Published in final edited form as:

Sci Transl Med. 2015 July 8; 7(295): 295re6. doi:10.1126/scitranslmed.aab1287.

Absence of Sperm Rna Elements Correlates With Idiopathic Male Infertility

Meritxell Jodar^{1,2}, Edward Sendler^{1,2}, Sergey I. Moskvovtsev^{5,6}, Clifford L. Librach^{5,6}, Robert Goodrich^{1,2}, Sonja Swanson⁵, Russ Hauser^{3,4}, Michael P. Diamond^{1,7}, and Stephen A. Krawetz^{1,2,*}

¹Department of Obstetrics and Gynecology, Wayne State University, Detroit, Michigan 48201

²Center for Molecular Medicine and Genetics, Wayne State University, Detroit, Michigan 48201

³Vincent Memorial Obstetrics and Gynecology Service, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts

⁴Departments of Environmental Health and Epidemiology, T.H. Chan School of Public Health, Boston, Massachusetts 02115

⁵CReATe Fertility Centre, Toronto, Ontario, Canada

⁶Department of Obstetrics and Gynaecology, University of Toronto, Ontario, Canada, Department of Obstetrics and Gynaecology, Women's College Hospital, Toronto, Ontario, Canada M5G-1N8

⁷Department of Obstetrics and Gynecology, Georgia Regents University, Augusta, Georgia, 30912

Abstract

Semen parameters have been used to diagnose male infertility and specify clinical interventions.. In idiopathic infertile couples, an unknown male factor could be the cause of infertility even when the semen parameters are normal. Next Generation Sequencing of spermatozoal RNAs has provided an objective measure of the paternal contribution that may be able to help guide the care of these couples. Spermatozoal RNAs from 96 couples presenting with idiopathic infertility were assessed in the context of fertility treatment and final reproductive outcome and sperm RNA elements (SREs) reflective of fecundity status were identified. The absence of required SREs reduced the probability to achieve live birth by Timed Intercourse (TIC) or Intrauterine Insemination (IUI) from 73% to 27%. However, the absence of these same sperm RNA elements does not appear to be critical when assisted reproductive technologies (ART) such as *In Vitro* Fertilization (IVF) with or without Intracytoplasmic Sperm Injection (ICSI) are employed.

*Corresponding Author: C.S. Mott Center, 275 E. Hancock Ave., Detroit, MI, 48201, U.S.A., steve@compbio.med.wayne.edu, Phone: 313-577-6770, Fax: 313-577-8554.

Author contributions: M.P.D., S.I.M., C.L.L., S.S. and R.H provided and characterized the clinical samples used in the study. M.J and R.G. isolated the sperm RNA and prepared the deep sequencing libraries. E.S., M.J., and S.A.K contributed to the biocomputational analysis of the RNA-seq data and the interpretation of the results. The manuscript was collaboratively written by E.S., M.J. S.A.K., M.P.D., S.I.M., C.L.L., R.H, and R.G. S.A.K. and M.P.D. directed the data analysis, writing and editing of the manuscript. All authors critically reviewed and approved the final version of the manuscript.

Competing interest: We declare no competing interests.

Data and materials availability: RNA-seq data sets are available at the GEO (NCBI) repository (GSE65683).

Approximately 30% of the idiopathic infertile couples presented an incomplete set of required SREs suggesting a male component as the cause of their infertility. Similarly, analysis of couples that failed to achieve a live birth when presented with a complete set of SREs suggested that a female factor was perhaps involved as confirmed by their diagnosis. The data presented from this study suggests that SRE analysis has the potential to inform on the individual success rate of different fertility treatments to reduce the time to achieve live birth.

Keywords

idiopathic male infertility; ART; sperm RNA; Next Generation Sequencing

Introduction

Approximately 13% of the general reproductive age population are challenged with fertility problems(1). The American Society for Reproductive Medicine estimates that male factors and female factors contribute approximately equally to this condition, with the remaining one-quarter likely a combination of factors from both partners. After 12 months of unprotected intercourse without pregnancy, affected couples typically begin to seek care and explore the possibility of fertility treatments(2).

Today upwards of one-percent of the children born in the United States of America are conceived using Assisted Reproductive Techniques (ART)(3). Typically, to establish the appropriate clinical treatment and minimize the risk of failure, an extensive evaluation of the female, and to a lesser extent the male, is undertaken. If no severe male or female factors are detected, fertility treatments such as Timed Intercourse (TIC) or Intrauterine Insemination (IUI) are recommended in combination with ovarian stimulation. After 3 or 4 unsuccessful IUI cycles or if either severe male or female factor is detected *In Vitro* Fertilization (IVF) with or without Intracytoplasmic Sperm Injection (ICSI) are suggested.

Initial male factor assessment includes a review of reproductive history (time of subfertility, existence of previous pregnancies and sexual function), family history (consanguinity and familiar infertility history), relevant diseases (diabetes and mumps among others) and exposure to factors that negatively impact fertility (drugs, lifestyle and occupation) along with a comprehensive physical examination. The male contribution is further evaluated by semen analysis, with intra-individual variation gauged through the results of two semen analyses separated by a period of up to one month(2). Assessment primarily relies on a defined series of semen parameters that include volume, sperm concentration, sperm motility, and sperm morphology. Other specific measures that may complement the workup include DNA fragmentation, the presence of antisperm antibodies, endocrine status, genetic and cytogenetic markers such as AZFa or AZFb Y microdeletions representative of azoospermia. Although the evaluation of general semen parameters like sperm count, motility, and morphology may be useful in the diagnosis of obvious cases of male infertility where specific etiologic factors may be apparent, no single or set of semen parameters are highly predictive of male fertility status within the general population(4). Current clinical practice focuses on whether there are sufficient spermatozoa with satisfactory motility and

morphology to reach and likely fertilize the oocyte. Their utility in selecting the least invasive fertility treatment for idiopathic infertile couples appears limited(5).

Spermatozoa are not just a vehicle that simply delivers the male genomic contribution to the oocyte. Upon fertilization, the spermatozoon provides a complete, highly structured and epigenetically marked genome, that together with a defined complement of RNAs and proteins, play a distinct role in early embryonic development(6, 7). While several studies have explored the effect of genetic variants such as SNP's(8), copy number variants(9), differential genome packaging(10), differential methylation(11), proteomic changes(12), and differential sperm RNAs(13, 14) in male infertility, comparatively few have examined their effect within the context of the reproductive clinic(15-19).

Characterization of the RNAs retained in sperm by Next Generation Sequencing (NGS) has recently been reported(20-22). In contrast with earlier array-based approaches, RNA-seq has revealed a rich and complex population of unique coding and non-coding transcripts such as sperm-specific isoforms, intronic retained and otherwise unannotated elements, and long and small non-coding RNAs(20-22). The large number of unique sperm transcripts is suggestive of regulatory roles(20, 22) influencing fertilization, early embryogenesis, and the phenotype of the offspring(20, 23). The utility of spermatozoal microarray-based approaches to predict the outcome of different fertility treatments has met with varying degrees of success(17, 18). The intricacies of spermatozoal RNAs as revealed by NGS analysis(22) suggest that this technology is much better suited to the task. The objective of this initial study was to evaluate the diagnostic potential of NGS as a prognostic assay of spermatozoal RNAs that can predict the birth outcome after different fertility treatments.

Results

Identifying SREs, sperm RNA elements required for natural conception

The ability of spermatozoal RNAs to predict the Live Birth (LB) for various fertility treatments was assessed within the context of the idiopathic infertile couple to ascertain if the underlying cause could be attributed to a male factor. As summarized in Table 1A, no significant differences between the choice of treatment modality as a function of the different patient variables including age or any of the semen parameters were observed, consistent with idiopathic infertility (one tailed ANOVA or Kruskal–Wallis test, $p>0.05$). Female age was significantly higher in couples that did not achieve pregnancy (Table 1 B; two tailed t-test, $p=0.024$) and this could attributed to unsuccessful IVF/ICSI (Table 1 C; two tailed t-test, $p=0.004$). A set of SREs required for live birth by natural conception were identified within positive group control I (LB by TIC during the first spermatogenic cycle and first attempt). Of the 278,605 sperm RNA elements surveyed, only elements that ranked above the 99th percentile and were essentially at equivalent levels across group control I samples (IQR negative; outside InterQuartile Range $1.5X (Q3-Q1)$) were defined as SREs required for natural conception. A total of 648 met the stringent criteria to be classified as required SREs (above the 99th percentile rank present at a constant level in the control group). Nine of these 648 SREs corresponded to intergenic regions, 12 to sperm-specific intronic elements, and 42 were within 24 different non-coding RNAs all of which are likely regulatory. However the majority (585) were within exonic regions of 262 different genes,

of which 40% were ontologically classified as associated with spermatogenesis, sperm physiology, fertility and early embryogenesis prior to implantation (Figure 1).

Ability of SREs to predict fertility treatment outcome

To discern whether SREs were indicative of fertility treatment outcome, all Group II test samples were assessed as a function of the required SRE abundance within these samples. As shown in Figures 2A (left and middle panels) all 648 required SRE were present in the control group and in 37 of the 56 samples within group II. As summarized in Figure 2B, the samples presenting all SREs have a 72% rate of success to achieve a LB within the first two treatment cycles (6 months). In comparison, as in Figures 2A (right panel), one third of the samples from group II have at least one SRE absent (Supplemental Figure 1). Although the proportion of male or female factor ascribed to idiopathic infertility remains to be established, the proportion of patients identified with some SRE absent is similar to the expected rate (24, 25). No correlation between the number of SREs absent and semen parameters and age of partners were observed (Supplemental Figure 2).

Samples with all SREs present have a similar high rate of LB for both TIC/IUI and ART, which is 77 and 75% respectively (Figure 3A). However the absence of some of the SREs reduces the LB rate by TIC/IUI from 77% to 27%, while the LB rate remains similar for ART (78%; Figure 3B). As observed in Figure 3C, patients with all SREs are more likely to achieve LB by TIC/IUI as compared to those with some SRE(s) absent (two tailed Fisher exact test, $p = 0.012$). These significant differences were supported by a power of 0.7 and α error of 0.029. In comparison, when treated by ART, no differences in the number of SREs that were absent were observed when the LB and NLB groups were compared (two tailed Mann-Whitney, $p=0.783$). This is consistent with the view that ART may be able to “rescue” some otherwise impaired sperm functions such as transit to oocyte and/or fertilization as represented by functions of the genes corresponding to the missing SREs. Interestingly, six of the Group II couples failed to achieve a LB even by ART. A significantly higher female age was observed in this group ($p=0.004$ two tailed t test; Table 1C) suggesting a potential age related female factor but only three of the six subjects were over 35 years of age. Of the couples that failed to achieve LB by ART with partners ≥ 35 years of age two did not have a complete set of SREs. These included *NDRG1* (stress response), *TESK1* (kinase), *DBN1* (stabilizes gap junctions) and *CAMTA2* (calcium dependent transcription factor) that are associated with embryogenesis or implantation.

Testing SREs between clinics and against couples with a known female factor

The results of an independent group of test samples (group III) are summarized in Table 2. Group III-i was comprised of five couples from an independent fertility clinic. The presence of all SREs was confirmed and in all cases LB was achieved. In two cases, LB was achieved spontaneously or by IUI while the remaining three cases directly employed ART. The presence of all SREs in the samples suggests that SREs may provide a measure of male fecundity across clinics. Group III-ii was comprised four samples in which an unknown female factor was suggested. This was supported by the observation that in three of the couples (Table II samples 6-8) since after a series of failed fertility treatments, a gestational carrier (GC) yielded successful results. The causative female factor was confirmed in

Sample 9 (Table 2 ii) by the presentation of stage 2 endometriosis. All SREs were present in three (samples 7-9) of the four samples (samples 6-9). Two SREs were lacking in the sample 6 suggesting that they could be rescued by ART. The spectrum of test set analyses directly speaks to the utility of sperm RNA as a marker of fecundity.

Discussion

Except for cases where examination of sperm reveals gross deficiencies in count, motility, or morphology, it is clear that current standard tests have a limited capacity to discern male factor infertility for couples presenting with idiopathic infertility(26) and thereby be predictive of fertility treatments. This observation emphasizes the need to develop alternative strategies like NGS affords. In comparison to somatic cells RNA-seq data from sperm is characterized by the variability of absolute transcript abundance between samples due to the physiological fragmentation of spermatozoal RNAs as part of the extrusion of the cytoplasm as each spermatozoon is formed. However, this did not affect the 648 comparatively small exon-sized SREs that show no variability in 32 of the 35 TIC/IUI cases that achieved LB. It appears that the SREs that were absent in the 3 LB TIC/IUI cases were not diagnostic and perhaps could be removed from consideration. Their validation in a larger population will clarify which are essential for diagnosis and may be reflective of their role towards the birth of a healthy child.

Standard clinical practice often utilizes TIC or IUI as the initial treatment for couples with idiopathic infertility. It is not until after 2-4 IUI cycles have failed that ART is recommended. The results shown in this study suggest that with all 648 SREs, couples that present for reproductive counselling have a significant high success rate (77%) using TIC/IUI in comparison when some SREs are absent (27%: two tailed Fisher tests $p=0.012$). When validated and if implemented as part of the clinical assessment, the absence of an SRE may suggest the earlier use of ART that given current practice could reduce the time to achieve LB.

Approximately 40% of the SREs are within exonic regions of genes that are known to be involved in spermatogenesis, sperm motility, fertilization and the first steps of embryogenesis prior to implantation. This corresponds to 262 genes that when impaired would not be bypassed by TIC or IUI, but could likely be remediated using ART. This may reflect sperm no longer having to transit the woman's upper reproductive tract to reach the oocyte and that only those embryos that have shown successful fertilization and initial development are transferred to women using ART. However, some SREs are absent in patients where ART was not successful. Their contribution to the mechanism of successful fertilization and early embryogenesis remains to be elucidated. It is possible that these RNAs are critical for implantation and/or embryogenesis, and thus in these instances even ART cannot lead to a viable pregnancy as is perhaps exemplified in the absence of transcription factor *CAMTA2*.

RNA-seq data also affords the opportunity to SNP genotype each population of sperm RNAs that perhaps can reflect a series of health modifiers (27). For example, within this initial study 102 SREs were derived from 35 genes that have been associated with a spectrum of

genetic disorders from enolase deficiency to Parkinson disease. This is of note given the global allelic imbalance in the gene expression favoring the paternal expression of these genes associated with complex diseases(28) that may be compounded when the paternal effect of diet and environment on the future health of the child is considered (29). Continued development of this sperm RNA-seq methodology is expected to reveal genomic variants from this data(27) that will better inform on the underlying origin of male infertility and possibly predict the future health of the child.

The use of spermatozoal RNA NGS identified a set of molecular biomarkers that shows its potential to predict the success rate of the fertility treatments used. In comparison to a smaller microarray based study where 26 differential mRNAs were identified (18), NGS analysis identified 648 SREs suggesting that RNA-seq technology may more completely resolve variances in RNA profiles for the complex sperm cell.. The statistically significant differences observed in non-invasive treatment outcome based on the presence or absence of the complete set of SREs (two tailed Fisher test $p=0.012$) supports the view that sperm RNA analysis has the potential to effect clinical care for idiopathic infertile couples when used to assess the likelihood of success of TIC/IUI prior to employing ART. This may permit an informed choice towards a treatment paradigm that would help the female partner to avoid undergoing invasive procedures such as egg collection. The results of this 0.7 powered study should encourage a larger, blinded and controlled prospective analysis of patients using non-invasive treatments to ensure its utility prior to its introduction into the fertility clinic. Nevertheless, with the rapidly decreasing cost of NGS, deep sequencing of sperm RNA has the distinct potential to afford clinical benefit while enhancing our understanding of the father's contribution to the birth and life of a healthy child.

Materials and Methods

Experimental design

A retrospective study was designed to investigate whether spermatozoal RNAs could predict the outcome of various fertility treatment options used in the care of idiopathic infertile couples. RNA sequencing was used to obtain the spermatozoal RNA profile of patients included in the study. Several sperm RNA elements required for natural conception were defined and tested using the sperm of idiopathic infertile males of couples that underwent non-invasive or invasive treatments. Sample size was dictated by the availability of patients matching the strict criteria in the fertility clinics powering the study to 0.7 with an α error of 0.029.

Study subjects

Semen samples were collected after IRB approval and informed consent from a total of 96 patients from the CReATe Fertility Centre, Toronto, Canada (Site 1) and the Harvard School of Public Health, Boston, USA (Site 2), then processed and frozen at -80°C . Only couples presenting with unexplained infertility by standard procedures as confirmed by a reproductive endocrinologist and andrologist were recruited to the study. All participants underwent some reproductive treatment ((TIC), (IUI), (IVF), or (ICSI)) due to their inability to achieve a successful live birth by spontaneous conception. Non-invasive treatments (TIC

and IUI) were the first fertility treatment used in approximately 80% of the patients with a success rate of 65% in the two first cycles. After some unsuccessful IUI cycles (2.6 cycles in average) 26% of patients underwent ART treatments achieving 85% success in the first cycle. In contrast, 20% of the patients made the initial personal treatment decision to directly undergo ART with a success rate of 50% in first cycle. Couples were excluded from this study when the male partner considerably deviated from the standard semen parameters. This included males presenting with less than 10 million motile sperm indicative of a negative association of TMC and IUI outcomes (30). In addition, a female partner exhibiting low ovarian reserve validated by Anti-Müllerian Hormone (AMH) level (<15.7 pmol/L) or known history of hormonal disorders, showing evidence of stage 3 or 4 endometriosis, a history of chemotherapy or pelvic irradiation, as well as patients unable or unwilling to consent were excluded from the study. Semen parameters were assessed as described in the Supplemental Methods and this designated day 0 of a 90 day cycle corresponding to one complete spermatogenic cycle. The deidentified frozen samples were processed and analyzed at Wayne State University under the IRB protocols H-06-67-96 and HIC 095701MP2F.

RNA sequencing

Sperm RNA from 96 samples was isolated, quality assessed, deep sequenced, and analyzed as described in the Supplementary Methods. The 72 samples passing all the sequence quality measures were divided into 3 groups for post hoc statistical analysis. Group I samples were used to determine the required sperm RNA elements (SREs) for “natural conception” live birth. This positive control population was derived from seven couples that summarily achieved a live birth by controlling the optimal fertility window (TIC) during the first cycle monitored for intercourse timing. This was considered the same as natural conception. The abundance of all members of this composite group of SREs was assessed within the remaining 65 samples that represented various fertility treatments. Group II test samples were comprised of (i) 29 couples treated by TIC after the first spermatogenic cycle or IUI (ii) 13 couples that were unsuccessful with IUI that subsequently utilized ART and (iii) 14 couples that personally decided to undergo ART after semen assessment. The Group III test samples were comprised of (i) samples from five couples from independent fertility clinic, that achieved a live birth and (ii) four likely female factor couples, three of whom only achieved a live birth with the use of a gestational carrier suggestive of a female factor, and the fourth presenting with stage 2 endometriosis, a known female factor. The results corresponding to the 72 RNA-seq data sets are available at the GEO (NCBI) repository (GSE65683).

Statistical analysis

Statistical analyses were performed using SigmaPlot, version 11 (Systat Software, San Jose, CA) and statistical tests were considered significant at $p < 0.05$. According to the normality of the parameter tested (male and female age, sperm motility, DFI index are normally distributed in contrast to total millions of sperm cells, sperm morphology and number of SREs absent), a parametric one-way analysis of variance (ANOVA) test or non-parametric Kruskal–Wallis one-way analysis of variance by ranks ($\alpha = 0.05$) was used to detect differences in the average of different seminal parameters or SREs that were absent between

the groups based on the treatment used. According to the normality of the parameter tested, a parametric Student's t-test or non-parametric Mann-Whitney U test (two tailed, $\alpha=0.05$) was used to detect differences in the average of different seminal parameters or SREs absent between the samples that achieved live birth (LB) or failed (NLB) within each group. Two tailed Fisher's exact test ($\alpha=0.05$) was used to compare the success rate of different fertility treatments in the presence or absence of the complete set of SREs. Gpower version 3.1 was used to calculate the power of the two tailed Fisher's exact test ($\alpha=0.05$) (31).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: The work was supported by a Collaborative Translational Research Project (CTRP) grant from EMD Serono to SAK and MPD and funds to SAK through Charlotte B. Failing Professorship and NIH grants ES017285 and ES009718 to R.H.

References and Notes

1. Boivin J, Bunting L, Collins JA, Nygren KG. International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. *Human reproduction*. 2007; 22:1506–1512. [PubMed: 17376819]
2. A. Male Infertility Best Practice Policy Committee of the American Urological, M. Practice Committee of the American Society for Reproductive. Report on optimal evaluation of the infertile male. *Fertility and sterility*. 2006; 86:S202–209. [PubMed: 17055823]
3. CDC. 2004 Assisted Reproduction Technologies Success Rates. National Summary and Fertility Clinic Reports. 2006
4. Buck Louis GM, Sundaram R, Schisterman EF, Sweeney A, Lynch CD, Kim S, Maisog JM, Gore-Langton R, Eisenberg ML, Chen Z. Semen quality and time to pregnancy: the Longitudinal Investigation of Fertility and the Environment Study. *Fertility and sterility*. 2014; 101:453–462. [PubMed: 24239161]
5. van den Hoven L, Hendriks JC, Verbeet JG, Westphal JR, Wetzels AM. Status of sperm morphology assessment: an evaluation of methodology and clinical value. *Fertility and sterility*. 2015; 103:53–58. [PubMed: 25450299]
6. Krawetz SA. Paternal contribution: new insights and future challenges. *Nature reviews Genetics*. 2005; 6:633–642.
7. Johnson GD, Lalancette C, Linnemann AK, Leduc F, Boissonneault G, Krawetz SA. The sperm nucleus: chromatin, RNA, and the nuclear matrix. *Reproduction*. 2011; 141:21–36. [PubMed: 20876223]
8. Aston KI, Carrell DT. Genome-wide study of single-nucleotide polymorphisms associated with azoospermia and severe oligozoospermia. *Journal of andrology*. 2009; 30:711–725. [PubMed: 19478329]
9. Tuttelmann F, Simoni M, Kliesch S, Ledig S, Dworniczak B, Wieacker P, Ropke A. Copy number variants in patients with severe oligozoospermia and Sertoli-cell-only syndrome. *PloS one*. 2011; 6:e19426. [PubMed: 21559371]
10. Hammoud SS, Nix DA, Hammoud AO, Gibson M, Cairns BR, Carrell DT. Genome-wide analysis identifies changes in histone retention and epigenetic modifications at developmental and imprinted gene loci in the sperm of infertile men. *Human reproduction (Oxford, England)*. 2011; 26:2558–2569.
11. Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR. Distinctive chromatin in human sperm packages genes for embryo development. *Nature*. 2009; 460:473–478. [PubMed: 19525931]

12. Martinez-Heredia J, de Mateo S, Vidal-Taboada JM, Ballesca JL, Oliva R. Identification of proteomic differences in asthenozoospermic sperm samples. *Human reproduction*. 2008; 23:783–791. [PubMed: 18281682]
13. Jodar M, Kalko S, Castillo J, Ballesca JL, Oliva R. Differential RNAs in the sperm cells of asthenozoospermic patients. *Human reproduction*. 2012; 27:1431–1438. [PubMed: 22353264]
14. Platts AE, Dix DJ, Chemes HE, Thompson KE, Goodrich R, Rockett JC, Rawe VY, Quintana S, Diamond MP, Strader LF, Krawetz SA. Success and failure in human spermatogenesis as revealed by teratozoospermic RNAs. *Human molecular genetics*. 2007; 16:763–773. [PubMed: 17327269]
15. Klaver R, Tuttelmann F, Bleiziffer A, Haaf T, Kliesch S, Gromoll J. DNA methylation in spermatozoa as a prospective marker in andrology. *Andrology*. 2013; 1:731–740. [PubMed: 23970452]
16. Azpiazu R, Amaral A, Castillo J, Estanyol JM, Guimera M, Ballesca JL, Balasch J, Oliva R. High-throughput sperm differential proteomics suggests that epigenetic alterations contribute to failed assisted reproduction. *Human reproduction*. 2014; 29:1225–1237. [PubMed: 24781426]
17. Garcia-Herrero S, Garrido N, Martinez-Conejero JA, Remohi J, Pellicer A, Meseguer M. Differential transcriptomic profile in spermatozoa achieving pregnancy or not via ICSI. *Reproductive biomedicine online*. 2011; 22:25–36. [PubMed: 21123116]
18. Garrido N, Martinez-Conejero JA, Jauregui J, Horcajadas JA, Simon C, Remohi J, Meseguer M. Microarray analysis in sperm from fertile and infertile men without basic sperm analysis abnormalities reveals a significantly different transcriptome. *Fertility and sterility*. 2009; 91:1307–1310. [PubMed: 18367176]
19. Camprubi C, Pladevall M, Grossmann M, Garrido N, Pons MC, Blanco J. Semen samples showing an increased rate of spermatozoa with imprinting errors have a negligible effect in the outcome of assisted reproduction techniques. *Epigenetics*. 2012; 7:1115–1124. [PubMed: 22885410]
20. Jodar M, Selvaraju S, Sendler E, Diamond MP, Krawetz SA. N. Reproductive Medicine, The presence, role and clinical use of spermatozoal RNAs. *Human reproduction update*. 2013; 19:604–624. [PubMed: 23856356]
21. Krawetz SA, Kruger A, Lalancette C, Tagett R, Anton E, Draghici S, Diamond MP. A survey of small RNAs in human sperm. *Human reproduction*. 2011; 26:3401–3412. [PubMed: 21989093]
22. Sendler E, Johnson GD, Mao S, Goodrich RJ, Diamond MP, Hauser R, Krawetz SA. Stability, delivery and functions of human sperm RNAs at fertilization. *Nucleic acids research*. 2013; 41:4104–4117. [PubMed: 23471003]
23. Ostermeier GC, Miller D, Huntriss JD, Diamond MP, Krawetz SA. Reproductive biology: delivering spermatozoan RNA to the oocyte. *Nature*. 2004; 429:154. [PubMed: 15141202]
24. Nardelli AA, Stafinski T, Motan T, Klein K, Menon D. Assisted reproductive technologies (ARTs): evaluation of evidence to support public policy development. *Reprod Health*. 2014; 11:76. [PubMed: 25376649]
25. M. Practice Committee of the American Society for Reproductive. Effectiveness and treatment for unexplained infertility. *Fertility and sterility*. 2006; 86:S111–114. [PubMed: 17055802]
26. Malcolm CE, Cumming DC. Follow-up of infertile couples who dropped out of a specialist fertility clinic. *Fertility and sterility*. 2004; 81:269–270. [PubMed: 14967354]
27. Piskol R, Ramaswami G, Li JB. Reliable identification of genomic variants from RNA-seq data. *American journal of human genetics*. 2013; 93:641–651. [PubMed: 24075185]
28. Crowley JJ, Zhabotynsky V, Sun W, Huang S, Pakatci IK, Kim Y, Wang JR, Morgan AP, Calaway JD, Aylor DL, Yun Z, Bell TA, Buus RJ, Calaway ME, Didion JP, Gooch TJ, Hansen SD, Robinson NN, Shaw GD, Spence JS, Quackenbush CR, Barrick CJ, Nonneman RJ, Kim K, Xenakis J, Xie Y, Valdar W, Lenarcic AB, Wang W, Welsh CE, Fu CP, Zhang Z, Holt J, Guo Z, Threadgill DW, Tarantino LM, Miller DR, Zou F, McMillan L, Sullivan PF, Pardo-Manuel de Villena F. Analyses of allele-specific gene expression in highly divergent mouse crosses identifies pervasive allelic imbalance. *Nature genetics*. 2015
29. Soubry A, Hoyo C, Jirtle RL, Murphy SK. A paternal environmental legacy: evidence for epigenetic inheritance through the male germ line. *Bioessays*. 2014; 36:359–371. [PubMed: 24431278]

30. Ombelet W, Dhont N, Thijssen A, Bosmans E, Kruger T. Semen quality and prediction of IUI success in male subfertility: a systematic review. *Reproductive biomedicine online*. 2014; 28:300–309. [PubMed: 24456701]
31. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*. 2007; 39:175–191. [PubMed: 17695343]
32. WHO. WHO Laboratory Manual Examination of Human Semen and Sperm-Cervical Mucus Interaction (4th). 1999
33. Wolff H. The biologic significance of white blood cells in semen. *Fertility and sterility*. 1995; 63:1143–1157. [PubMed: 7750580]
34. WHO. WHO Laboratory Manual for Examination and Processing of Human Semen. 5th. World Health Organization; Switzerland: 2010.
35. Kruger TF, Ackerman SB, Simmons KF, Swanson RJ, Brugo SS, Acosta AA. A quick, reliable staining technique for human sperm morphology. *Archives of andrology*. 1987; 18:275–277. [PubMed: 2445315]
36. Evenson DP, Larson KL, Jost LK. Sperm chromatin structure assay: its clinical use for detecting sperm DNA fragmentation in male infertility and comparisons with other techniques. *Journal of andrology*. 2002; 23:25–43. [PubMed: 11780920]
37. Moskovtsev SI, Willis J, White J, Mullen JB. Sperm DNA damage: correlation to severity of semen abnormalities. *Urology*. 2009; 74:789–793. [PubMed: 19643462]
38. Mao S, Goodrich RJ, Hauser R, Schrader SM, Chen Z, Krawetz SA. Evaluation of the effectiveness of semen storage and sperm purification methods for spermatozoa transcript profiling. *Systems biology in reproductive medicine*. 2013
39. Goodrich R, Johnson G, Krawetz SA. The preparation of human spermatozoal RNA for clinical analysis. *Archives of andrology*. 2007; 53:161–167. [PubMed: 17612875]
40. Mao S, Sendler E, Goodrich RJ, Hauser R, Krawetz SA. A comparison of sperm RNA-seq methods. *Systems biology in reproductive medicine*. 2014; 60:308–315. [PubMed: 25077492]
41. Lambard S, Galeraud-Denis I, Martin G, Levy R, Chocat A, Carreau S. Analysis and significance of mRNA in human ejaculated sperm from normozoospermic donors: relationship to sperm motility and capacitation. *Molecular human reproduction*. 2004; 10:535–541. [PubMed: 15100385]
42. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology*. 2009; 10:R25. [PubMed: 19261174]
43. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 2009; 25:1105–1111. [PubMed: 19289445]

Figure 1. Genomic localization (exon, intron, intergenic, non-coding RNAs) and function of the 648 required sperm RNA elements (SREs)

The majority of the required SREs are located in exons of annotated genes (585 of 648; 90.3%), while the remaining are in intronic regions (12 of 648; 1.9 %), intergenic regions (9 of 648; 1.4%) or match to non-coding transcriptional elements including small nuclear RNAs, miRNAs and long non-coding RNAs (42 of 648; 6.5%) with potential regulatory function. Approximately 40% of the genes which contain one or more SREs have a known role in spermatogenesis, sperm physiology (sperm energy production or acrosome reaction), fertilization, and/or early embryogenesis. Additionally, 20% have a known role in cellular process such as in transcription regulation, protein transport, Ubiquitin-like conjugation pathway and lipid metabolism. The potential of the remaining transcripts have yet to be defined.

Figure 2. Distribution of the 648 required sperm RNA elements (SREs)

The percentile rank distribution of the SREs in control (Group I) and test set (Group II: (i) exclusively use TIC or IUI, (ii) unsuccessful IUI followed by ART, (iii) directly use ART) and outcome prediction is presented. **A.** Left panel, 7 TIC individuals used to determine the SREs considered as natural conception presenting all 648 SREs; middle panel 37 samples from Group II (25 Group II-i, 5 Group II-ii and 7 Group II-iii) with the complete set of SREs; right panel, 19 samples From Group II (4 Group II-i, 8 Group II-ii and 7 Group II-iii) with at least one SRE is absent. The percentile rank of each element is indicated (green, >90th; yellow, 75th < yellow < 85th; and red, "outlier" elements 0th). The first 33 sperm elements are at least absent in one sample from the test set (right panel). The remaining SREs (34- 648) were present in all samples from groups I and II, showing that the vast majority of SRE's are uniformly abundant in all samples surveyed. **B.** Using TIC/IUI or ART couples with all SREs have a success rate of 72% during the first two treatment cycles (6 months) after sperm RNA evaluation. LB: Live Birth; NLB: No Live Birth; TIC: Timed Intercourse; IUI: Intrauterine Insemination; IVF: *In Vitro* Fertilization; ICSI: Intracytoplasmic Sperm Injection; SRE: Sperm RNA Element.

Figure 3. Analysis of treatment outcome as function of required sperm RNA elements (SREs)

A. The majority of the 37 group II couples with all SREs present underwent TIC/UI, achieving a LB rate of 77%. The remaining samples reflect patient preference along with the previously unsuccessful TIC/UI cases were treated by ART achieving a 75% LB rate. **B.** RNA analysis from samples with at least one SRE absent. Note that only 3 of the 11 TIC/UI samples with a SRE absent achieved LB. The success ratio of live birth using ART is similar to the ratio observed in samples with all SREs. **C.** The percentage of LB using a non-invasive treatment for couples presenting with the complete set of SREs is higher compared to those with some SRE absent (Fisher's exact test, two-tailed P-value= 0.012) LB: Live Birth; NLB: No Live Birth; IUI: Intrauterine Insemination; TIC: Timed Intercourse; ART: Assisted Reproductive Technologies. SRE: Sperm RNA Elements.

Characteristics of the Study Population

Table 1

The distribution and characteristics of the study population in relation to fertility treatment employed and procedural outcome - Live Birth (LB) vs. No Live Birth (NLB) is detailed. Group I achieved a live birth pregnancy in their first attempt using Timed Intercourse (TIC) during first spermatogenic cycle after semen assessment (90 day cycle), and were considered as a natural conception. Samples from test set (II) include different subgroups based on treatment: (i) Intrauterine Insemination (IUI) or TIC delayed past the first 90 day cycle; (ii) ART preceded by unsuccessful IUI or TIC (iii) ART. The independent set of samples (III) was composed from 2 subgroups; (i) samples from an independent fertility clinic; (ii) patients that never achieved LB, subsequent to which a female factor was diagnosed. Sample characteristics include male age, female age, and semen parameters comprising Total Million Sperm Cell per sample, Sperm motility (%), Sperm Morphology (% of Normal forms (NF)), and Sperm DNA fragmentation (DNA Fragmentation Index (DFI)). A. No correlation of the type of fertility treatment employed as compared to any individual or sperm sample parameter evaluated was observed. B. When considered as a group, female age showed a negative correlation with LB (two tailed t-test, $p=0.024$) C. Female age was significantly higher in patients that were unsuccessful when treated by ART (two tailed t-test, $p=0.004$) but not in patients that were unsuccessful when treated by TIC/IUI. .

A						
Group	Natural conception		Test set		Independent set	
	I		II		III	
Subgroup		I	ii	iii	i	ii
Male age	35	34.6	33.6	34.2	37	36.2
Female age	32.3	32	31.1	33.6	35.2	32
Total million sperm	193.4	172.9	235.8	159	349.9	160.2
Sperm motility (%)	54	52.3	50	50	39.5	52.5
TMC	113.9	93.9	125.5	83.8	175.5	86.7
Sperm morphology (%NF)	10.8	10.6	5.7	9.2	4.5	10.2
DNA fragmentation (DFI)	14.7	16.4	17.3	19.4	-	18.4

P-value

0.683
0.304
0.090
0.441
0.342
0.271
0.734

Author Manuscript

B	Final outcome			P-value
	LB	NLB		
N	62	10		
Male age	34.6	35		0.770
Female age	32	34.9		0.024*
Total million sperm	192.4	212.4		0.122
Sperm motility (%)	50.4	52.8		0.583
TMC	103	116.7		0.182
Sperm morphology (%NF)	8.8	10.9		0.342
DNA fragmentation (DFI)	17.5	14.6		0.303
C	TIC/IUI			ART
	LB	NLB		
N	35	17		
	P-value	27	6	
Female age	32.3	31.5		0.377
				31.5
				37
				0.004*

Table 2
Required Sperm RNA Elements (SREs) in the test group of samples (Group III)

The 648 RNA elements describing the fertile sperm by natural conception were tested in 9 samples. (i) All samples were obtained from an independent fertility clinic and achieved LB presenting all required sperm RNA elements. LB was achieved spontaneously or by IUI (samples 4,5) while the remaining three cases (1-3) directly employed ART. A single instance (sample 6) with known female factor shows the absence of 2 SREs. It is possible that with a gestational carrier pregnancy was rescued using ART despite the absence of these 2 SREs. LB: Live Birth; NLB: No Live Birth; TIC: Timed Intercourse; IUI: Intrauterine Insemination; IVF: *In Vitro* Fertilization; ICSI: Intracytoplasmic Sperm Injection; GC: Gestational Carrier.

Treatment			Final outcome	Required SREs absent
i- Samples from independent fertility clinic				
1	IVF (LB)		LB	0
2	IVF (NLB)	ICSI (LB)	LB	0
3	IVF (NLB)	ICSI (LB)	LB	0
4	IUI (LB)		LB	0
5	Natural conception		LB	0
ii- Samples with known female factor				
6	ICSI (LB with GC)		LB	2
7	ICSI (NLB)	ICSI (LB with GC)	LB	0
8	ICSI (NLB)	ICSI (LB with GC)	LB	0
9	TIC (NLB)	IUI (NLB)	NLB	0