



Published in final edited form as:

Syst Biol Reprod Med. 2023 August ; 69(4): 296–309. doi:10.1080/19396368.2023.2195964.

Exploring the Internal Exposome of Seminal Plasma with Semen Quality and Live Birth: A Pilot Study

Emily Houle^{1,†}, YuanYuan Li^{2,†}, Madison Schroder², Susan L McRitchie², Tayyab Rahil³, Cynthia K Sites³, Susan Jenkins Sumner^{2,*}, J. Richard Pilsner^{1,*}

¹Department of Obstetrics and Gynecology, School of Medicine, Wayne State University, 275 E. Hancock Street, Detroit, MI, USA

²Department of Nutrition, Nutrition Research Institute, 500 Laureate Way, University of North Carolina at Chapel Hill School of Public Health, Kannapolis, NC 28010, USA

³Department of Obstetrics and Gynecology, Division of Reproductive Endocrinology and Infertility, University of Massachusetts Medical School, Baystate Medical Center, 759 Chestnut Street, Springfield, MA, USA

Abstract

Infertility is clinically defined as the inability to achieve pregnancy within 12 months of regular unprotected sexual intercourse and affects 15% of couples worldwide. Therefore, identification of novel biomarkers that can accurately predict male reproductive health and couples' reproductive success is of major public health significance. The objective of this pilot study is to test whether untargeted metabolomics is capable of discriminating reproductive outcomes and understand associations between the internal exposome of seminal plasma and the reproductive outcomes of semen quality and live birth among ten participants undergoing assisted reproductive technology (ART) in Springfield, MA. We hypothesize that seminal plasma offers a novel biological matrix by which untargeted metabolomics is able to discern male reproductive status and predict reproductive success. The internal exposome data was acquired using UHPLC-HR-MS on randomized seminal plasma samples at UNC at Chapel Hill. Unsupervised and supervised multivariate analyses were used to visualize the differentiation of phenotypic groups classified by men with normal or low semen quality based on World Health Organization guidelines

* Corresponding Authors J. Richard Pilsner, PhD, Professor, Department of Obstetrics and Gynecology, School of Medicine, Wayne State University, 275 E. Hancock Street, Detroit, MI, USA, rpilsner@wayne.edu, Susan CJ Sumner, PhD, Professor, Department of Nutrition, Nutrition Research Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27514, USA, Tel: 919-622-4456, Susan_sumner@unc.edu.

† co-first authors

Author Contributions

Assisted with study design and analyzing and interpreted data, drafted and completed manuscript: EH; analyzed and interpreted data, drafted and completed manuscript: YYL; performed exposome measurements and assisted with data analysis: MS; assisted with data analyses and interpretation: SLM; oversaw recruitment of patients for study and infertility diagnoses: TR, CKS; assisted with study design, and interpreted data: SJS; oversaw overall study design and interpreted data: JRP.

Disclosure

The authors report there are no conflicts of interests to declare.

Ethics Approval

Consent from eligible participants interested in participating was obtained by attending physicians by signed consent forms. This study was approved by the institutional review boards at Baystate Medical Center and at the University of Massachusetts Amherst (reference BH-12–190).

as well as by successful ART: live birth or no live birth. Over 100 exogenous metabolites, including environmentally relevant metabolites, ingested food components, drugs and medications, and metabolites relevant to microbiome-xenobiotic interaction, were identified and annotated from the seminal plasma samples, through matching against the NC HHEAR hub in-house experimental standard library. Pathway enrichment analysis indicated that fatty acid biosynthesis and metabolism, vitamin A metabolism, and histidine metabolism were associated sperm quality; while pathways involving vitamin A metabolism, C21-steroid hormone biosynthesis and metabolism, arachidonic acid metabolism, and Omega-3 fatty acid metabolism distinguished live birth groups. Taken together, these pilot results suggest that seminal plasma is a novel matrix to study the influence of the internal exposome on reproductive health outcomes. Future research aims to increase the sample size to validate these findings.

Keywords

seminal plasma; exposomics; male infertility; metabolomics; environmental exposure

Introduction

The World Health Organization (WHO) defines infertility as a disease characterized by the failure to achieve pregnancy after 12 months of unprotected sexual intercourse (WHO 2021). Infertility was primarily considered a female-related disease; however, it is now recognized that male-factor infertility contributes up to 50% of cases (Esteves, Miyaoka et al. 2011). Moreover, in 30% to 40% of cases, the etiology of infertility remains unknown (Bracke, Peeters et al. 2018). Diagnosis of male infertility continues to rely on semen parameter guidelines set by the WHO (WHO 2021); however, semen parameters have been shown to be less predictive of overall male fertility and display high temporal variability (Minhas, Bettocchi et al. 2021). Thus, there is a growing interest in identifying new biomarkers in sperm and seminal fluid that could improve the diagnosis of male infertility and predict couples' reproductive success.

Seminal plasma is the medium produced by several accessory sex glands of the male reproductive tract. Approximately 90% of semen is comprised of seminal plasma (Duncan and Thompson 2007), which protects and nourishes sperm during transit in the male reproductive tract and, subsequently, in the female reproductive tract. The components of seminal plasma are diverse and include fructose, putrescine, spermine, and spermidine, which protect sperm from the acidic female reproductive tract. Other components include proteins, extracellular vesicles, RNAs, and antioxidants (Dobrakowski, Kasperczyk et al. 2017, Ayaz, Houle et al. 2021). These components have been shown to play important roles in sperm development and function, suggesting that the seminal plasma is not just a medium for sperm transfer and protection but can also be utilized as a biospecimen matrix to study spermatogenesis and male infertility (Panner Selvam, Finelli et al. 2021). In addition to the endogenous metabolites and proteins related to the host metabolism, many xenobiotic compounds originating from environmental and other exogenous exposures (e.g., nutrition, pharmaceuticals, etc.) can also be identified in seminal plasma (Laws, Neff et al. 2021, Rodprasert, Toppari et al. 2021).

The CDC defines the exposome as the measure of all exposures an individual accumulates over a life time and how these exposures then relate to health ((NIOSH) 2022). Exposomics is a comprehensive exposure assessment term that is divided into two main branches: the internal and external exposome. The external exposome captures an individual's exposures to the built environment, environmental stressors, and lifestyle behaviors through surveys and sensor technology. Alternatively, the assessment of the internal exposome relies primarily on quantitative -OMICS approaches including genomics, transcriptomics, proteomics, lipidomics, etc., to generate large-scale biomarker datasets. Of the -OMICS techniques currently available to interrogate the internal exposome, metabolomics is emerging as a powerful tool to understand associations between the endogenous and exogenous metabolites in biological matrices and human health (Panner Selvam, Finelli et al. 2021). There are two primary approaches within metabolomics, untargeted and targeted analyses (Cajka and Fiehn 2016), both of which have advantages and limitations. As the name suggests, untargeted metabolomics provides an unbiased/discovery approach by analyzing all detectable metabolites/signals, known or unknown, within the biological sample. Alternatively, targeted metabolomics measures an *a priori* investigator-selected group or panel of metabolites.

The use of metabolomics to understand the etiology of male infertility is still in its infancy but has gained traction over the last decade. In 2014, Zhang et al. identified 10 urinary metabolites that distinguished oligozoospermic and idiopathic infertile men, suggesting these metabolites could be used as biomarkers of fertility (Zhang, Huang et al. 2014). In a follow-up study, the same group identified 37 urinary metabolites that differentiated fertile and normozoospermic infertile men (Zhang, Mu et al. 2014). Others have investigated metabolomics within spermatozoa, but this approach has its limitations owing to the presence of seminal plasma and washing reagents during sperm processing (Panner Selvam, Finelli et al. 2021). In 2015, Pavia et al. were the first group to identify metabolites (n=69) within sperm cells (Paiva, Amaral et al. 2015). In 2018, Zhao et al. identified 33 metabolites; 27 of which were decreased in asthenozoospermic compared to normozoospermic men (Zhao, Zhang et al. 2018). Metabolomic profiles via proton magnetic resonance spectroscopy has also been used to distinguish different fractions of sperm after percoll-based density gradient centrifugation, such that choline and lactate levels were lower in sperm found in the '80%' compared to '40%' gradient layers (Reynolds, Calvert et al. 2017).

Untargeted metabolomics of seminal plasma has also been employed to understand associations with reproductive health outcomes. For instance, carnitine, lactate, and choline were associated with sperm physiology and defects in sperm morphology; whereas the metabolite panel of alanine, tyrosine, and phenylalanine classified men with low sperm concentration (Gupta, Mahdi et al. 2011). Another study identified 19 metabolites that distinguished asthenozoospermic and healthy men; these included several amino acids as well as changes in lipids, phospholipids, and cholesterol metabolism. For the later, the oxysterols of 5 α -cholesterol and 7-ketocholesterol were elevated in asthenozoospermic men indicating higher oxidative stress (Zhang, Diao et al. 2015).

To our knowledge, the use of untargeted internal exposomics to identify both endogenous and exogenous metabolites in seminal plasma and understand their relationship to

reproductive outcomes is limited. We hypothesize that seminal plasma offers a novel biological matrix by which metabolomic profiles are able to discern male reproductive status and predict reproductive success. To test this hypothesis, we have utilized untargeted metabolomics to uncover metabolomic profiles associated with semen quality and successful live birth. This proof-of-concept research will support using seminal plasma as an ideal biomatrix for metabolomics analysis to understand the interactions between host metabolism, life-stage exposures, and reproductive outcomes.

Results

Demographics, semen quality and reproductive outcomes

The demographics and reproductive characteristics of the SEEDS participants are shown in Table 1. The low sperm quality (LSQ) group (n=5) had a mean age (range) of 36.6 (32.2 – 42.3) years and BMI of 35 (22.0 – 48.1), while the normal semen quality (NSQ) group (n=5) had a mean age (range) of 43 (32.0 – 61.1) years and BMI of 26.8 (24.4 – 29.2). By design, sperm concentration, motility, and morphology were significantly lower in the LSQ group ($p<0.01$) compared with the NSQ group. We also stratified the seminal plasma samples according to whether men of couples had successful birth outcomes after receiving infertility treatment (live birth vs. not-live birth). As shown in Table 1, the live birth (LB) group (n=4) had a mean (range) age of 36.7 (32.0 – 38.6) years and a BMI of 31.6 (22.0 – 48.1), and those in the no live birth (NLB) group (n=6) had a mean (range) age of 42 (32.2 – 61.1) and BMI of 31.2 (24.4 – 40.3). Of note, 50% of men had normal semen parameters in both the LB (n=2) and NLB (n=3) groups and suggests that sperm concentration, motility and morphology were not predictive of live birth success.

Exogenous metabolites determined in seminal plasma

After matching seminal plasma signals to the IESL, over 100 exogenous metabolites were identified at the OL1 and OL2a ontology levels, or annotated (OL2b) (Table S1), including 30 OL1 high confidence metabolites (Table 2). These confidently identified exogenous compounds included environmentally relevant metabolites (phthalates, tobacco-related metabolite, phenols, and benzene metabolites), ingested food components (hippuric acid and derivatives, benzaldehyde/benzoic acid metabolites, purine derivatives, tryptophan-indole metabolite, and pyridine carboxylic acids), drugs and medications (acetaminophen, ibuprofen metabolite, naproxen), and metabolites relevant to microbiome-xenobiotic interaction (dipeptide, sugar amide, tyrosine metabolite).

Styrene, a chemical used as an additive in food packaging, is a benzene derivative and is suspected to impact sperm quality (Kolstad, Bonde Jp Fau - Spano et al. 1999). However, styrene was significantly ($p<0.05$) higher in the LB group than in the NLB group. A polyphenol metabolite related to fruit consumption, 4-hydroxyhippuric acid, was higher in the NSQ than the LSQ ($p<0.1$) and higher in the LB group than the NLB group ($p<0.1$). Mercapturic acids, (R,S)-N-Acetyl-S-(2-hydroxy-3-buten-1-yl)-L-cysteine and N-Acetyl-S-(2-hydroxy-3-propionamide)-L-cysteine (OL2a) were identified in seminal plasma, with higher relative abundance in the LSQ group compared with the NSQ ($p<0.1$). Monoethyl phthalate (OL2a) and N-heptylparaben (OL2b) were higher in the LSQ group

than the NSQ ($p < 0.1$) group. Exposure to phthalates and parabens can occur during ingesting of foods and beverages and use of personal care products.

Metabolic profile and pathways differentiating NSQ and LSQ groups

Unsupervised and supervised multivariate analyses were used to visualize the differentiation of metabolic profiles between NSQ and LSQ. The model statistics for the unsupervised PCA ($R^2X=0.54$) and the supervised OPLS-DA ($R^2X=0.44$, $R^2Y=0.918$, $Q^2=-0.41$) indicated a distinct metabolic profile for NSQ and LSQ (Figure 1 A and B). The negative Q^2 in this discovery analysis indicates that a larger sample size is needed for replication. Using a cutoff at $p < 0.1$, 811 signals significantly differentiated the NSQ and LSQ groups. Endogenous metabolites identified or annotated by the IESL (OL1, OL2a, OL2b) and public databases (PDa) are shown in Table S2. Endogenous metabolites significantly perturbed between the NSQ and LSQ ($p < 0.1$) included mono- and polyunsaturated fatty acids, dicarboxylic acid derivatives, acylcarnitines, eicosanoids, di- or tripeptides, cholic acids and hormones (Table S1).

Long-chain acylcarnitines (myristoyl carnitine, tetradecenoyl carnitine, and 3-hydroxyhexadecanoylcarnitine) and short-chain acylcarnitines (glutaryl carnitine, succinyl carnitine) were higher in the LSQ than the NSQ, while the free carnitine, deoxy carnitine, and the medium-chain acylcarnitine (methylglutaryl carnitine) were perturbed in the opposite direction. The immune-related eicosanoids (15-deoxy-¹²¹⁴-Prostaglandin J2, prostaglandin E2, 11-dehydro TXB2, 11-dehydrothromboxane B2, 5,8,11-eicosatriynoic acid, and 5S,12R-Dihydroxy-6Z,8E,10E,14Z-eicosatetraene-1,20-dioic acid) were all lower in the LSQ compared to NSQ (Table S2).

For pathway analysis, we first used Metabolyte 5.0 to identify the overall perturbations between LSQ and NSQ using all the normalized signals after filtering. We found that six pathways involving fatty acid biosynthesis and metabolism, vitamin A metabolism, and histidine metabolism were associated with the differentiation of LSQ and NSQ (Figure 1). We verified the empirical findings from the Metaboanalyt analysis by also matching signals to the IHPSL and PD. As shown in Figure 2, in the carnitine shuttle pathway, all the identified long-chain acylcarnitine (C14-C18) and short-chain acylcarnitine (C3-C5) were elevated in the LSQ compared to NSQ, while the medium-chain acylcarnitine (e.g., methylglutaryl carnitine), free carnitine, and deoxy carnitine were decreased. As to the vitamin A metabolism (Figure 3), retinol, retinoic acid, retinal, 3,4-didehydroretinoate, 18-hydroxyretinoic acid, and 4-hydroxyretinoic acid were verified by the MS and MS/MS spectra information. We did not observe a significant difference in the level of retinol between the LSQ and NSQ, although most of the other vitamin A-related metabolites, including retinoic acids, hydroxy retinoic acids, and didehydroretinoates were lower in the LSQ than the NSQ.

Metabolic profile and pathways differentiating LB and NLB groups.

Next, seminal plasma samples were classified into the LB ($n=4$) and NLB ($n=6$) groups. Differentiation of LB and NLB groups was shown in the unsupervised PCA and the supervised OPLS-DA plot (Figure 4 A and B) with strong model

statistics in the PCA ($R^2X=0.473$) and OPLS-DA ($R^2X=0.82$, $R^2Y=1$, $Q^2=0.84$).

There were 701 signals significantly differentiating the NLB and LB ($p<0.1$) with over 50 endogenous metabolites identified or annotated by matching against the IESL (OL1, OL2a, OL2b) and public databases (PDa), as shown in Table S3.

The immune response-related eicosanoids [e.g., prostaglandins, arachidonoyl glycerol, hydroxyeicosatetraenoic acids (HETE), and heptadecatrienoic acids (HHE)], hormones (e.g., 19-nor-4-androstene-3,17-dione, nandrolone, 5-androsten-3 beta.-ol-17-one), medium-chain acylcarnitines (e.g., octanoyl carnitine, decanoyl carnitine, and dodecenoylcarnitine), phospholipid metabolites (e.g., phosphorylcholine, 1-palmitoyl-2-myristoyl-sn-glycero-3-phosphocholine, 1-stearoyl-2-linoleoyl-sn-glycero-3-phosphocholine), polyunsaturated fatty acids (9,12-octadecadiynoic acid and pinolenic acid) were all lower in the NLB compared to the LB.

Pathway analysis identified 8 pathways involving vitamin A metabolism, C21-steroid hormone biosynthesis and metabolism, arachidonic acid and prostaglandin, omega-3 fatty acid metabolism, pyrimidine metabolism, glycan biosynthesis, and glycolysis and gluconeogenesis, were NLB and LB groups. Figure 2 shows the vitamin A metabolism pathway was downregulated in NLB compared to LB, as the metabolites of retinal and retinoic acids were dramatically reduced in the NLB. As for the carnitine shuttle pathway (Figure 3), most acylcarnitines, including free carnitine, long-chain acylcarnitines, and medium-chain acylcarnitines, were downregulated in the NLB compared to LB.

Discussion

Infertility is one of the most common reproductive health disorders affecting 10–16% of couples globally and male reproductive health issues are responsible in roughly 50% of cases (Jungwirth, Giwercman et al. 2012, Thoma, McLain et al. 2013). Semen parameters analyses remain the most prevalent diagnostic tool to identify male infertility; however, men with normal semen parameters may still be unable to achieve pregnancy (Table 1). Therefore, new biomarkers and a better understanding of infertility pathophysiology to improve the diagnosis of male infertility are needed. Seminal plasma, a male biofluid produced by the male reproductive tract containing many biomolecules related to the host metabolism and exogenous metabolites originating from environmental exposures, is an ideal non-invasive biospecimen to examine the effects of the internal exposome on reproductive outcomes. This knowledge could help us identify novel biomarkers to predict birth outcomes. More importantly, it will help us recognize and reduce the modifiable risk factors (such as environmental or dietary exposures) for male infertility and identify targets to guide interventions. The current proof of concept research has demonstrated the potential of using seminal plasma as a novel biospecimen in internal exposome research for large cohort studies on reproductive health.

Exogenous metabolites identified/annotated from seminal plasma

Growing evidence has demonstrated that environmentally relevant exposures, especially endocrine-disrupting chemicals (EDCs), such as phthalates, phenols, parabens, etc., are associated with diminished fecundity for men and women in reproduction (Laws, Neff et

al. 2021, Rodprasert, Toppari et al. 2021). Amongst publications focused on male exposures and reproductive health, the majority analyzed the exposure metabolites through serum/urine samples to determine their association with male reproductive health outcomes (Jeannet M Fau - Vassalli, Vassalli P Fau - Magnin et al. 1974, Chung, Buck Louis et al. 2019, Deng, Yang et al. 2022), while only a few used seminal plasma as a matrix (Xia, Zhu et al. 2009, Buck Louis, Smarr et al. 2018, Wang, Liu et al. 2018, Engel, Baumann et al. 2021). In this pilot study, we identified (or annotated) over 100 exogenous metabolites from seminal plasma (Table S1). Many EDCs and pollutants, including benzene, phthalates, tobacco metabolites, phenols, fluorene, and mercapturic acids, have been analyzed in seminal plasma using the targeted analysis and demonstrated associations with adverse reproductive outcomes in previous research (Sharpe 2012, Jurewicz, Radwan M Fau - Sobala et al. 2013). These metabolites were identified or annotated by matching with MS, RT, and/or MS/MS. In addition, monoethyl phthalate and the N-heptyl paraben metabolite were significantly higher ($p < 0.1$) in men from the LSQ group compared to the NSQ group (Figure S1). Interestingly, we have also found that mercapturic acids [N-Acetyl-S-(2-hydroxy-3-propionamide)-L-cysteine, and (R,S)-N-Acetyl-S-(2-hydroxy-3-buten-1-yl)-L-cysteine] were significantly higher ($p < 0.1$) in the LSQ compared to the NSQ group. Mercapturic acids can be formed after exposures to pollution (e.g., car exhaust), combustion of plant matter (e.g., tobacco products), or ingestion of foods heated to high temperatures (e.g. from acylamide in fried foods) (Nauhaus, Fennell et al. 1996, Rietjens, Dussort et al. 2018, Frigerio, Campo et al. 2020). The underlying mechanisms for the correlation between these environmental exposures and low sperm quality are not fully understood, but studies implicate a role for oxidative stress (Martin-Hidalgo, Bragado et al. 2019, Poli, Andreoli et al. 2020, Tran-Guzman and Culty 2022).

In addition to exogenous metabolites relevant to environmental exposure, we also identified the existence of over-the-counter medications (acetaminophen, naproxen, and carboxyibuprofen) with the forms of the parent compound and/or their metabolites. Our results are consistent with the perspective that many drugs/medications can enter the male genitourinary tract and be excreted into semen, which may interfere with semen characteristics and cause infertility and/or adverse birth outcomes. Therefore, there is a need to conduct pharmacokinetic- and therapeutic-drug-monitoring studies to identify and reduce the risk of medications the male reproductive system. Beyond the environmental relevant metabolites and medications, we have also identified/annotated several nutritional metabolites related to coffee consumption and intake of polyphenol-enriched food and metabolites involved in microbiome-xenobiotic interaction (Pichini, Zuccaro et al. 1994, Li, Rushing et al. 2022). A fruit consumption-related metabolite, 4-hydroxyhippuric acid, was associated with normal sperm quality and live birth. These results demonstrated that in addition to studying the toxic exposures related to adverse health outcomes, seminal plasma could also be used to study beneficial exposures, including dietary intake and supplements and their relations to reproductive health.

Metabolic profile and pathway differentiated between LSQ vs. NSQ

We found that fatty acid metabolism, especially for eicosanoids (Table S2) and carnitine/acylcarnitine profiles (Figure 2), differed between LSQ and NSQ groups. Our findings

overlap with results from a cohort study that included 100 men where the polyunsaturated fatty acids (PUFA) metabolism was positively associated with sperm parameters (Huang, Liu et al. 2019). It is recognized that PUFAs are important constituents composed of the sperm membrane and the essential precursors to synthesizing eicosanoids, including prostaglandins and thromboxanes, significantly influencing sperm morphology and function in normal and pathological conditions (Van Tran, Malla et al. 2017, Collodel, Castellini et al. 2020). Our finding of decreased eicosanoids in the LSQ compared to the NSQ (Table S2) aligns with most published results (Esmaeili, Shahverdi et al. 2015, Wang, Wu et al. 2019). Carnitine shuttle and fatty acid oxidation are pathways highly associated with sperm parameters and functions (Jeulin and Lewin 1996, Agarwal and Said 2004). The observed decrease of carnitine and deoxycarnitines in the LSQ (Table S2 and Figure 2) is in line with the importance of carnitine with normal sperm parameters and function. Interestingly, the long-chain acylcarnitines (LCFA-acylcarnitine) (Figure 2) were increased in the LSQ compared to NSQ, which is opposite to most literature that increased LCFA-acylcarnitines are associated with active mitochondrial activity, energy production, maturation, and antioxidant properties (Ruiz-Pesini, Diez et al. 1998). We speculate that the elevated LCFA-acylcarnitines in the LSQ might be related to impaired sperm mitochondrial functions due to the overload of fatty acids and the reduced flexibility in metabolic functions, which was found in the scenario of insulin resistance (Koves, Ussher Jr Fau - Noland et al. 2008).

Metabolic profile and pathway differentiated between No Live Birth (NLB) vs. Live birth (LB)

Although sperm parameters are conventionally used for male infertility diagnosis, sperm parameters were not associated with a successful live birth. For our live birth analyses, we observed more profound perturbations in vitamin A, PUFA, eicosanoids/arachidonic acid/prostaglandin, and steroid hormone metabolism compared to the perturbations between LSQ vs. NSQ. Most metabolites in these pathways showed higher levels in the LB group than in the NLB group. Sex hormones serve as a source of nutrients for sperm development and maturation, protecting sperm from infection and enabling them to overcome the immunological and chemical environment of the female reproductive tract and participate in ovum fertilization (Vitku, Kolatorova et al. 2017). Research has demonstrated that low circulating testosterone level is associated with abnormal sperm parameters and adverse birth outcomes in assisted reproduction treatment (Trussell, Coward et al. 2019). In addition to the inherited genetic reasons, exposure to the EDCs might also cause low testosterone levels in men and therefore cause failures in pregnancy. Although prostaglandins are the primary components in seminal plasma, their role in facilitating pregnancy is not fully understood. The positive correlation between the prostaglandin levels and live birth outcomes might be due to the functions of prostaglandins and other eicosanoids in maintaining sperm motility transport and subsequent fertilization for a successful pregnancy outcome (Bromfield 2014). We also found that the Vitamin A pathway was positively associated with successful live birth and sperm quality (Figure 3). Vitamin A and its downstream metabolites are essential nutrients needed for spermatogenesis (Chung and Wolgemuth 2004, Wang, Liang et al. 2020) and a part of the antioxidant defense system within the body and, more specifically, the male reproductive tract (Li, Long et al. 2019). Disturbing this system can have many downstream effects, including capacitation and

acrosome reaction (Malivindi, Rago et al. 2018). We also studied the perturbation of carnitine shuttle pathways regarding NLB vs. LB. In contrast to our findings for sperm quality, almost all acylcarnitines, regardless of long-chain, short-chain, or free carnitine, were lower in the NLB group. This result is consistent with most publications that carnitine and acylcarnitines are essential nutrients for reproductive health, and supplements with carnitine and acylcarnitines are beneficial to reproductive outcomes (Agarwal and Said 2004, Khaw, Wong et al. 2020).

Overlap of metabolic profiles and pathways perturbed in semen quality and live birth analyses.

Metabolic profiles perturbed between LSQ vs. NSQ and between NLB vs. LB were quite distinctive, which can be reflected in the Venn diagram (Figure S2). There were 811 and 701 signals that distinguished LSQ vs. NSQ and NLB vs. LB groups, respectively; however, only 6 signals overlapped between these two pairwise comparisons. Such low overlap between these two comparison groups further suggest that these two reproductive outcomes have distinct metabolomic profiles. Although pathway analyses were enriched in some common pathways in sperm quality and live birth analyses, individual metabolites within the shared pathway that were significantly changed ($p < 0.1$) in each pairwise comparison were different. In general, carnitine shuttle and fatty acid metabolism were more related to the differentiation of sperm quality (concentrations, motility, and cell number), while vitamin A, sexual hormone, USFA, eicosanoids/arachidonic acid/prostaglandin metabolism were positively associated with the successful live birth. The limited overlap between these two comparison groups highlights the need for further research to uncover novel biomarker(s) to integrated with (or replace) sperm parameters to predict ART outcomes and to identify targets to develop interventions to improve successful live births.

A main limitation of this pilot study is the small sample size. As presented herein, when seminal plasma samples were classified into the LB (n=4) and NLB (n=6) groups, strong OPLS0-DA model statistics ($R^2X=0.82$, $Q^2=0.84$) were found, with a positive Q^2 that indicates that the study may be replicated with a similar sample size. However, when the classification of NSQ and LSQ were used, the OPLS-DA model statistics were not as strong ($R^2X=0.44$, $Q^2= -0.41$), such that the negative Q^2 for the NSQ vs LSQ comparison indicated that a larger sample size was necessary for replication. Based on the current analysis, we believe that the LB vs NLB could be replicated with the study size provided herein, and that the NSQ vs LSQ would require at least 10 biospecimens per group. An appropriate followup study would require at a minimum 10 samples in each group to provide 80% power ($\alpha=0.05$) to detect effect sizes of approximately 1.4 or higher using a two-side Wilcoxon Rank Sum Test (Faul, Erdfelder E Fau - Lang et al. 2007). Following replication, specific hypothesis can be developed for assessment by quantitative targeted metabolomics analysis.

Future research should include untargeted exposome analyses for replication of this study with higher sample sizes, and to include multiple timepoints to determine the intrapersonal variability of the internal exposome across seminal plasma collected at multiple time-points. In addition, future studies should include the impact of the exposome on sperm, including

sperm capacitation, acrosome reaction, and fertilization. Research should also focus on determining the concordance of the exposome of seminal plasma with less invasive body fluids, such as serum and urine, that may find more clinical utility for biomarker assessment.

Conclusion

This exploratory study used an untargeted platform to simultaneously capture the internal exposome profile from seminal plasma covering endogenous metabolites related to the host metabolism and the circulating exogenous metabolites from environmental/lifestyle exposures. With this approach, we discovered metabolites and metabolic perturbations in seminal plasma related to sperm quality and live birth. This highlights the promise of untargeted exposomics of seminal plasma as a novel approach to advance our understanding of the male contribution to reproductive success.

Methods

Study Population and Design:

This study comprised of 10 male partners of couples recruited at the Baystate Reproductive Medicine in Springfield, MA, USA as part of the Sperm Environmental Epigenetics and Development Study (SEEDS) (Wu, Olmsted et al. 2017). Male participants had to meet the following inclusion criteria: 18–55 years old without vasectomy and fresh ejaculate sperm that was used for assisted reproductive technology (ART) treatment. Clinic personnel recorded participant information including demographic information, fertilization, embryo data, and birth outcomes (i.e., live birth).

For male semen quality analysis, the 10 samples were evaluated using three semen parameters (concentration, total motility, and percent normal morphology). Using the WHO reference cutoffs (Cooper, Noonan et al. 2010), individuals were identified as having normal semen quality or low semen quality (normal semen quality/NSQ n=5, low semen quality/LSQ n=5). To be considered NSQ, an individual had to have all mentioned semen parameters above WHO cutoffs. LSQ individuals had to have 2 or more parameters below WHO cutoffs. The 10 semen specimens used in this analysis were from individuals who either had or did not have a live birth. Live birth information was provided by Baystate Medical to denote whether participants had a successful live birth (LB; n=4) or no live birth (NLB, n= 6) after the ART cycle.

Seminal Plasma Collection:

Semen samples were collected in a sterile plastic specimen cup after a 2–3-day abstinence period. To separate motile sperm from the seminal plasma, semen samples were processed using a two-step (80 and 40%) gradient fractionation. The seminal plasma was removed from the top of the gradient and placed in a sterile tube and frozen at -80°C until analyses.

Metabolomics Analysis:

Sample Preparation.—Seminal plasma samples were prepared according to the published protocols for plasma and urine with modifications (Li, Ghanbari et al. 2020,

Harville, Li et al. 2021). Seminal plasma (50- μ L) was mixed with methanol (400- μ L containing 500 ng/ml L-tryptophan-d5 and vortexed (5,000 rpm for 2 min). Quality control pools (QC pools) were made by pooling 7- μ L plasma from each study sample, and the LC-MS grade water was used to create blanks. QC pools and blanks were prepared identical to the individual study sample. All samples were centrifuged at 16,000 rcf for 10 min at 4°C. The supernatant (350- μ L) was dried and reconstituted with 100 μ l water-methanol (95:5, v/v) for the untargeted analysis. The individual study samples were randomized and QC pools, blanks, and NIST plasma reference material were interspersed. The overlap total ion chromatograms (TICs) of the study samples within different phenotypic groups were shown in Figure S3.

Analysis and data acquisition and preprocessing.—The untargeted internal exposome data was acquired on a Vanquish UHPLC system coupled with a Q Exactive™ HF-X Hybrid Quadrupole-Orbitrap Mass Spectrometer (UHPLC-HR-MS; ThermoFisher Scientific, San Jose, CA). A volume of 5 μ l was injected into the instrument for analysis. Metabolites were separated via an HSS T3 C18 column (2.1 \times 100 mm, 1.7 μ m, Waters Corporation) at 50 °C with mobile phases of water (A) and methanol (B), each containing 0.1% formic acid (v/v). The UHPLC linear gradient began with 2% B, increased to 100% B in 16 min, and then held for 4 min, with a flow rate of 400 μ l/min. The untargeted data were acquired from 70 to 1050 m/z under the data-dependent acquisition mode. The UPLC-HRMS data was processed by Progenesis QI (version 2.1, Waters Corporation) for peak picking and alignment. Background signals were excluded if the mean intensity across blanks were higher than that across the quality control study pools based on the non-normalized data. The remaining peaks were normalized by Progenesis QI using the “normalize to all” feature. The peaks that were highly varied (RSD>50%) across QC pools were excluded for further analysis.

Metabolite identification and annotation.—Peaks were identified or annotated by Progenesis QI through matching to an in-house experimental standards library (IESL) developed by the North Carolina HHEAR Hub at UNC Chapel Hill. The IESL was developed using approximately 2,400 compounds covering endogenous metabolites associated with host metabolism and microbiome metabolism and exogenous compounds and their metabolites or conjugates that are derived from environmentally relevant chemicals, ingested food components, drugs and medications. All spectra and chromatographic data for compounds in the IESL were acquired under identical conditions to the study samples. In addition to matching peaks to the IESL, peaks were also matched to public databases (PD), including NIST and METLIN, to derive annotations. Ontology levels (OL or PD) for the individual peaks are provided to indicate the evidence, including retention time (RT), exact mass (MS), MS/MS fragmentation pattern, and/or isotopic ion pattern, which supported the identification or annotation (Li, Douillet et al. 2020, Harville, Li et al. 2021). Metabolites that matched to the IESL by RT (\pm 0.5 min), MS (<5 ppm), and MS/MS (similarity score > 30) were labeled with OL1, whereas those that match by RT and MS were labeled OL2a. Metabolites identification with OL1 and OL2a are considered highly confident in this study. An OL2b label was provided for signals matched by MS and MS/MS to the IESL but had a RT outside the allowed window (\pm 0.5 min).

The OL2b matches indicate the conjugates/isomers that share similar moieties with the matched compound in IESL. The label of PDa was provided for signals matched by MS (<5 ppm) and experimental MS/MS (similarity score > 30) to public databases (PD). Annotated signals with a PDa label could be the listed compound, isomer, or derivative of the listed compound. For this study, we report metabolites identified within OL1 or OL2a and annotated within OL2b or PDa (Table 2, S1–S3).

Statistical, Multivariate, and Pathway Analyses.

Statistical analysis was conducted using the phenotypes classified by semen quality (LSQ versus NSQ) and live birth outcomes (NLB versus LB), respectively. The normalized and filtered data were visualized by SIMCA 15.0 (Umetrics, Umeå, Sweden) using unsupervised principal component analysis (PCA) and using supervised orthogonal partial least squares discriminate analysis (OPLS-DA). Wilcoxon Rank Sum test and fold change (FC, based on median) for each pairwise comparison were calculated by R version 4.0.3 (2020–10-10). Peaks/metabolites with the calculated p-value < 0.1 were considered significant to the phenotypic groups for comparison. Due to the small sample size, we used the Wilcoxon Rank Sum test. Using a p-value threshold of 0.1 allowed us to reduce the detectable effect size from approximately 2.1 to 1.8, with 80% power (Faul, Erdfelder E Fau - Lang et al. 2007). The p-value, direction of change, and fold change for the metabolite comparisons between the study groups are presented in Table S1–S3.

Pathway analysis was conducted by MetaboAnalyst 5.0 using the Functional Analysis module (MS Peaks), which was built on the Mummichog algorithm (Li, Ghanbari et al. 2020, Li, Li et al. 2021, Li, Li et al. 2022, Rushing, McRitchie et al. 2022) and GSEA. All the normalized peaks (m/z and rt), 15648 peaks, after data preprocessing and quality control were uploaded to the pathway analysis tool, together with the p-value and fold-change calculated from the pairwise comparison. All possible metabolites matched by m/z were searched in the Homo sapiens (human) [MFN] pathway library. A 3-ppm mass accuracy tolerance was used in signal annotations to identify candidate pathways. This software uses a p-value cutoff of 0.05 when evaluating the size of the permutation group to select significant signals to match all possible metabolites. Significance of the perturbed pathways was determined using a combined p-value < 0.05 (Mummichog and GSEA). We verified the enriched empirical annotations made through the Mummichog algorithm by matching the corresponding signals against the in-house physical standard library and public databases (NIST, HMDB).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

This research was funded in part by the NIEHS HHEAR Untargeted Analysis Resource (U2CES030857). This work was also supported by grants R01-ES028298 and R01-ES030942 from the National Institute of Health (NIH).

Data sharing:

Metabolomics data and relevant deidentified metadata will be uploaded to the National Metabolomics Data Repository (NMDR, metabolomicsworkbench.org) and can be accessed via <http://dx.doi.org/10.21228/M8ZM62>

Abbreviations:

WHO	World Health Organization
ART	Assisted reproductive technology
SEEDS	Sperm Environmental Epigenetics and Development Study
NSQ	Normal semen quality
LSQ	Low semen quality
LB	Live birth
NLB	No live birth
QC	Quality Control
PD	Public databases
OL	Ontology levels
RT	Retention time
MS	Exact mass
PCA	Principal component analysis
OPLS-DA	Orthogonal partial least squares discriminate
IESL	In-house experimental standards library
EDCs	Endocrine-disrupting chemicals

References

- (NIOSH), T. N. I. f. O. S. a. H. (2022, August 19, 2022). "Exposome and Exposomics." Retrieved February 13, 2023, from <https://www.cdc.gov/niosh/topics/exposome/default.html>.
- Agarwal A and Said TM (2004). "Carnitines and male infertility." *Reproductive Biomedicine Online* 8(4): 376–384. [PubMed: 15149558]
- Ayaz A, Houle E and Pilsner JR (2021). "Extracellular vesicle cargo of the male reproductive tract and the paternal preconception environment." *Systems biology in reproductive medicine* 67(2): 103–111. [PubMed: 33630671]
- Bracke A, Peeters K, Punjabi U, Hoogewijs D and Dewilde S (2018). "A search for molecular mechanisms underlying male idiopathic infertility." *Reproductive BioMedicine Online* 36(3): 327–339. [PubMed: 29336995]
- Bromfield JJ (2014). "Seminal fluid and reproduction: much more than previously thought." *Journal of Assisted Reproduction and Genetics* 31: 627–636. [PubMed: 24830788]

- Buck Louis GM, Smarr MM, Sun L, Chen Z, Honda M, Wang W, Karthikraj R, Weck J and Kannan K (2018). "Endocrine disrupting chemicals in seminal plasma and couple fecundity." *Environmental Research* 163: 64–70. [PubMed: 29426029]
- Cajka T and Fiehn O (2016). "Toward Merging Untargeted and Targeted Methods in Mass Spectrometry-Based Metabolomics and Lipidomics." *Analytical Chemistry* 88(1): 524–545. [PubMed: 26637011]
- Chung MK, Buck Louis GM, Kannan K and Patel CJ (2019). "Exposome-wide association study of semen quality: Systematic discovery of endocrine disrupting chemical biomarkers in fertility require large sample sizes." *Environment International* 125: 505–514. [PubMed: 30583854]
- Chung SSW and Wolgemuth DJ (2004). "Role of retinoid signaling in the regulation of spermatogenesis." *Cytogenetic and Genome Research* 105(2–4): 189–202. [PubMed: 15237207]
- Collodel GA-O, Castellini C, Lee JC and Signorini CA-O (2020). "Relevance of Fatty Acids to Sperm Maturation and Quality." *Oxidative Medicine and Cellular Longevity* 2020.
- Cooper TG, Noonan E, Von Eckardstein S, Auger J, Baker HWG, Behre HM, Haugen TB, Kruger T, Wang C, Mbizvo MT and Vogelsohn KM (2010). "World Health Organization reference values for human semen characteristics*†." *Human Reproduction Update* 16(3): 231–245. [PubMed: 19934213]
- Deng YL, Yang P, Wang YX, Liu C, Luo Q, Shi T, Zeng JY, Lu TT, Chen PP, Miao Y, Zhang M, Cui FP, Lu WQ and Zeng QA-OX (2022). "Urinary concentrations of polycyclic aromatic hydrocarbon and phthalate metabolite mixtures in relation to semen quality among men attending an infertility clinic." *Environmental Science and Pollution Research* 29: 81749–81759. [PubMed: 35737263]
- Dobrakowski M, Kasperczyk S, Horak S, Chyra-Jach D, Birkner E and Kasperczyk A (2017). "Oxidative stress and motility impairment in the semen of fertile males." *Andrologia* 49(10): e12783.
- Duncan MW and Thompson HS (2007). "Proteomics of semen and its constituents." *PROTEOMICS – CLINICAL APPLICATIONS* 1(8): 861–875. [PubMed: 21136740]
- Engel KM, Baumann S, Blaurock J, Rolle-Kampczyk U, Schiller J, von Bergen M and Grunewald S (2021). "Differences in the sperm metabolomes of smoking and nonsmoking men." *Biology of Reproduction* 105(6): 1484–1493. [PubMed: 34554205]
- Esmaili V, Shahverdi AH, Moghadasian MH and Alizadeh AR (2015). "Dietary fatty acids affect semen quality: a review." *Andrology* 3(3): 450–461. [PubMed: 25951427]
- Esteves SC, Miyaoka R and Agarwal A (2011). "An update on the clinical assessment of the infertile male." *Clinics* 66(4): 691–700. [PubMed: 21655766]
- Faul F, Erdfelder A-G, Fau - Lang E, Lang A, Fau - Buchner Ag and Buchner A (2007). "G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences." *Behavior Research Methods* 39(2): 175–191. [PubMed: 17695343]
- Frigerio GA-O, Campo LA-O, Mercadante R, Miel y ska-Švach D, Pavanello S and Fustinoni SA-O (2020). "Urinary Mercapturic Acids to Assess Exposure to Benzene and Other Volatile Organic Compounds in Coke Oven Workers." *International Journal Environmental Research and Public Health* 17(5): 1801. [PubMed: 32164281]
- Gupta A, Mahdi AA, Ahmad MK, Shukla KK, Jaiswer SP and Shankhwar SN (2011). "1H NMR spectroscopic studies on human seminal plasma: A probative discriminant function analysis classification model." *Journal of Pharmaceutical and Biomedical Analysis* 54(1): 106–113. [PubMed: 20719458]
- Harville EW, Li YY, Pan K, McRitchie S, Pathmasiri W and Sumner S (2021). "Untargeted analysis of first trimester serum to reveal biomarkers of pregnancy complications: a case–control discovery phase study." *Scientific Reports* 11(1).
- Huang Q, Liu L, Wu Y, Wang X, Luo L, Nan B, Zhang J, Tian M and Shen H (2019). "Seminal plasma metabolites mediate the associations of multiple environmental pollutants with semen quality in Chinese men." *Environment International* 132.
- Fau - Vassalli Jeannet M, P., P, Fau - Magnin C, Vassalli, Fau - Cruchaud A, Magnin C and Cruchaud A (1974). "The LDA assay, a new test for the detection of auto, allo and specific anti-leukemic lymphocytotoxic antibodies." *Bollettino Dell'istituto Sieroterapico Milanese* 53(1).

- Jeulin C and Lewin LM (1996). "Role of free L-carnitine and acetyl-L-carnitine in post-gonadal maturation of mammalian spermatozoa." *Human Reproduction* 2(2): 87–102.
- Jungwirth A, Giwercman A, Tournaye H, Diemer T, Kopa Z, Dohle G, Krausz C and European I Association of Urology Working Group on Male (2012). "European Association of Urology guidelines on Male Infertility: the 2012 update." *European Urology* 62(2): 324–332. [PubMed: 22591628]
- Jurewicz J, Radwan W Fau - Sobala M, Sobala S Fau - Brze nicki W, Fau - Ligocka D. Brze nicki S, Fau - Radwan P. Ligocka D, Fau - Bochenek M. Radwan P, Fau - Hanke W. Bochenek M and Hanke W (2013). "Association between a biomarker of exposure to polycyclic aromatic hydrocarbons and semen quality." *International Journal of Occupational Medicine and Environmental Health* 26: 790–801. [PubMed: 24310922]
- Khaw SC, Wong ZZ, Anderson R and Martins da Silva S (2020). "l-carnitine and l-acetylcarnitine supplementation for idiopathic male infertility." *Reproduction and Fertility* 1(1): 67–81. [PubMed: 35128424]
- Kolstad HA, Bonde M Fau - Spano Jp, Fau - Giwercman A. Spano M, Fau - Zschiesche W. Giwercman A, Fau - Kaae D. Zschiesche W, Fau - Larsen S. B. Kaae D, Fau - Roeleveld N. Larsen Sb and Roeleveld N (1999). "Change in semen quality and sperm chromatin structure following occupational styrene exposure." *International Archives of Occupational and Environmental Health* 72: 135–141. [PubMed: 10392560]
- Koves TR, Ussher RC Fau - Noland Jr, Fau - Slentz D. Noland Rc, Fau - Mosedale M. Slentz D, Fau - Ilkayeva O. Mosedale M, Fau - Bain J. Ilkayeva O, Fau - Stevens R. Bain J, Fau - Dyck J. R. B. Stevens R, Fau - Newgard C. B. Dyck Jr, Fau - Lopaschuk G. D. Newgard Cb, Fau - Muoio D. M. Lopaschuk Gd and Muoio DM (2008). "Mitochondrial overload and incomplete fatty acid oxidation contribute to skeletal muscle insulin resistance." *Cell Metabolism* 7(1): 45–56. [PubMed: 18177724]
- Laws MJ, Neff AM, Brehm E, Warner GR and Flaws JA (2021). "Endocrine disrupting chemicals and reproductive disorders in women, men, and animal models." *Advances in Pharmacology* 92: 151–190. [PubMed: 34452686]
- Li SA-O, Li Y, Rushing BA-O, Harris SA-O, McRitchie SA-O, Jones JC, Dominguez D, Sumner SA-O and Dohlman HA-O (2021). "Multi-omics analysis of glucose-mediated signaling by a moonlighting G β protein Asc1/RACK1." *PLOS Genetics* 17(7).
- Li SA-O, Li Y, Rushing BR, Harris SA-O, McRitchie SL, Dominguez D, Sumner SJ and Dohlman HA-O (2022). "Multi-Omics Analysis of Multiple Glucose-Sensing Receptor Systems in Yeast." *Biomolecules* 12(2): 175. [PubMed: 35204676]
- Li X, Long X.-y., Xie Y.-j., Zeng X, Chen X and Mo Z.-c. (2019). "The roles of retinoic acid in the differentiation of spermatogonia and spermatogenic disorders." *Clinica Chimica Acta* 497: 54–60.
- Li Y-Y, Douillet C, Huang M, Beck R, Sumner SJ and Styblo M (2020). "Exposure to inorganic arsenic and its methylated metabolites alters metabolomics profiles in INS-1 832/13 insulinoma cells and isolated pancreatic islets." *Archives of Toxicology* 94(6): 1955–1972. [PubMed: 32277266]
- Li YY, Ghanbari R, Pathmasiri W, McRitchie S, Poustchi H, Shayanrad A, Roshandel G, Etemadi A, Pollock JD, Malekzadeh R and Sumner SCJ (2020). "Untargeted Metabolomics: Biochemical Perturbations in Golestan Cohort Study Opium Users Inform Intervention Strategies." *Frontiers in Nutrition* 7.
- Li YY, Rushing B, Schroder M, Sumner S and Kay CA-O (2022). "Exploring the Contribution of (Poly)phenols to the Dietary Exposome Using High Resolution Mass Spectrometry Untargeted Metabolomics." *Mol Nutr Food Res* 66(21): e2100922. [PubMed: 35106906]
- Malivindi R, Rago V, De Rose D, Gervasi MC, Cione E, Russo G, Santoro M and Aquila S (2018). "Influence of all- trans retinoic acid on sperm metabolism and oxidative stress: Its involvement in the physiopathology of varicocele-associated male infertility." *Journal of Cellular Physiology* 233(12): 9526–9537. [PubMed: 29943815]
- Martin-Hidalgo D, Bragado MJ, Batista AR, Oliveira PF and Alves MG (2019). "Antioxidants and Male Fertility: from Molecular Studies to Clinical Evidence." *Antioxidants (Basel)* 8(4).
- Minhas S, Bettocchi C, Boeri L, Capogrosso P, Carvalho J, Cilesiz NC, Cocci A, Corona G, Dimitropoulos K, Gül M, Hatzichristodoulou G, Jones TH, Kadioglu A, Martínez Salamanca

- Jl, Milenkovic U, Modgil V, Russo GI, Serefoglu EC, Tharakan T, Verze P and Salonia A (2021). "European Association of Urology Guidelines on Male Sexual and Reproductive Health: 2021 Update on Male Infertility." *European Urology* 80(5): 603–620. [PubMed: 34511305]
- Nauhaus SK, Fennell TR, Asgharian B, Bond JA and Sumner SC (1996). "Characterization of urinary metabolites from Sprague-Dawley rats and B6C3F1 mice exposed to [1,2,3,4-13C]butadiene." *Chem Res Toxicol* 9(4): 764–773. [PubMed: 8831821]
- Paiva C, Amaral A, Rodriguez M, Canyellas N, Correig X, Ballescà JL, Ramalho-Santos J and Oliva R (2015). "Identification of endogenous metabolites in human sperm cells using proton nuclear magnetic resonance (1H-NMR) spectroscopy and gas chromatography-mass spectrometry (GC-MS)." *Andrology* 3(3): 496–505. [PubMed: 25854681]
- Panner Selvam MK, Finelli R, Agarwal A and Henkel R (2021). "Proteomics and metabolomics — Current and future perspectives in clinical andrology." *Andrologia* 53(2).
- Pichini S, Zuccaro P and Pacifici R (1994). "Drugs in semen." *Clin Pharmacokinet* 26(5): 356–373. [PubMed: 8055681]
- Poli D, Andreoli R, Moscato L, Pelà G, de Palma G, Cavallo D, Petyx M, Pelosi G, Corradi M and Goldoni M (2020). "The Relationship Between Widespread Pollution Exposure and Oxidized Products of Nucleic Acids in Seminal Plasma and Urine in Males Attending a Fertility Center." *Int J Environ Res Public Health* 17(6).
- Reynolds S, Calvert SJ, Paley MN and Pacey AA (2017). "1H Magnetic Resonance Spectroscopy of live human sperm." *MHR: Basic science of reproductive medicine* 23(7): 441–451. [PubMed: 28431025]
- Rietjens I, Dussort P, Günther H, Hanlon P, Honda H, Mally A, O'Hagan S, Scholz G, Seidel A, Swenberg J, Teeguarden J and Eisenbrand G (2018). "Exposure assessment of process-related contaminants in food by biomarker monitoring." *Arch Toxicol* 92(1): 15–40. [PubMed: 29302712]
- Rodprasert W, Toppari J and Virtanen HE (2021). "Endocrine Disrupting Chemicals and Reproductive Health in Boys and Men." *Front Endocrinol (Lausanne)* 12: 706532. [PubMed: 34690925]
- Ruiz-Pesini E, Diez C, Lapeña AC, Pérez-Martos A, Montoya J, Alvarez E, Arenas J and López-Pérez MJ (1998). "Correlation of sperm motility with mitochondrial enzymatic activities." *Clin Chem* 44(8 Pt 1): 1616–1620. [PubMed: 9702947]
- Rushing BR, McRitchie S, Arbeeva L, Nelson AE, Azcarate-Peril MA, Li YY, Qian Y, Pathmasiri W, Sumner SCJ and Loeser RF (2022). "Fecal metabolomics reveals products of dysregulated proteolysis and altered microbial metabolism in obesity-related osteoarthritis." *Osteoarthritis Cartilage* 30(1): 81–91. [PubMed: 34718137]
- Sharpe RM (2012). "Sperm counts and fertility in men: a rocky road ahead. Science & Society Series on Sex and Science." *EMBO Rep* 13(5): 398–403. [PubMed: 22491033]
- Thoma ME, McLain AC, Louis JF, King RB, Trumble AC, Sundaram R and Buck Louis GM (2013). "Prevalence of infertility in the United States as estimated by the current duration approach and a traditional constructed approach." *Fertility and Sterility* 99(5): 1324–1331.e1321. [PubMed: 23290741]
- Tran-Guzman A and Culty M (2022). "Eicosanoid Biosynthesis in Male Reproductive Development: Effects of Perinatal Exposure to NSAIDs and Analgesic Drugs." *Front Toxicol* 4: 842565. [PubMed: 35295224]
- Trussell JC, Coward RM, Santoro N, Stetter C, Kunselman A, Diamond MP, Hansen KR, Krawetz SA, Legro RS, Heisenleder D, Smith J, Steiner A, Wild R, Casson P, Coutifaris C, Alvero RR, Robinson RB, Christman G, Patrizio P, Zhang H and Lindgren MC (2019). "Association between testosterone, semen parameters, and live birth in men with unexplained infertility in an intrauterine insemination population." *Fertility and Sterility* 111(6): 1129–1134. [PubMed: 30982604]
- Van Tran L, Malla BA, Kumar S and Tyagi AK (2017). "Polyunsaturated Fatty Acids in Male Ruminant Reproduction - A Review." *Asian-Australas J Anim Sci* 30(5): 622–637. [PubMed: 26954196]
- Vitku J, Kolatorova L and Hampl R (2017). "Occurrence and reproductive roles of hormones in seminal plasma." *Basic Clin Androl* 27: 19. [PubMed: 29046808]

- Wang GS, Liang A, Dai YB, Wu XL and Sun F (2020). "Expression and localization of retinoid receptors in the testis of normal and infertile men." *Molecular Reproduction and Development* 87(9): 978–985. [PubMed: 32770619]
- Wang YX, Liu C, Chen YJ, Chen HG, Yang P, Wang P, Huang LL, Ai SH, Duan P, Pan A, Zeng Q and Lu WQ (2018). "Predictors and correlations of phthalate metabolite concentrations in urine and seminal plasma among reproductive-aged men." *Environ Res* 161: 336–344. [PubMed: 29190551]
- Wang YX, Wu Y, Chen HG, Duan P, Wang L, Shen HQ, Lu WQ, Sun B, Wang Q, Zhang B, Chavarro JE, Zhang J and Pan A (2019). "Seminal plasma metabolome in relation to semen quality and urinary phthalate metabolites among Chinese adult men." *Environ Int* 129: 354–363. [PubMed: 31150977]
- WHO. (2021). "World Health Organization Laboratory Manual for the Examination and Processing of Human Semen. 6th ed." Retrieved accessed on 3 December 2021, from <https://www.who.int/publications/i/item/9789240030787>.
- Wu H, Olmsted A, Cantonwine DE, Shahsavari S, Rahil T, Sites C and Pilsner JR (2017). "Urinary phthalate and phthalate alternative metabolites and isoprostane among couples undergoing fertility treatment." *Environmental research* 153: 1–7. [PubMed: 27875712]
- Xia Y, Zhu P, Han Y, Lu C, Wang S, Gu A, Fu G, Zhao R, Song L and Wang X (2009). "Urinary metabolites of polycyclic aromatic hydrocarbons in relation to idiopathic male infertility." *Hum Reprod* 24(5): 1067–1074. [PubMed: 19202139]
- Zhang J, Huang Z, Chen M, Xia Y, Martin FL, Hang W and Shen H (2014). "Urinary metabolome identifies signatures of oligozoospermic infertile men." *Fertil Steril* 102(1): 44–53.e12. [PubMed: 24746742]
- Zhang J, Mu X, Xia Y, Martin FL, Hang W, Liu L, Tian M, Huang Q and Shen H (2014). "Metabolomic analysis reveals a unique urinary pattern in normozoospermic infertile men." *J Proteome Res* 13(6): 3088–3099. [PubMed: 24796210]
- Zhang X, Diao R, Zhu X, Li Z and Cai Z (2015). "Metabolic characterization of asthenozoospermia using nontargeted seminal plasma metabolomics." *Clinica Chimica Acta* 450: 254–261.
- Zhao K, Zhang J, Xu Z, Xu Y, Xu A, Chen W, Miao C, Liu S, Wang Z and Jia R (2018). "Metabolomic Profiling of Human Spermatozoa in Idiopathic Asthenozoospermia Patients Using Gas Chromatography-Mass Spectrometry." *BioMed Research International* 2018: 1–8.

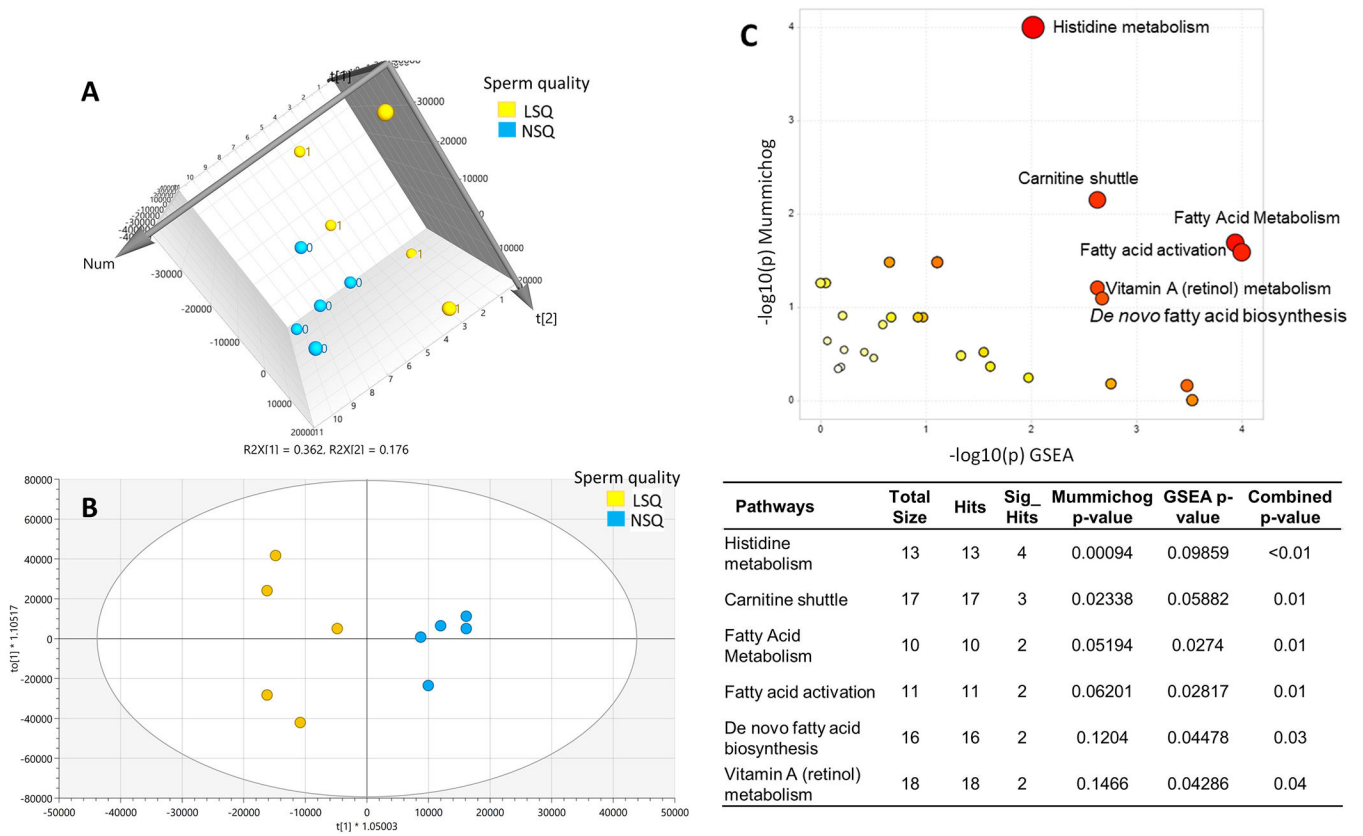


Figure 1. Metabolic profiles (A and B) and pathways (C) differentiated between Low Sperm quality (LSQ, yellow) and Normal Sperm Quality (NSQ, blue).

Score scatter plots of **A**) unsupervised principle component analysis (PCA), $R2X=0.54$, and **B**) supervised orthogonal projections to latent structures discriminant analysis (OPLS-DA), $R2X=0.44$, $R2Y=0.918$, $Q2=-0.41$, for LSQ vs. NSQ. **C**) The enriched metabolic pathways perturbed between LSQ vs. NSQ, conducted by Metaboanalyses 5.0. The scatter plot (upper right) visualizes the analytics of pathway enrichment, while the table (down right) indicates the pathways and p -values of the pathway analysis.

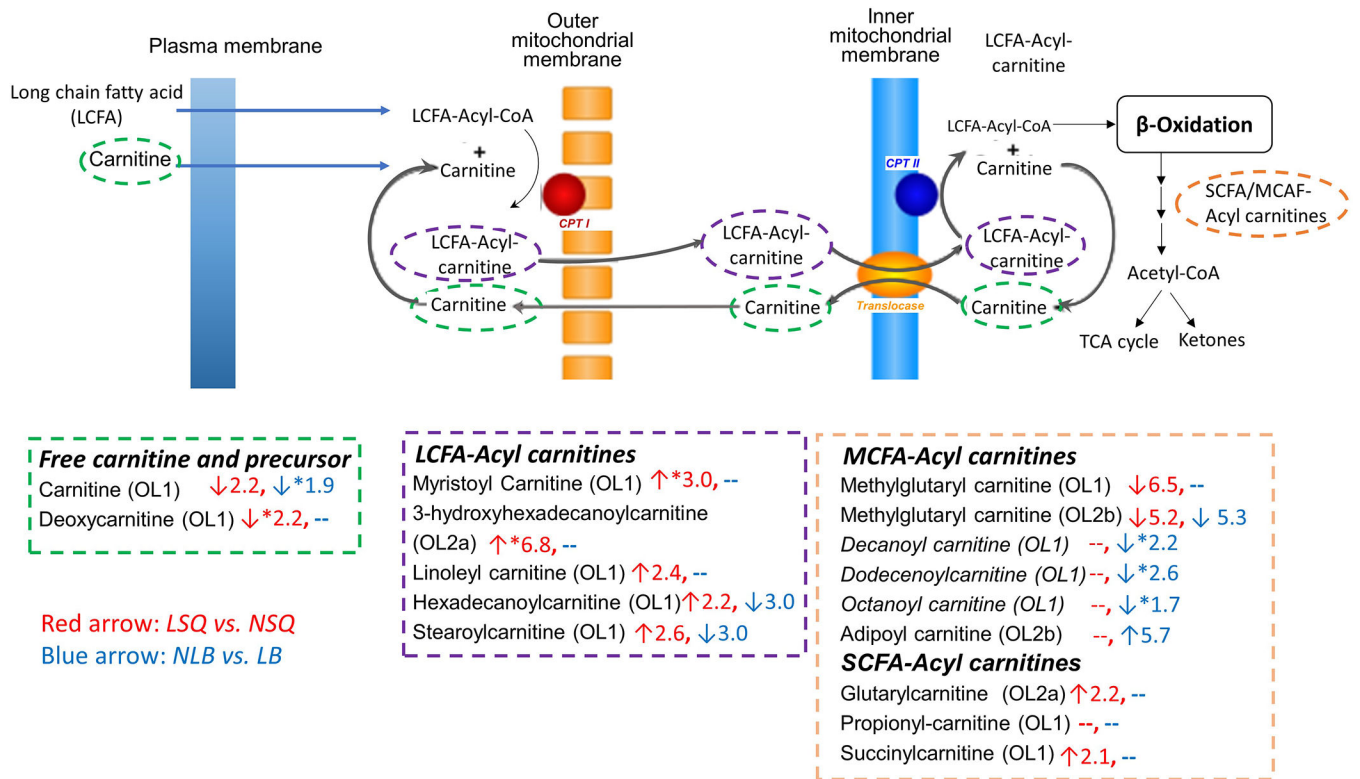


Figure 2. Carnitine shuttle pathway, sperm quality, and birth outcome.

LSQ vs. NSQ: ↓, decreased in LSQ; ↑, increased in LSQ. NLB vs. LB: ↓, decreased in NLB; ↑ increased in NLB. Number indicates the fold change for each of the pairwise comparison. Metabolites with *p<0.1 or |Fold change|>2.0 are labeled with increase or decrease for the comparison. --, indicates no change. **OL1-OL2b**: metabolites in the enriched pathway were verified by matching against the in-house physical standard library via MS, RT, MS/MS (OL1), MS, RT (OL2a), and MS, MS/MS (OL2b)

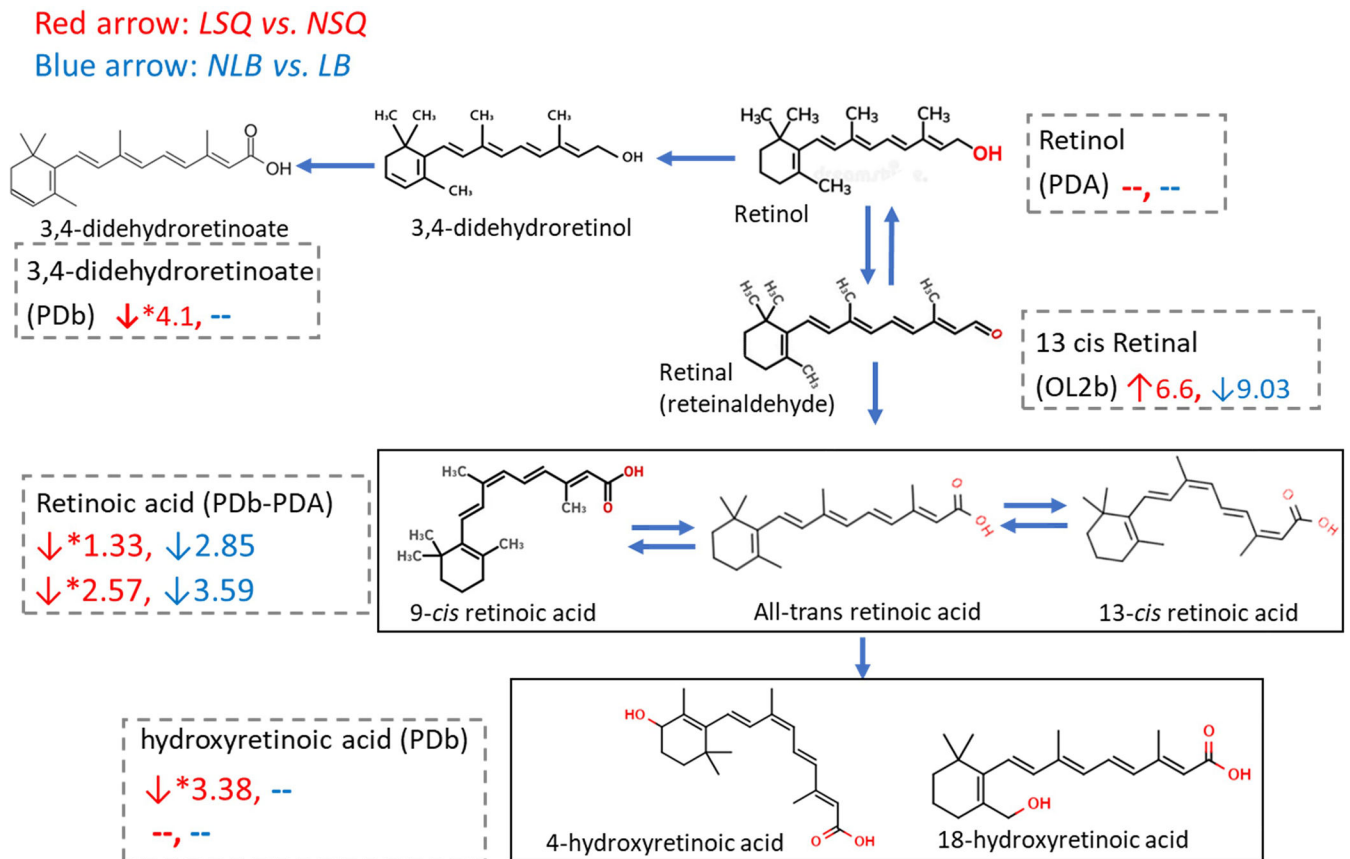


Figure 3. Vitamin A metabolism pathway, sperm quality, and birth outcome.

LSQ vs. NSQ: ↓, decreased in LSQ; ↑, increased in LSQ. NLB vs. LB: ↓, decreased in NLB; ↑ increased in NLB. Number indicates the Fold change for each of the pairwise comparison. Metabolites with * $p < 0.1$, or $|\text{Fold change}| > 2.0$ are labeled with increase or decrease for the comparison. --, indicates no change. **OL2b, PDA, PDb**: metabolites in the enriched pathway were verified by matching against the in-house physical standard library with MS, MS/MS (OL2b), or matching against the public database with MS, experimental MS/MS (PDA), MS, theoretical MS/MS (PDb)

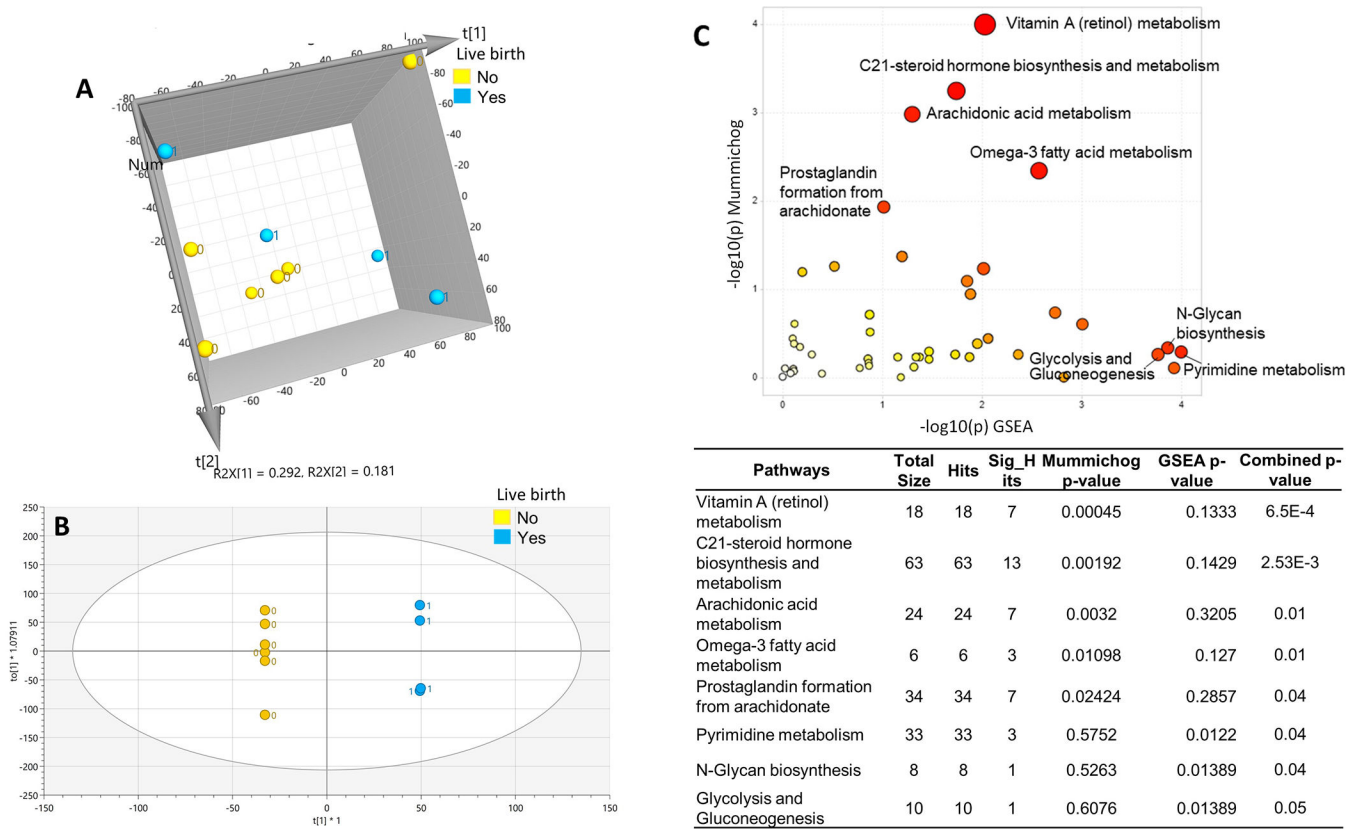


Figure 4. Metabolic profiles (A and B) and pathways (C) differentiated between no live (no, yellow) and live birth outcomes (yes, blue). Score scatter plots of **A**) unsupervised PCA, $R^2X=0.47$, and **B**) supervised OPLS-DA, $R^2X=0.82, R^2Y=1, Q^2=0.84$, for not-live birth vs. live birth. **C**) The enriched metabolic pathways perturbed between no live birth vs. live birth, conducted by Metaboanalyses 5.0. The scatter plot (upper right) visualizes the analytics of pathway enrichment, while the table (down right) indicates the pathways and p -values of the pathway analysis.

Table 1:

Demographic Data

Demographics by analysis	Semen Quality		Live Birth	
	NSQ (n= 5) Mean (Range)	LSQ (n= 5) Mean (Range)	LB (n= 4) Mean (Range)	NLB (n= 6) Mean (Range)
Male Age (years)	43 (32.0 – 61.1)	36.6 (32.2 – 42.3)	36.7 (32.0 – 38.6)	42 (32.2 – 61.1)
Female Age (years)			35.0 (31.9 – 37.9)	37.0 (31.0 – 40.7)
Male BMI (kg/m ²)	26.8 (24.4 – 29.2) [‡]	35 (22.0 – 48.1)	31.6 (22.0–48.1)	31.2 (24.4 – 40.3) [‡]
Female BMI (kg/m ²)			25.4 (20.1–31.8)	32.3 (20.1 – 46.1)
Race	Count			
Male				
Non-Hispanic White	5	3	4	4
Refused	0	2	0	2
Female				
Non-Hispanic White			3	6
Asian			1	0
Baseline semen parameters	Median (Range)		Median (Range)	
Sperm concentration (×10 ⁶ /mL)	89.2 (42.7–131)	14.3 (4.5–24.5) [*]	51.7 (17.4–94)	51.8 (4.5–131)
Normal morphology (%) ³	67.2 (44–85)	27 (9–43.3) [*]	48.1 (28–77)	46.5 (9–85)
Sperm count (10 ⁶ /ejaculate)	8.3 (6.5–11)	3.1 (0.5–6) [*]	3.6 (0.5–7)	7 (3–11)
Oocytes/Embryos	Mean ± SD			
Oocytes Retrieved			8.5 ± 4.1	9.7 ± 9.1
Fertilized Embryos			5.5 ± 2.6	5.2 ± 5.6

[‡]Indicates one BMI is missing

^{*}p<0.01

NSQ: Normal Semen Quality

LSQ: Low Semen Quality

LB: Live Birth

NLB: No Live Birth

Table 2. Exogenous Metabolites Identified in Seminal Plasma via Match against the NC HHEAR Hub In-house Physical Standard Library within MS, RT, and MS/MS

^a Compound Name	^b Ontology	^c Source of Exposure	Structure category
styrene	OL1, OL2b	Environmental	Benzene metabolite
3,5-Di-tert-butyl-4-hydroxybenzaldehyde	OL1	Environmental	BHA metabolite
1,3-xylene	OL1, OL2b (3)	Environmental	Dimethylbenzene
9-Hydroxyfluorene	OL1, OL2a	Environmental	Fluorene
Oxybenzone	OL1	Environmental	Oxybenzone
Mono-(2-ethylhexyl) Terephthalate	OL1	Environmental	Phthalate
Mono-n-octyl phthalate	OL1	Environmental	Phthalate
Mono-2-ethylHydroxyhexyl Terephthalate (MEHHTP)	OL1, OL2a	Environmental	Phthalate
Monoisobutyl phthalic acid	OL1, OL2a	Environmental	Phthalate
Butyl benzyl phthalate	OL1, OL2b	Environmental	Phthalate
Cotinine	OL1	Environmental	Tobacco metabolite
4-hydroxy-3-methoxybenzaldehyde	OL1	Ingested food	Benzaldehyde
4-Hydroxyhippuric acid	OL1 (2), OL2b	Ingested food	Hippuric acid and derivatives
1,3-dimethyluric acid	OL1	Ingested food	Purine derivatives
1,3-dimethylxanthine	OL1	Ingested food	Purine derivatives
1,7-dimethylxanthine	OL1	Ingested food	Purine derivatives
7-methylxanthine	OL1	Ingested food	Purine derivatives
Caffeine	OL1	Ingested food	Purine derivatives
1-methylxanthine	OL1, OL2a	Ingested food	Purine derivatives
3,7-dimethylxanthine	OL1, OL2a	Ingested food	Purine derivatives
4-pyridoxic acid	OL1	Ingested food	Pyridinecarboxylic acids
1-methyl-3-pyridiniumcarboxylic acid	OL1	Ingested food	Pyridinecarboxylic acids
indole-3-acetic acid	OL1, OL2b	Ingested food	Tryptophan-indole derivatives
Acetaminophen	OL1 (2)	Med/Drug	Analgesic
Naproxen	OL1	Med/Drug	Nonsteroidal anti-inflammatory drug
Carboxybupropfen	OL1, OL2a	Med/Drug	Nonsteroidal anti-inflammatory drug metabolite
leucyl-proline	OL1, OL2a(2)	Microbiome	Dipeptide

^a Compound Name	^b Ontology	^c Source of Exposure	Structure category
N-Acetylglucosamine	OL1, OL2a, OL2b(2)	Microbiome	Sugar amide
Indolepropionate	OL1	Microbiome	Tryptophan-indole derivatives
4-Hydroxyphenyllactate	OL1, OL2a	Microbiome	Tyrosine metabolite

^aTable 2 lists the most confident (OL1) exogenous metabolites identified from seminal plasma via matching against the NC HHEAR Hub In-house physical standard library (IPSL). The assigned compound name is based on the name of the reference standards run on the UHPLC-HRMS platform, and some of the isomers (e.g., D, L isomers or R, S isomers) may not be differentiated.

^bOntology levels: OL1, highly confident identification based on matching with IPSL via retention time (RT, with RT error |0.5|), exact mass (MS, with mass error<5ppm), and tandem mass similarity (MS/MS, with similarity >30); OL2a, confident identification based on matching with IPSL via MS and RT; OL2b, annotations for the isomer or derivatives of the compound listed but not the compound itself, based on matching with IPSL via MS and MS/MS. In situations where multiple peaks correspond to a compound via OL1 and (or) OL2a, it may indicate closely eluting isomers or the presence of split peaks and also means high confidence that the assigned compound exists within the biological sample, even if it is unclear which of the multiple peaks best represented that compound. In situations where one or multiple peaks match to a compound via level 2b, it indicates the presence of derivative(s), isomer(s), or conjugate(s) related to the matched compound.

^cThe source of exposure indicates the possible sources where the exogenous parent compounds, or their metabolites derived in the human body after the exposure, might be originated. One exogenous compound or metabolites might originate from multiple sources. Identification and annotation of other exogenous metabolites within less confident levels (OL2 and OL2b) were shown in table S1.