Introduction

Infertility is defined as inability to conceive within 12 months of regular unprotected sexual intercourse (1) and affects 15% of couples (2). A male factor alone contributes to 30% of these cases, and to another 20% of cases when combined with female factors (3). In addition to the medical history and physical exam, the workup of male factor infertility centers on the endocrine evaluation and semen analysis (SA). Though the SA remains a mainstay of the infertility workup, it is an imperfect test with significant variability and only modest predictive value for fertility (4). While additional serum blood tests such as testosterone, luteinizing hormone, follicle stimulating hormone and prolactin may point the clinician towards an underlying etiology for male factor infertility, these tests in combination with the SA still may not provide a diagnosis. This has led to the need to develop adjunct tests to assist with diagnosis of the underlying etiology of infertility and subfertility which may be reflected in an abnormal SA (5).

Genetics have long been recognized as playing a role in male factor infertility. The most common genetic causes include azoospermia factor (AZF) deletions of the Y chromosome, Klinefelter syndrome with an extra copy of the X chromosome, and monogenic disorders including Kallman syndrome and cystic fibrosis. However, these genetic conditions account for only 30% of cases of male infertility and subfertility (6,7), with approximately 50% of male factor infertility still classified as idiopathic, after controlling for hormonal abnormalities or other anatomic etiologies (6). The proportion of these cases that are due to genetic causes remains unknown (7). Spermatogenesis is a complex process involving the coordination of hormones, environmental factors, and over 2,000 different genes (8). Animal models have demonstrated that the knockout of numerous genes leads to subfertility (9) and the desire to
identify an etiology of infertility has driven investigation into identification of the genes involved, epigenetic changes to these genes, the role of these changes on subfertility and infertility, and identification of biomarkers which may improve the diagnostic capability of semen parameters.

The aim of this review is to outline the role of biomarkers in the diagnosis of male infertility, and to define the current role and future use of genomics, epigenetics, metabolomics and proteomics in the workup of the infertile or subfertile male.

Biomarkers in male infertility

Biomarkers

A biomarker is defined as a quantifiable and measurable biological parameter which serves as an index for certain disease states (10). The ideal biomarker would be cost-effective, minimally invasive, accurate, and able to detect disease in the early stage. In the field of infertility, the SA acts as a basic biomarker for male infertility. However, the SA remains an imperfect test as it may be affected by physiologic parameters including abstinence interval, systemic illness, medical comorbidities, lifestyle factors including diet and activity, and environmental exposures (11-17). Given the high variability in semen parameters, multiple adjunct biomarkers have been developed over the years, including antisperm antibody (ASA) testing, acrosome reaction testing, sperm penetration assays, sperm-zona pellucida binding tests, hyaluronan binding assays, and deoxyribonucleic acid (DNA) damage testing (18-21).

ASA testing was one of the first widely used biomarkers in the field of male infertility. ASAs develop with disruption of the blood-testis barrier as in the case of trauma, infection, testicular torsion, or even following vasectomy (22). ASAs have been identified in approximately 12% of infertile men, versus 2.5% of fertile men (22). While ASAs are thought to have a negative impact on fertility by impacting the quality of sperm, many studies on the effect of ASA on semen parameters have shown conflicting results (18). A recent meta-analysis by Cui et al. suggested that ASA may have a negative effect on sperm concentration and progressive motility (18). However, other studies have shown that the presence of ASAs does not translate into consistently altered rates of spontaneous, intrauterine insemination (IUI)-assisted, in vitro fertilization (IVF)-assisted, or intracytoplasmic sperm injection (ICSI)-assisted pregnancies (23-25), which brings into question the utility of ASA as a biomarker for male infertility.

Numerous sperm function tests have been developed, including assays for detecting the acrosome reaction, sperm penetration, sperm-zona pellucida binding, and hyaluronan binding. In the era of ICSI, functional sperm tests have fallen out of favor as the functional aspects of the sperm can be bypassed with direct intracytoplasmic injection. However, these tests retain some utility as they may allow the provider to direct a patient away from futile rounds of IUI or IVF in certain situations (19). Sperm function tests have been shown to be highly predictive of the sperm’s ability to fertilize an egg, with the downside of these tests being the requirement for intact sperm or eggs, as well as the time consuming and technically challenging nature of these tests, thus limiting their utility as a biomarker (19).

DNA fragmentation has been associated with impairments in natural conception, IUI, and IVF (3,26,27), with inconsistent effects on outcomes with ICSI (28). Recent guidelines suggest performing DNA fragmentation screening on men with a clinical varicocele and borderline to normal semen parameters to help identify optimal surgical candidates, men with unexplained infertility, recurrent spontaneous pregnancy loss, recurrent IUI, IVF or ICSI failure, or patients with lifestyle risk factors including smoking, advanced age, obesity, or environmental exposures (29). Based on these guidelines, O’Neill et al. proposed a treatment algorithm for patients with poor IUI outcomes and suggested that these patients undergo DNA fragmentation testing (28). Patients with a normal DNA fragmentation assay proceed to IVF, and patients with an abnormal assay proceeded to ICSI. If subsequent IVF or ICSI was unsuccessful, patients were encouraged to pursue epididymal or testicular sperm extraction, as studies have shown that levels of DNA damage are higher in ejaculated sperm (30). This group suggested that this algorithm could be used to guide the decision regarding the assisted reproductive technology (ART) treatment approach, in hopes of decreasing the number of futile treatment cycles and shorten the time to conception (28). When determining the utility of proceeding with multiple rounds of ART versus preceding directly to testicular sperm extraction for patients with high DNA fragmentation, the provider must utilize shared decision making with the patient and the partner, and fully consider the costs of multiple rounds of ART and risks incurred by the female partner weighed against the risks to the male patient undergoing extraction. A recent review of the literature by Awaga et al. showed that the use of testicular spermatozoa, rather than ejaculated...
spermatozoa, results in a higher probability of pregnancy in patients with high DNA fragmentation index, but only in the setting of oligospermia (31). Given this, it may be beneficial to proceed with sperm extraction and ICSI for men found to have high levels of DNA fragmentation in the setting of oligospermia, rather than attempting ICSI with ejaculated spermatozoa.

The limited utility of the historically used biomarkers in identifying the underlying etiology of male infertility has prompted the search for markers which would be easier to employ, more cost effective, and more predictive of reproductive outcomes. Since the mapping of the human genome, much interest has been paid to the field of genetic medicine, with treatment modalities tailored to a patient's specific genetic makeup. The study of genomics, including epigenetics, metabolomics, and proteomics, holds promise as a potential area of targeted diagnostics in the setting of male infertility. The following sections will review the recent advances in, and future directions of, genomics and its sub-fields.

**Genomics**

The field of genomics centers on the systematic evaluation of an organism’s complete DNA sequence (Figure 1) (32). Since the advent of karyotyping, the hunt for genetic causes of infertility has gained interest. Following the initial cloning of the Y chromosome in 1992, the first genetic marker of infertility, AZF, was identified (33-35). It was soon determined that the presence of specific sub-regions of AZF is critical for proper spermatogenesis. Deletion of the AZFa region leads to deletion of the USP9Y and DBY/DDX3Y genes resulting in Sertoli cell only syndrome. Deletion of the AZFb region results in deletion of the RBMY1 and PRY genes which causes arrest of spermatogenesis. Though the AZFc region also contains genes important in spermatogenesis, microdeletions in this region typically allow for limited spermatogenesis, therefore making these patients candidates for testicular sperm extraction (7,36). These microdeletions are typically assessed using polymerase chain reaction (PCR).

Another common application of the field of genomics as it relates to male infertility is in the workup of congenital bilateral absence of the vas deferens (CBAVD). CBAVD occurs in approximately 1% of infertile men, is due to mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene, and results in obstructive azoospermia (7). Standard CFTR screening via PCR is...
recommended for patients presenting with CBAVD. In cases where standard CFTR screening is negative, and the female partner is a carrier, full gene sequencing may be considered (37).

In its most basic form, genomics includes standard karyotyping. It has been reported that chromosomal abnormalities are 8–10 times more common in infertile men, with anomalies occurring in approximately 3% of patients with oligospermia, and 19% of patients with azoospermia (7,38). The karyotype is one of the most common tests recommended for the workup of male infertility and is recommended by the American Urological Association (AUA) in patients with severe oligospermia of <5 million sperm/mL and non-obstructive azoospermia (37). However, studies have suggested that chromosomal abnormalities are identified in patients with less severe oligospermia, and that this stringent cutoff may fail to capture a significant number or patients with chromosomal abnormalities in the setting of unexplained infertility (39). A recent study of over 16,000 men showed that the frequency of major chromosomal anomalies in men with primary infertility was 5.5%, and 2.5% of men with secondary infertility, with fewer abnormalities noted in patients with normal semen analyses (SA) (40). Patients with a history of adverse pregnancy outcomes were found to have chromosomal abnormalities in 1.8% of cases regardless of SA results. Out of all patients evaluated, 4% were noted to have chromosomal abnormalities. 64% of these cases involved changes in autosomal chromosome number or structure, and 31% involved changes in sex chromosome number or structure. Even in patients with a normal SA, chromosomal abnormalities were still detected in 0.02% of patients with primary infertility and 0.3% of patients with secondary infertility, suggesting that chromosomal anomalies may contribute to infertility even in the presence of a normal SA (40). In addition to well-described abnormalities in sex chromosomes in patients with infertility, this group found a high occurrence of chromosomal polymorphic variations. These variations have been shown to be associated with poor spermatogenesis and infertility, though the exact mechanism of these remains unclear (40).

While cytogenetic karyotyping of peripheral blood lymphocytes can provide basic information regarding overall chromosome number and structure, fluorescent in situ hybridization (FISH) is gaining popularity in the field of genomics to identify specific chromosomal regions which may be associated with deleterious effects on semen quality and fertility. FISH has been used since the 1990s to assess sperm for aneuploidy of chromosomes X, Y, 13, 18 and 21, which are thought to be most frequently involved in spontaneous miscarriage (41). The downfall of FISH analysis on sperm is the high cost associated with the testing, and the fact that the sperm which are tested are not able to be utilized for ART. However, knowledge of the presence of spermatic aneuploidy may help in counseling of patients regarding the risks of ICSI and whether to proceed with preimplantation genetic testing (3).

Mayeur et al. have suggested that a criterion of <5x10^6 progressive spermatozoa retrieved after sperm preparation may be a better threshold for genetic screening in patients with infertility, to better capture patients with Robertsonian and reciprocal translocations which may be playing a role in infertility (39). Their studies, along with others, have identified numerous chromosomes (mostly commonly chromosomes 1, 2, 6, 9, 16, and 17) with redundant regions or polymorphisms which were associated with infertility and poor or very poor progressive spermatozoa after processing. This suggests that these regions may contain genes which are important for proper sperm function (39,40).

As additional genes are being identified as contributing to male infertility, FISH may begin to play a bigger role in the assessment of peripheral lymphocytes for Robertsonian translocations and reciprocal translocations which may impact fertility, without having to test, and thus destroy, viable sperm which may be present in small numbers and utilized in ART. One challenge to using FISH as a diagnostic tool, however, is that the DNA sequence of interest must be known prior to the application of the fluorescent label. It should be recognized that due to the complex interplay of the thousands of genes involved, the phenotype of infertility or subfertility may not be uniform, which makes the search for the genomics involved more difficult (7).

Thousands of genes are involved in spermatogenesis. With the advent of whole-genome assessments using comparative genomic hybridization (CGH), high-throughput sequencing technologies or next generation sequencing, and analysis of single nucleotide polymorphism (SNP) arrays, it is becoming possible to study large amounts of genetic material for relatively small sequences of abnormalities which may impact fertility. These technologies have shown promise in terms of effectiveness and affordability in identifying potential biomarkers and candidate genes in male infertility. An excellent review by Krausz and Riera-Escamilla detailed these techniques, described their utility, and several of the genes implicated in
male infertility (42). Okutman et al. have recently compiled a list of 31 genes for which the data is strong to support their role in abnormal spermatogenesis or fertilization failure (43). Though their study was published in November of 2018, Cannarella et al. have even more recently published a list of 60 genes involved in spermatogenic failure (44), which speaks to the exponential expansion of our knowledge of the genes involved in male infertility. Similarly, Patel et al. recently demonstrated the utility of using next generation sequencing to identify genes known to impact male fertility, suggesting that next generation sequencing may be used to more rapidly identify a larger number of genes thought to be related to male infertility with greater accuracy and at significantly reduced cost compared to traditional standard analysis (45). However, though we continue to identify genes involved in infertility, these genes remain primarily of diagnostic interest, and few have resulted in therapeutic targets for the treatment of male infertility. As we gain a deeper understanding of the genes involved, karyotyping, FISH, and gene sequencing may be able to further stratify male infertility into phenotypic subsets.

Epigenetics:

Genetic epigenesis is the process of modifying the DNA of a cell to ultimately impact cell differentiation. Epigenetic changes to DNA, including DNA methylation, heterochromatization, histone modification, timing of DNA replication, and nucleosome positioning cause heritable changes to cells without altering the DNA sequence (Figure 1) (46). This process is well known to impact gene expression in somatic cells, and it is being shown to have a role in the sperm as well. Spermatozoa are particularly susceptible to oxidative stress, and reactive oxygen species (ROS) can induce epigenetic changes in sperm DNA (47). It is thought that sperm histone modifications, DNA methylation, and protamination may impact transcription during embryogenesis (48), and some evidence suggests that these changes could result in miscarriage (49).

The use of epigenetic changes in DNA as a marker to identify male factor infertility is difficult, as these changes may have a subtle or even insignificant biological impact, and multiple different changes may lead to similar infertility phenotypes. While determining causal relationships between these changes and embryologic outcomes remains challenging, several spermatoc epigenetic changes have been associated with sperm abnormalities. For example, individual sperm with low levels of protamine-1 and protamine-2 content have been shown to have increased susceptibility to DNA damage and decreased viability (50). Spermatic DNA hypomethylation of the gene H19, and hypermethylation of genes MEST and SRNPN have been associated male infertility (51). Additionally, alterations in DNA methylation at specific genomic loci have been associated with difficulty conceiving (52), and DNA methylation and gene transcription patterns may be predictive of success rates of IUI and IVF (53,55). It has also been shown that patients with a history of oligoasthenoteratozoospermia (OAT) have higher rates of DNA methylations and altered gene expression in embryos that have undergone preimplantation genetic screening following IVF (49). DNA methylation as a marker of male infertility is an attractive target, as DNA methylation remains stable throughout spermatogenesis, unlike RNA transcription and histone modification (48). Whether epigenetic changes in sperm represent causative etiologies of infertility and subfertility, or whether these changes reflect an alternate underlying pathology, remains to be determined.

With the growing concern about the decline in male fertility (54), much research is being targeted towards the impact of environmental factors on sperm quality. It is known that oxidative stress and urinary phthalates and phthalate alternatives impair sperm DNA methylation (55). There is concern that environmental factors, particularly plastic-derived compounds, may generate oxidative stress leading to heritable abnormal DNA methylation (56). In addition to ingested environmental contaminants, the increase in rates of obesity and diabetes have been associated with male infertility and decreased sperm quality (55,57). Though this is related in part to obesity-induced endocrine disruption, obesity is thought to also result in epigenetic changes in sperm, the impacts of which may lead to health impairments in offspring (58). Currently, the only commercially available screening tool for epigenetic changes in sperm is “Seed” developed by Episona Inc., developed in 2016. With the growing interest in epigenetics as a marker for male infertility, it is expected that additional companies will develop testing for these factors.

Post-transcriptional epigenetics may also play a role in male infertility. This is largely performed by short non-coding ribonucleic acids (sRNAs), including microRNAs, which modulate protein translation (59). These sRNAs are expressed in germ cells and are required for normal spermatogenesis, and numerous microRNAs have been found to be up- or down-regulated in the sperm and seminal plasma of men with subfertility and infertility.
Additional modifications to the sperm epigenome take place in the epididymis, as the spermatozoa undergo further maturation. Modification of sperm sRNAs during maturation in the epididymis has been linked to normal fertilization potential, embryonal development and offspring phenotype (61). It has been hypothesized that the sRNA profile of spermatozoa is altered by interaction with epididymosomes, membrane-bound extracellular vesicles which encapsulate complexes of sRNAs and proteins and affect the luminal contents of the epididymis (61). Multiple environmental stressors have been shown to alter the sRNA profile of sperm, and it is thought that this occurs through packaging of altered sRNAs under stressful conditions into epididymosomes, effecting which sRNAs are delivered to maturing sperm (61). This is of clinical interest, as it suggests that using testicular-derived sperm for embryo fertilization may bypass the negative downstream effects that altered sRNAs may have on normal fertilization and embryo phenotype.

Proteomics

Proteomics involves the study of the structure and function of the expressed complement of proteins (Figure 1) (62). This concept is important, as gene expression does not necessarily translate into protein expression, and the protein complement varies from cell to cell. This makes proteomic analysis an attractive target for biomarker development. However, the protein complement of the ejaculate is complex and is added to by the epididymis, seminal vesicles, prostate and bulbourethral glands (63). A systematic review by Amaral et al. in 2014 identified over 6,000 different sperm proteins involved in multiple functional pathways (64), and the location of these proteins on the sperm have been linked to their site of origin in the testis or epididymis (65). With the development of mass spectrometry systems, the sensitivity and specificity of, as well as the ease of, protein identification has improved (65). These techniques can identify proteins within the seminal plasma, sperm membrane, various sperm regions, or mitochondria, or perhaps more importantly protein-protein interaction networks (65). It is important to understand the individual proteins involved in various fertility processes, as well as the way these proteins interact with each other, as alterations in individual proteins may alter the downstream effect of their function and interaction, resulting in various phenotypes of infertility. Many studies have been conducted and have identified a wide array of protein complement patterns present in the setting of recurrent IVF failures, asthenozoospermia, OAT, NOA, globozoospermia, teratospermia, elevated levels of DNA fragmentation and ROS, and clinical varicoceles (66-73).

The utility of analyzing the semen proteome as a biomarker for infertility is complicated, as semen specimens are highly variable within individuals and are influenced by environmental factors, as well as post-ejaculation physiological changes to sperm (74-76). In a model of mice exposed to endocrine disruptor chemicals including compounds found in plastics, numerous proteins were found to be up- or down-regulated in the testis, the majority of which were involved in cell death (77). Additionally, it has been difficult to find proteins and peptides in sperm and semen which are reproducibly altered in various phenotypes of infertility (78). Barrachina et al. recently described a novel strategy of analysis of quantitative proteomic data based on the establishment of stable-protein pairs, which allowed for identification of highly stable proteome signatures in men with normal semen parameters. However, when they applied this analysis technique to infertile men classified according to abnormal semen parameters, a high level of heterogeneity was identified, suggesting the current classification system may not be adequate to describe various infertility phenotypes (78).

A review of the literature published as recently as January of 2019 resulted in numerous articles in the field of proteomics, demonstrating the growing interest in proteomics in the study of male infertility. One recent study comparing the proteome of the ejaculate from fertile men with high ROS levels to that of fertile men with normal ROS levels found that numerous antioxidant proteins were present in higher levels in men with elevated ROS levels, suggesting that there are specific proteins which may be upregulated in fertile men with high ROS levels acting in a protective fashion to mitigate the negative effects of ROS and preserve fertility (79). Additionally, recently heat shock protein A4L has found to be downregulated in sperm in patients with asthenozoospermia, resulting in decreased sperm motility and sperm-oocyte penetration (80). This is just a small sample of the number of studies with promising results targeting specific proteins which appear to be located in the sperm and seminal plasma which may be developed as biomarkers for male infertility in the future.

Metabolomics

Metabolomics is the study of cellular metabolic products
As metabolic products are the downstream products of gene expression, they may be a closer representation of an actual phenotype than genomics, proteomics or epigenetics. As previously discussed, oxidative stress can induce epigenetic changes in sperm, and leads to formation of ROS. Elevated ROS levels are seen in infertile men, and are associated with impaired sperm morphology, concentration, and motility, as well as increased levels of DNA fragmentation. ROS have gained interest as potential biomarkers of male infertility, and the study of ROS in semen is likely the most widely known application of metabolomics in male infertility. Serum samples have also been assessed for metabolomic alterations to identify a pattern in patients with NOA. Interestingly, a recent study by Gilany et al. used Raman spectroscopy to identify ROS in the ejaculate of 20 patients with non-obstructive azoospermia (NOA) to detect spermatogenesis. Their assay was able to distinguish patients who had spermatogenesis on testicular sperm extraction (TESE) from those without spermatogenesis and was further able to sub-classify patients with a negative TESE into patients with hypospermatogenesis, sperm maturation arrest, and Sertoli-only syndrome. Another study by Gilany et al. analyzed seminal plasma of men with NOA who underwent subsequent TESE. Gas chromatography-mass spectrometry (GC-MS) was used to perform untargeted metabolomics profiling, and this approach was able to distinguish a differential metabolomic profile between NOA men who had sperm identified on TESE, and those with no sperm identified on TESE. If these results are determined to be reproducible, this technology could impact the pre-operative screening and counseling of patients with NOA.

Other metabolic products of spermatogenesis are being investigated as biomarkers for male infertility. A study by Zhao et al. utilized GC-MS to identify sperm metabolites associated with idiopathic asthenozoospermia. They identified numerous metabolites which were either upregulated or downregulated, involved in amino acid, energy, and nucleoside metabolism. Qiao et al. also analyzed seminal plasma using GC-MS and identified 44 metabolites that were expressed in altered amounts in patients with unexplained infertility with normal SA as compared to fertile controls, predominantly involving alterations in amino acid metabolism. Their metabolite profiling platform distinguished 82% of infertile patients from fertile controls and based on these results they suggested that their profiling platform may be used for diagnosis of idiopathic infertility. Similarly, a recent meta-analysis by Guo et al. identified 16 metabolites in seminal plasma which were significantly different between men with normal BMI and normal semen parameters versus obese men with abnormal semen parameters. Specifically, the pathways of arginine, proline, beta-alanine and glutathione metabolism were significantly altered, with significantly higher concentrations of spermidine and spermine identified in obese men with abnormal semen parameters. Spermidine and spermine have been shown to be important in normal fertility. Overall, metabolomic profiling of the ejaculate is a rapidly growing field, and shows promise for the development of accurate biomarkers for infertility phenotypes.

While analysis of the proteomic and metabolomic profiles of sperm and semen remain an attractive minimally invasive way to diagnose male infertility with the potential to distinguish infertility phenotypes, clinical application of these profile panels remains limited, in part due to the wide range of proteins, peptides, and metabolites found in semen samples which may be affected by various environmental factors to varying degrees between individuals. Significantly more research must be done before these profile panels can implemented clinically on a large scale.

Conclusions

Male infertility is a complex disease involving the interplay of thousands of genes, proteins, and metabolites. The study of these factors, along with epigenetic changes to the sperm DNA, remains an attractive target for the development of specific biomarkers which may be able to identify the etiology of male infertility in a minimally invasive way. Though many factors have been shown to be involved in infertility, developing a profile of biomarkers that are consistent across infertility phenotypes remains elusive. As our understanding of the factors involved grows, it will become possible to implement our knowledge of these factors into the clinical decision-making process.

Acknowledgments

None.

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.
**Ethical Statement:** The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**References**


