R2Y (0.824 and 0.944) and Q2 (0.430 and 0.494) parameters (Figures 3A, B). In addition, the 200 iterations permutation test was conducted to prevent overfitting of the

OPLS-DA model (Figures 3C, D). 129 differentially expressed metabolites (62 downregulated and 67 upregulated) were identified following multivariate and univariate analyses based on the statistical significance criteria (VIP>1 and P<0.05). Significant differential metabolites are listed in Table SI, and the cluster heatmap of the significant differential metabolites is shown in Figure 3E. KEGG pathway analysis showed significant enrichment in a variety of metabolic pathways, including the citrate cycle (TCA cycle), Parkinson’s disease, cholesterol metabolism, glycerophospholipid metabolism, oxidative phosphorylation, amino acid metabolism, amino sugar and nucleotide sugar metabolism, glutathione metabolism, protein digestion and absorption pathways (Figure 3F). Results of



quantitative analysis of important metabolites are shown in Figure 3G, which were mainly related to energy metabolism and oxidative stress.

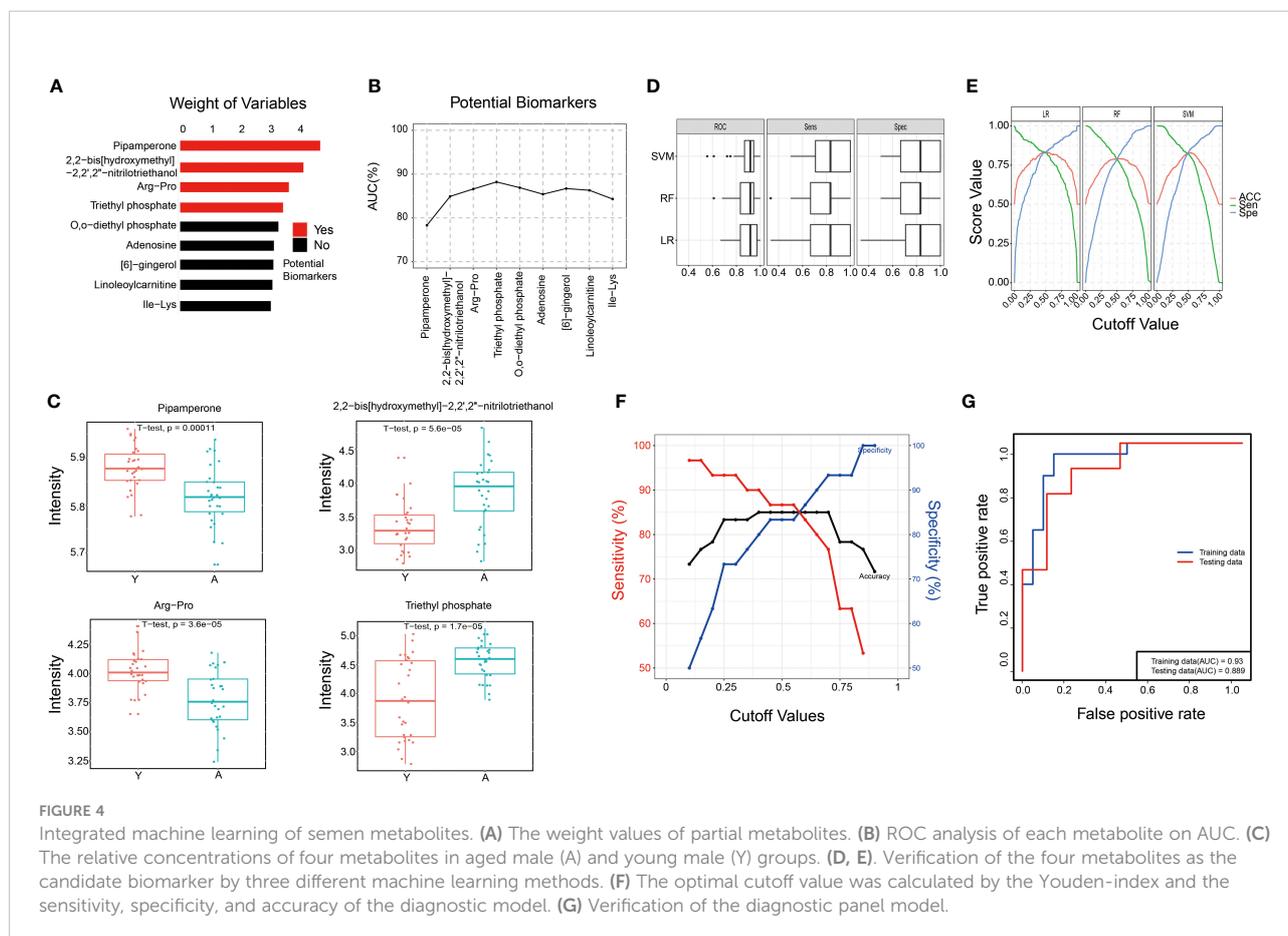
3.3 Integrated machine learning of semen metabolites

The comprehensive weight value of each metabolite was calculated, and the candidate biomarkers were sorted according to their weight values (Figure 4A). Then, we performed ROC analysis to evaluate the effect of each metabolite on AUC values and found that the top four metabolites (Pipamperone, 2,2-Bis(hydroxymethyl)-2,2',2''-nitrioltriethanol, Arg-Pro, Triethyl phosphate) exhibited the best performance for predicting decreased fertility (Figure 4B). The relative concentrations of these four metabolites in two groups were presented in Figure 4C, and Pearson's correlation analysis of the candidate biomarkers was also performed to show the low correlation between these metabolites (Figure S1). To validate the effects of the four metabolites as the candidate biomarker, three different machine learning methods were used and compared, including Logistic Regression (LR), Random Forest (RF), and Support

Vector Machine (SVM). The sensitivity, specificity, and accuracy of the three differential models were all greater than 0.8, which substantiated the good predictive performance of these four metabolites for aging semen (Figures 4D, E). At last, a diagnostic model was constructed by LR to distinguish aging semen from normal semen. The optimal cutoff value of the diagnostic model determined by the Youden-index was 0.58, with a sensitivity, specificity, and accuracy of 0.87, 0.87 and 0.87, respectively (Figure 4F). To test the diagnostic panel model, ROC analysis was applied to 70% of the metabonomic data to train the models, and the trained models were then used for prediction of the remaining 30% metabonomic data, which yielded an AUC value of 0.93 in the training set and 0.889 in the test set (Figure 4G).

3.4 Proteomic analysis of sperm

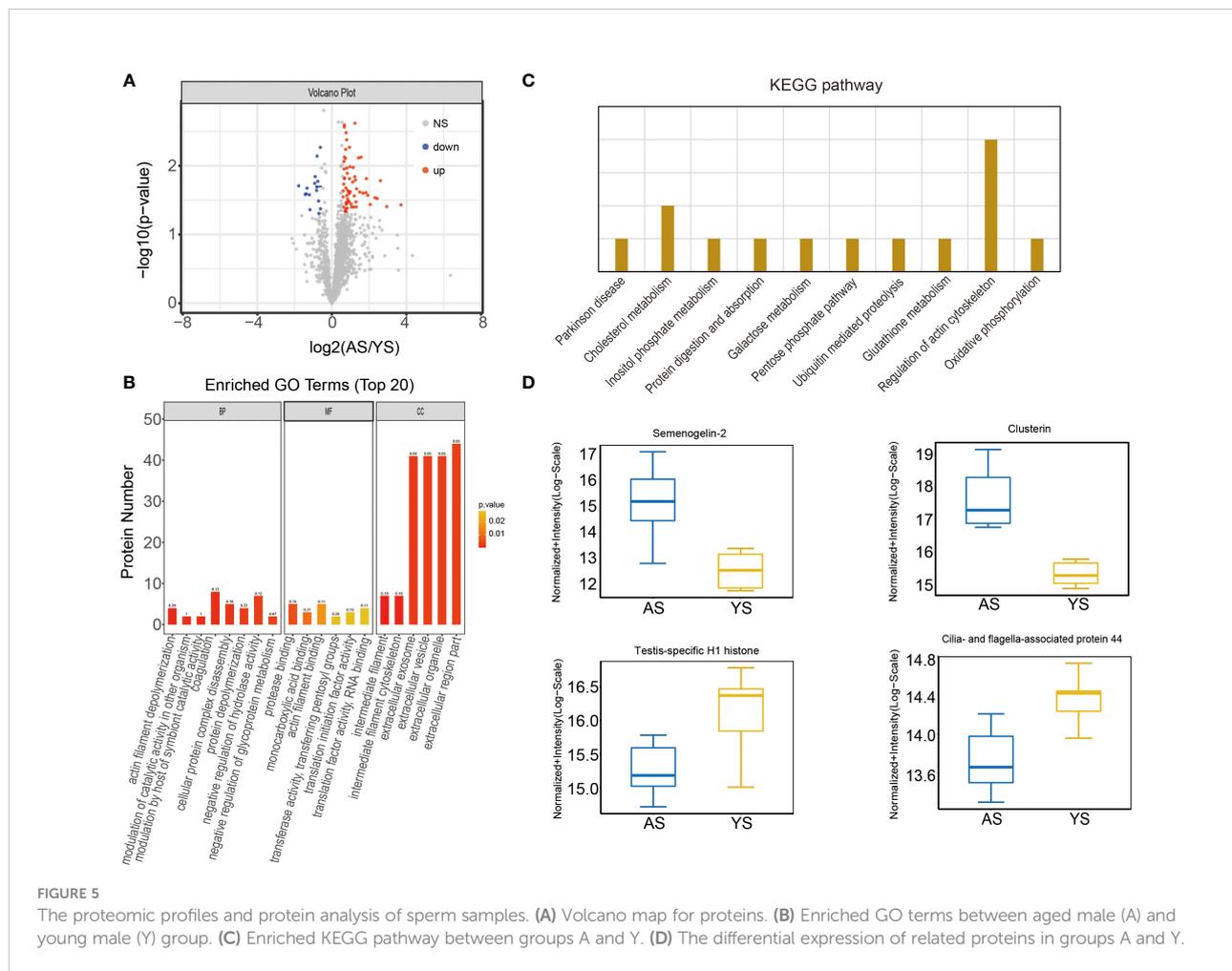
A total of 2369 proteins were identified in both groups using the DIA quantitative proteomics approach, of which 80 proteins were significantly altered (fold change > 1.5 or < 0.67, $p < 0.05$) (Table SII). Specifically, 64 upregulated and 16 downregulated proteins were detected in the aged group



compared to the young group (Figure 5A). GO analysis was conducted to identify the enriched GO terms of the DEPs. The top 20 GO terms are shown in Figure 5B, including actin filament depolymerization, cellular protein complex disassembly and protein depolymerization in biological processes, protease binding and actin filament binding in molecular functions in extracellular region part, intermediate filament and intermediate filament cytoskeleton in cellular components. KEGG analysis showed that DEPs were significantly enriched in protein digestion and absorption, galactose metabolism, pentose phosphate pathway, ubiquitin-mediated proteolysis, oxidative phosphorylation, glutathione metabolism, Parkinson's disease, regulation of actin cytoskeleton, cholesterol metabolism and inositol phosphate metabolism (Figure 5C). The expression level of semenogelin-2, clusterin, testis-specific H1 histone and cilia- and flagella-associated protein 44 were significantly differentially expressed in the aged group (Figure 5D).

3.5 Integrated metabolomics and proteomics analysis

An integrated analysis of differential metabolites and proteins in both groups was conducted to further evaluate the influence of male age on semen quality. Seventeen KEGG pathways were found to be involved in both metabolomics and proteomics, including oxidative phosphorylation, glutathione metabolism, protein digestion and absorption, galactose metabolism, cholesterol metabolism, Parkinson's disease, pantothenate and CoA biosynthesis and the Sphingolipid signaling pathway (Figure 6A). Pearson correlation hierarchical cluster analysis was performed to objectively reflect the expression patterns of differential proteins and metabolites, and similar expression patterns involved in the similar biological process were grouped into the same cluster (Figure 6B). We then constructed a correlation network to assess the association between differential proteins and metabolites (Figure 6C).



positively associated with sperm concentration and DFI but negatively associated with semen volume, total sperm count, progressive motility and progressive motile sperm count, which was consistent with a recent large retrospective study (20). Elevated sperm DFI with male age is well-recognized as an important indicator of sperm DNA damage. The decline of total sperm count and progressive motile sperm count is widely thought to significantly impair sperm fertilization capacity. In addition, we found that the sperm concentration increased while the semen volume decreased with age. We speculated that the decrease in semen volume might be due to the impaired function of the accessory glands of the male reproductive tract as a result of aging. Indeed, when the total sperm count did not change significantly, the decrease in semen volume would inevitably lead to an increase in sperm concentration. Albeit the mechanisms of age-dependent semen quality decline remain largely unclear, current evidence suggests that aging-induced oxidative stress could be a major cause. Accordingly, exploring the involved mechanisms and discovering the biomarkers of aging semen is paramount.

Semen reportedly contains many metabolites including nucleosides, lipids, monosaccharides, amino acids, minerals and electrolytes which are essential in regulating the motility and fertilization of sperm (21). As an easily accessible noninvasive material, semen provides a quick and objective assessment of dysregulated metabolism, suggesting it hold much promise as a new biomarker to predict fertility in advanced-aged males (22). In addition, the dysregulated sperm protein expression is significantly related to male infertility (23). In this study, we constructed the sperm proteomic and semen metabolomic profiles of aged and young males and then analyzed the potential functions of differential proteins and metabolites, which were helpful for further exploration of the mechanism involved. We analyzed sperm proteomics using DIA technology, which could record information on all ions for higher data coverage and improve the quantitative analysis efficiency of the identified proteins and reproducibility (24).

16 downregulated and 64 upregulated proteins were found in the aged sperm proteome compared to the young group. These differential proteins were enriched in various pathways, including protein digestion and absorption, galactose metabolism, pentose phosphate pathway, ubiquitin-mediated proteolysis, oxidative phosphorylation, glutathione metabolism, Parkinson's disease, regulation of actin cytoskeleton, cholesterol metabolism and inositol phosphate metabolism. It has been established that clusterin and semenogelin-2, two members of the epididymal protease inhibitor (Eppin), are present on the human sperm surface to inhibit motility (25, 26). Clusterin expression is reportedly upregulated in many human diseases, including Alzheimer's, type 2 diabetes and prostate cancer (27). Consistently, we found that sperm clusterin was 4.96-fold higher in aged males than in the young group. The decreased

expression of Cilia and flagella-associated protein 44 (Cfap44) in the aged group may directly affect β -tubulin cleavage, preventing microtubule formation and eventually resulting in the decline of sperm motility (28). We also found that testis-specific H1 histone was downregulated expression in the aged group, which might be related to the lack of histone to protamine exchange, and eventually influence offspring through epigenetic effects (29).

UHPLC-Q-TOF/MS was adopted during metabolomic analysis since it can reportedly overcome the shortcomings of ultraviolet detectors and is adequate for analyzing the components of seminal plasma (30). 129 differential (62 downregulated and 67 upregulated) metabolites were found in the aged group and were enriched in 10 KEGG pathways, including citrate cycle (TCA cycle), glutathione metabolism, protein digestion and absorption pathways, Parkinson's disease, glycerophospholipid metabolism, amino acid metabolism, oxidative phosphorylation, amino sugar and nucleotide sugar metabolism as well as cholesterol metabolism. Among these, citrate, creatinine, phosphoric acid, oxidized glutathione, and carnitine have been associated with sperm energy metabolism, oxidative stress, and sperm membrane damage. In addition, spermatogenesis and sperm function are highly energy-dependent, and low sperm ATP correlates with decreased sperm motility and fertilization rate (31). The decline of citrate (the first product of the TCA cycle involved in aerobic metabolism) in seminal plasma has been related to oligozoospermia (14). Oxidative phosphorylation is the primary energy metabolism pathway that mediates sperm motility (32). In the present study, citrate and oxidative phosphorylation were downregulated in the aged group, suggesting that the synthesis of ATP was reduced. In addition, carnitine is highly expressed in the epididymis and has been established to be essential for fatty acid metabolism in sperms. In this respect, lower levels of carnitine have been reported in the seminal plasma of infertile patients, and a strong positive correlation between L-Carnitine expression and sperm motility (33). Our study also found that the aged group had lower levels of Linoleoylcarnitine and Oleoyl-L-carnitine than the young group, indicating that fatty acid metabolism and energy production were reduced. Current evidence suggests the detrimental effect of oxidative stress on male fertility is mediated by damage to the sperm plasma membrane and DNA integrity (34). Reduced glutathione (GSH) and oxidized glutathione (GSSG), two predominant forms of antioxidants that are critical for protecting the cells from oxidative stress, are reportedly upregulated in the aged group (35). Besides, the GSH/GSSG ratio, an indicator of cellular health, has been reported to decrease in neurodegenerative diseases, such as Parkinson's disease and Alzheimer's disease (36). There is a rich literature available suggesting that glutathione could protect sperm from oxidative stress in semen, and glutathione supplementation benefits sperm morphology and motility (37). In addition,

carnitine can protect sperm from oxidative damage and reduce spermatogenic cell apoptosis (38).

In addition to revealing possible mechanisms of decreased sperm motility and increased DFI in aged sperm, we established a biomarker panel that could identify aging semen through an integrated machine-learning method. Compared to one-dimensional statistical methods, machine learning is suitable for multi-dimensional variable analysis and can also select target molecules with high correlation, especially in constructing biomarker models to predict new samples. Furthermore, integrated machine-learning methods and integrating multiple machine learning algorithms were used to screen the most accurate biomarkers. In this study, four metabolites (Pipamperone, 2,2-Bis(hydroxymethyl)-2,2',2''-nitrioltriethanol, Arg-Pro, Triethyl phosphate) were used to construct an optimal biomarker panel and these four metabolic biomarkers were used to characterize aging semen. A simultaneous increase in these four metabolites indicates semen aging and poor sperm function, irrespective of age. This finding could help screen young males presenting with aging semen phenotypes and provide valuable insights to improve the fertility of such patients. However, a limitation of this study is the relatively small sample size used for the integrated machine-learning analysis. Indeed, a larger sample size is warranted to verify the accuracy and robustness of our biomarker panel for aging semen.

In conclusion, we corroborated that semen quality is related to male age and screened the differential metabolites and proteins between aged and young males, explaining the decline of semen quality in aged males to a certain extent. Besides, a biomarker panel was constructed to define aging semen. Our data improves currently available omics information on semen and provides the foothold for discovering novel biomarkers that can identify susceptible populations and improve semen quality and fertility of aged males.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Ethics statement

The studies involving human participants were reviewed and approved by Shanghai First Maternity and Infant Hospital's medical ethics committee. The patients/participants provided their written informed consent to participate in this study.

Author contributions

YG, XT and XJ conceived and designed the study. YY and HY contributed to the recruitment and characterization of the individuals. JL, QC and PK contributed to the sample collection, and FH contributed to analyze the data. YG and JL performed the experiments and wrote the manuscript. XT and WL critically revised the manuscript. All authors provided a critical review and approved the final manuscript.

Funding

This work was supported by grants from the National Natural Science Foundation of China (No. 82271632, 81720108017). Clinical science and technology innovation Program of Shanghai Hospital Development Center(SHDC12021115). Shanghai Rising-Star Program (19QA1407000) and Special Funds for Clinical Medical Research of Chinese Medical Association-Research and Development of Young Physicians in Reproductive Medicine (18010390768).

Acknowledgments

The authors sincerely thank all the participants for their help in collecting data.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fendo.2022.1058250/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

The correlation analysis of the candidate biomarkers.

References

- Attali E, Yogev Y. The impact of advanced maternal age on pregnancy outcome. *Best Pract Res Clin Obstet Gynaecol* (2021) 70:2–9. doi: 10.1016/j.bpobgyn.2020.06.006
- Halvaei I, Litzky J, Esfandiari N. Advanced paternal age: effects on sperm parameters, assisted reproduction outcomes and offspring health. *Reprod Biol Endocrinol* (2020) 18(1):110. doi: 10.1186/s12958-020-00668-y
- Janeczko D, Holowczuk M, Orzel A, Klatka B, Semczuk A. Paternal age is affected by genetic abnormalities, perinatal complications and mental health of the offspring. *BioMed Rep* (2020) 12(3):83–8. doi: 10.3892/br.2019.1266
- Bjorndahl L, Kirkman Brown J, other Editorial Board Members of the WHOLMfE. The sixth edition of the WHO laboratory manual for the examination and processing of human semen: ensuring quality and standardization in basic examination of human ejaculates. *Fertil Steril* (2022) 117(2):246–51. doi: 10.1016/j.fertnstert.2021.12.012
- Lewis SE. Is sperm evaluation useful in predicting human fertility? *Reproduction* (2007) 134(1):31–40. doi: 10.1530/REP-07-0152
- Kwon WS, Rahman MS, Pang MG. Diagnosis and prognosis of male infertility in mammal: the focusing of tyrosine phosphorylation and phosphotyrosine proteins. *J Proteome Res* (2014) 13(11):4505–17. doi: 10.1021/pr500524p
- Llavanera M, Bermudez A, Maynou J, Huetos A, Yeste M. A systematic review identifying fertility biomarkers in semen: a clinical approach through omics to diagnose male infertility. *Fertil Steril* (2022) 118(2):291–313. doi: 10.1016/j.fertnstert.2022.04.028
- Castillo J, Jodar M, Oliva R. The contribution of human sperm proteins to the development and epigenome of the preimplantation embryo. *Hum Reprod Update* (2018) 24(5):535–55. doi: 10.1093/humupd/dmy017
- Siva AB, Kameshwari DB, Singh V, Pavani K, Sundaram CS, Rangaraj N, et al. Proteomics-based study on asthenozoospermia: differential expression of proteasome alpha complex. *Mol Hum Reprod* (2010) 16(7):452–62. doi: 10.1093/molehr/gaq009
- Marchiani S, Tamburrino L, Benini F, Fanfani L, Dolce R, Rastrelli G, et al. Chromatin protamination and catsper expression in spermatozoa predict clinical outcomes after assisted reproduction programs. *Sci Rep* (2017) 7(1):15122. doi: 10.1038/s41598-017-15351-3
- Guijas C, Montenegro-Burke JR, Warth B, Spilker ME, Siuzdak G. Metabolomics activity screening for identifying metabolites that modulate phenotype. *Nat Biotechnol* (2018) 36(4):316–20. doi: 10.1038/nbt.4101
- Panner Selvam MK, Finelli R, Agarwal A, Henkel R. Proteomics and metabolomics - current and future perspectives in clinical andrology. *Andrologia* (2021) 53(2):e13711. doi: 10.1111/and.13711
- Xu Y, Lu H, Wang Y, Zhang Z, Wu Q. Comprehensive metabolic profiles of seminal plasma with different forms of male infertility and their correlation with sperm parameters. *J Pharm BioMed Anal* (2020) 177:112888. doi: 10.1016/j.jpba.2019.112888
- Mumcu A, Karaer A, Dogan B, Tuncay G. Metabolomics analysis of seminal plasma in patients with idiopathic oligoasthenoteratozoospermia using high-resolution NMR spectroscopy. *Andrology* (2020) 8(2):450–6. doi: 10.1111/andr.12707
- Gilany K, Jafarzadeh N, Mani-Varnosfaderani A, Minai-Tehrani A, Sadeghi MR, Darbandi M, et al. Metabolic fingerprinting of seminal plasma from non-obstructive azoospermia patients: Positive versus negative sperm retrieval. *J Reprod Infertil* (2018) 19(2):109–14.
- Soltow QA, Jones DP, Promislow DE. A network perspective on metabolism and aging. *Integr Comp Biol* (2010) 50(5):844–54. doi: 10.1093/icb/icc094
- Lv C, Wang X, Guo Y, Yuan S. Role of selective autophagy in spermatogenesis and Male fertility. *Cells* (2020) 9(11), 2523. doi: 10.3390/cells9112523
- Salmon-Divon M, Shrem G, Balayla J, Nehushtan T, Volodarsky-Perel A, Steiner N, et al. An age-based sperm nomogram: the McGill reference guide. *Hum Reprod* (2020) 35(10):2213–25. doi: 10.1093/humrep/deaa196
- Zhytnik L, Peters M, Tilk K, Simm K, Tonisson N, Reimand T, et al. From late fatherhood to prenatal screening of monogenic disorders: evidence and ethical concerns. *Hum Reprod Update* (2021) 27(6):1056–85. doi: 10.1093/humupd/dmab023
- Gao J, Yuan R, Yang S, Wang Y, Huang Y, Yan L, et al. Age-related changes in human conventional semen parameters and sperm chromatin structure assay-defined sperm DNA/chromatin integrity. *Reprod BioMed Online* (2021) 42(5):973–82. doi: 10.1016/j.rbmo.2021.02.006
- Jia B, Liang J, Lv C, Memon S, Fang Y, Wu G, et al. The characteristics of proteome and metabolome associated with contrasting sperm motility in goat seminal plasma. *Sci Rep* (2021) 11(1):15562. doi: 10.1038/s41598-021-95138-9
- 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. doi: 10.1371/journal.pone.0181115
- Bisconti M, Simon JF, Grassi S, Leroy B, Martinet B, Arcolia V, et al. Influence of risk factors for Male infertility on sperm protein composition. *Int J Mol Sci* (2021) 22(23):13164. doi: 10.3390/ijms222313164
- Eggers B, Eisenacher M, Marcus K, Uszkoreit J. Establishing a custom-fit data-independent acquisition method for label-free proteomics. *Methods Mol Biol* (2021) 2228 p:307–25. doi: 10.1007/978-1-0716-1024-4_22
- Mitra A, Richardson RT, O'Rand MG. Analysis of recombinant human semenogelin as an inhibitor of human sperm motility. *Biol Reprod* (2010) 82(3):489–96. doi: 10.1095/biolreprod.109.081331
- Yamasaki K, Yoshida K, Yoshiike M, Shimada K, Nishiyama H, Takamizawa S, et al. Relationship between semenogelins bound to human sperm and other semen parameters and pregnancy outcomes. *Basic Clin Androl* (2017) 27:15. doi: 10.1186/s12610-017-0059-6
- Wilson MR, Zoubeidi A. Clusterin as a therapeutic target. *Expert Opin Ther Targets* (2017) 21(2):201–13. doi: 10.1080/14728222.2017.1267142
- Tanaka AS, Tanaka M. A novel gene, cilia flagella associated protein 44, encoding an enzyme cleaving FtsZ and tubulin contributes to the regulation of secretory pathway. *Biochem Biophys Res Commun* (2017) 493(1):399–408. doi: 10.1016/j.bbrc.2017.09.010
- Cavalcanti MC, Rizgalla M, Geyer J, Failing K, Litzke LF, Bergmann M. Expression of histone 1 (H1) and testis-specific histone 1 (H1t) genes during stallion spermatogenesis. *Anim Reprod Sci* (2009) 111(2–4):220–34. doi: 10.1016/j.anireprosci.2008.03.018
- Dong Y, Jia G, Hu J, Liu H, Wu T, Yang S, et al. Determination of alkaloids and flavonoids in *sophora flavescens* by UHPLC-Q-TOF/MS. *J Anal Methods Chem* (2021) 2021, 9915027. doi: 10.1155/2021/9915027
- Moscattelli N, Spagnolo B, Pisanello M, Lemma ED, De Vittorio M, Zara V, et al. Single-cell-based evaluation of sperm progressive motility via fluorescent assessment of mitochondria membrane potential. *Sci Rep* (2017) 7(1):17931. doi: 10.1038/s41598-017-18123-1
- Hashemitabar M, Sabbagh S, Orzizadeh M, Ghadiri A, Bahmanzadeh M. A proteomic analysis on human sperm tail: comparison between normozoospermia and asthenozoospermia. *J Assist Reprod Genet* (2015) 32(6):853–63. doi: 10.1007/s10815-015-0465-7
- Agarwal A, Said TM. Carnitines and male infertility. *Reprod BioMed Online* (2004) 8(4):376–84. doi: 10.1016/S1472-6483(10)60920-0
- Bisht S, Dada R. Oxidative stress: Major executioner in disease pathology, role in sperm DNA damage and preventive strategies. *Front Biosci (Schol Ed)* (2017) 9(3):420–47. doi: 10.2741/s495
- Fratemale A, Paoletti MF, Casabianca A, Nencioni L, Garaci E, Palamara AT, et al. GSH and analogs in antiviral therapy. *Mol Aspects Med* (2009) 30(1–2):99–110. doi: 10.1016/j.mam.2008.09.001
- Owen JB, Butterfield DA. Measurement of oxidized/reduced glutathione ratio. *Methods Mol Biol* (2010) 648:269–77. doi: 10.1007/978-1-60761-756-3_18
- Beygi Z, Forouhari S, Mahmoudi E, Hayat SMG, Nourimand F. Role of oxidative stress and antioxidant supplementation in Male fertility. *Curr Mol Med* (2021) 21(4):265–82. doi: 10.2174/1566524020999200831123553
- Dobrakowski M, Kasprczyk S, Horak S, Chyra-Jach D, Birkner E, Kasprczyk A. Oxidative stress and motility impairment in the semen of fertile males. *Andrologia* (2017) 49(10):e12783. doi: 10.1111/and.12783