Asian Journal of Andrology (2019) 21, 121–130
www.asiaandro.com; www.ajandrology.com

Open Access

ORIGINAL ARTICLE

Sperm Biology

Proteomic analysis reveals dysregulated cell signaling in ejaculated spermatozoa from infertile men

Luna Samanta[1,2], Rakesh Sharma[1], Zhihong Cui[1,3], Ashok Agarwal[1]

Dysfunctional sperm maturation is the primary reason for the poor sperm motility and morphology in infertile men. Spermatozoa from infertile men were fractioned on three-layer density gradient (80%, 60%, and 40%). Fraction 1 (F1) refers to the least mature stage having the lowest density, whereas the fraction 4 (F4) includes the most dense and morphologically mature motile spermatozoa. Fraction 2 (F2) and fraction 3 (F3) represent the intermediate stages. Proteins were extracted and separated by 1-dimensional gel. Bands were digested with trypsin and analyzed on a LTQ-Orbitrap Elite hybrid mass spectrometer system. Functional annotations of proteins were obtained using bioinformatics tools and pathway databases. A total of 1585 proteins were detected in the four fractions of spermatozoa. A dysregulated protein turnover and protein folding may lead to accumulation of defective proteins or proteins that otherwise would have been eliminated during the process of maturation, resulting in the impairment of sperm function. Aberrant chaperone expression may be a major contributing factor to the defective sperm function. Androgen receptor was predicted as a transcription regulator in one of the networks and the affected pathways were chaperone-mediated stress response, proteosomal pathway, and sperm function. The downregulation of key pathways and proteins which compromises the fertilizing potential of spermatozoa may provide insight into the mechanisms that lead to male infertility.

Asian Journal of Andrology (2019) **21**, 121–130 doi: 10.4103/aja.aja_56_18; published online: 26 October 2018

Keywords: androgen receptor; chaperone; immature sperm; infertile men; proteasome; spermatozoa

INTRODUCTION

Male infertility is a multifactorial condition and there is no identifiable cause in 50% of the cases. The human testis produces spermatozoa at a rate of 1000 cells per second and these cells are highly differentiated and unique.[1] Spermatozoa originate from the complex process of spermatogenesis in three major steps as follows: (1) proliferation and differentiation of spermatogonia; (2) divisions during the spermatocyte stage; and (3) spermiogenesis. Spermiogenesis involves major morphological and molecular changes, including the removal of cytoplasm, formation of the acrosome and flagella, mitochondrial rearrangement, and nuclear remodeling. During mid-spermiogenesis, the nucleus of the round spermatid changes from spherical to a unique elongated and flattened shape. This reshaping protects the male genome during sperm transport and also facilitates the penetration of spermatozoa into ovum. Thus, spermatozoa are terminally differentiated and possess specialized organelles.[2] However, they undergo maturation during epididymal transit to acquire the ability to fertilize.[3]

Spermatozoa have different pathologies from those of somatic cells, which result in different sperm phenotypes in the ejaculated semen. Seven sperm phenotypes have been detected in human semen from electron microscopy, which include spermatozoa with dysplasia of the fibrous sheath, nonspecific flagellar defects, immotile cilia, acrosomal hypoplasia, defective chromatin condensation and compaction, pin head, and even sperm cells without heads.[4] These conditions cannot be identified by routine semen analysis or functional tests since the deficiencies demonstrated by these methods are secondary manifestations of an underlying pathology.

The generation of high-quality spermatozoa is governed by a number of selective mechanisms within the testes and epididymis.[5] The dramatic changes that occur during spermiogenesis, sperm maturation, and capacitation involve loss and gain of specific proteins.[6–8] Recently, we have reported that distinct proteomic signatures distinguish high-quality spermatozoa from their low-quality counterparts in fertile donors.[9] Improper spermatogenesis produces abnormal spermatozoa that are generally earmarked for elimination by apoptosis and appear in the ejaculate when they escape apoptosis. Furthermore, differential localization of Fas, a membrane receptor of the tumor necrosis factor family that initiates apoptosis, also segregates the spermatozoa into different subsets.[10,11] Thus, spermatozoa marked for apoptosis are of lower reproductive potential than their unmarked counterparts. Defects in epididymal maturation lead to increased morphological abnormalities in the spermatozoa and poor sperm motility.[2,12,13] In addition, immature spermatozoa exhibit metabolic alterations, presence of excess cytoplasm in the ejaculate, increased production of reactive oxygen species (ROS), lipid peroxidation, and DNA fragmentation.[5,14]

[1]American Center for Reproductive Medicine, Cleveland Clinic, Cleveland, OH 44195, USA; [2]Department of Zoology, School of Life Sciences, Ravenshaw University, Cuttack, Odisha 753003, India; [3]Institute of Toxicology, The Third Military Medical University, Chongqing 400038, China.
Correspondence: Dr. Ashok Agarwal (agarwaa@ccf.org)
Received: 20 December 2017; Accepted: 13 June 2018

Sperm preparation methods such as density gradient separation are routinely used to obtain highly motile and morphologically normal sperm for assisted reproductive technology (ART). Spermatozoa separated on a three-layer density gradient (40%, 60%, and 80%) demonstrate cell-to-cell variation in both fertile and infertile men.[5] Spermatozoa were separated into four fractions on the basis of their density and maturity. The lowest level of ROS production and DNA damage correlates with morphologically normal, motile spermatozoa obtained in the mature subset from fertile men compared with abnormal spermatozoa from infertile men.[14–16] Different subsets of spermatozoa obtained from the ejaculate of fertile men after separation on a three-layer density gradient differ in their proteome profile.[9] We have demonstrated an increasing trend in proteins involved in key biological processes during sperm maturation such as reproductive cellular process, gamete production, motility, oxidative phosphorylation, and energy metabolism. A decreasing trend was seen in the expression of proteins that were involved in key biological processes such as protein synthesis, protein transport, and response to oxidative stress.[9]

Division of spermatozoa into phenotypes or subsets is of importance in the evaluation of its true fertility potential, particularly when using testicular spermatozoa for intracytoplasmic sperm injection (ICSI). A recent review by Esteves *et al*.[17] documented that infertile couples may benefit from ICSI with testicular spermatozoa instead of ejaculated spermatozoa if the male partners exhibit high sperm DNA fragmentation (SDF) in the ejaculate. We have also demonstrated that proteins critical for sperm maturation, motility, and fertilization are involved in biological processes that are activated or suppressed in different subsets of spermatozoa from fertile men.[9] However, the underlying pathways and the distribution of proteins in immature and mature sperm from infertile men have not been explored utilizing a proteomic approach. The present study is a continuation of our previous report on fertile donors deciphering the proteomic signatures in the spermatozoa of infertile patients to understand the underlying mechanism(s) of defective sperm maturation.

PATIENTS AND METHODS

Patients

Following the approval of the study by the Institutional Review Board (IRB) of Cleveland Clinic (Cleveland, OH, USA), semen samples were collected from 11 infertile men. Men with leukocytospermia (Endtz test positive) and female factor infertility were excluded. All enrolled patients provided written consent to participate in the study.

Sample collection and semen analysis

Semen samples were examined according to 2010 World Health Organization (WHO) criteria.[18] All specimens were collected by masturbation after sexual abstinence of 48–72 h and were allowed to liquefy for 20 min at 37°C before further processing. Following liquefaction, manual semen analysis was performed, including evaluation of presence of round cells, presence of white blood cells by the peroxidase test (Endtz test), viability, and morphology as previously described.[9]

Separation of sperm phenotypes and proteomic analysis

For separating immature and mature spermatozoa, a three-layer density gradient was used as described earlier.[9] It consisted of 2 ml of 40%, 60%, and 80% of the upper layer, intermediate layer, and lower layer, respectively. The three layers were reconstituted from the stock (100%) solution of the gradient with the SpermRinse medium (Vitrolife, San Diego, CA, USA). The stock gradient was an antibiotic-free bicarbonate and HEPES-buffered medium containing silane-coated, colloid silica particles. The SpermRinse medium was a bicarbonate and HEPES-buffered medium containing human serum albumin and gentamycin as an antibiotic (Vitrolife, San Diego, CA, USA). The three-layer gradient is a slight modification of the 2-layer density gradient method routinely used for preparing sperm for ART techniques, especially intrauterine insemination.[9] Briefly, 1–2 ml of liquefied semen sample was carefully loaded on the 40% gradient and centrifuged at 300 *g* for 20 min. The resulting interfaces between the seminal plasma and 40% (fraction 1), 40% and 60% (fraction 2), 60% and 80% (fraction 3), and the 80% pellet (fraction 4, mature fraction) were carefully aspirated, resuspended in human tubal fluid media (HTF, Irvine Scientific, Santa Ana, CA, USA), and centrifuged at 300 g for 7 min. The pellets of each fraction were resuspended in 0.5–1 ml HTF, and the total sperm count, motility, and morphology were assessed again.

Sperm proteins from two individual samples and one pooled (from four individuals after normalization for spermatozoa number and protein content) sample were dissolved in RIPA lysis buffer (Sigma-Aldrich, St. Louis, MO, USA) supplemented with proteinase inhibitor cocktail (Roche, Indianapolis, IN, USA). Protein concentration was determined by the Bicinchoninic acid (BCA) kit (Thermo, Rockford, IL, USA). Proteins were extracted and separated by 1-dimensional gel electrophoresis. Bands were digested with trypsin and analyzed on a LTQ-Orbitrap Elite hybrid mass spectrometer (Thermo) system as described.[9] Each sample was run in triplicate and the average was taken.

Database searching and protein identification

Tandem mass spectra were extracted by Proteome Discoverer (version 1.4.1.288, Thermo Fisher Scientific, San Jose, CA, USA). Twelve tandem mass spectrometry or MS/MS samples (3 runs per sample) were analyzed by using Mascot (version 2.3.02, Matrix Science, London, UK), Sequest (version 1.4.0.288, Thermo Fisher Scientific), and X! Tandem (TheGPM, thegpm.org; version CYCLONE 2010.12.01.1). Mascot, Sequest, and X! Tandem were set up to search the human database (33 292 entries) assuming the digestion enzyme trypsin. To validate MS/MS-based peptide and protein identifications, Scaffold (version Scaffold 4.0.6.1, Proteome Software Inc., Portland, OR, USA) was used. Peptide identifications were accepted if they could be established at >95.0% probability by the PeptideProphet™ algorithm with Scaffold delta-mass correction.[19]

Quantitation of the relative abundance of protein in spermatozoa

The relative abundance of sperm proteins was determined by comparing the number of spectra, termed spectral counts (SC), used to identify each protein. The numerical values used in the quantitation correspond to the normalized spectral abundance factor (NSAF, SC/[ΣSC] × protein length). NSAF approach was applied before relative protein quantification.[20] Differentially expressed proteins (DEPs) were obtained by applying different constraints for significance tests or fold change cutoffs from the average SC of the protein from multiple runs.[9]

The categorization of overall abundance along with the filtering criteria used for differential expression analysis is summarized below:

1. Very low abundance: spectral count range 1.7–7; $P \leq 0.001$; and NSAF ratio ≥2.5 for overexpressed and ≤0.4 for underexpressed proteins
2. Low abundance: spectral count range 8–19; $P \leq 0.01$; and NSAF ratio ≥2.5 for overexpressed and ≤0.4 for underexpressed proteins
3. Medium abundance: spectral count range 20–79; $P \leq 0.05$; and NSAF ratio ≥2.0 for overexpressed and ≤0.5 for underexpressed proteins

4. High abundance: spectral counts >80; $P \leq 0.05$; and NSAF ratio ≥1.5 for overexpressed and ≤0.67 for underexpressed proteins.

Bioinformatic analyses

Functional bioinformatics analyses were done with publicly available tools such as Gene Ontology (GO) annotations from GO Term Finder (http://search.cpan.org/dist/GO-TermFinder/),[21] GO Term Mapper (http://go.princeton.edu/cgi-bin/GOTermMapper), UNIPROT (The UniProt Consortium; http://www.uniprot.org/), Database for Annotation, Visualization and Integrated Discovery (DAVID; http://david.niaid.nih.gov), and proprietary software packages such as Ingenuity Pathway Analysis (IPA from Ingenuity® Systems; https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/) and MetaCore™ (GeneGo Inc., Encinitas, CA, USA) to identify the cellular distribution of proteins and differentially affected processes and pathways.

Statistical analyses

The results were expressed as mean ± standard deviation (s.d.). To compare the differences between different fractions of the ejaculate, we used Jonckheere–Terpstra test (or Jonckheere's trend test). It is similar to the Kruskal–Wallis test where the null hypothesis is that several independent samples are from the same population. However, there is *no priori* ordering of the populations from which the samples are drawn. When there is *a priori* ordering, the Jonckheere–Terpstra test has more power than the Kruskal–Wallis test. In this test, there is no issue of normality and does not require log transformation of the data. The statistical analysis was performed using the MedCalc (version 17.9.7, MedCalc Software bvba, Ostend, Belgium). For IPA and MetaCore™, the right-tailed Fisher's exact test was used. Differences were considered statistically significant for $P < 0.05$.

RESULTS

Semen analysis

Total sperm count (TSC), presence of round cells, motility, and total motile sperm count (TMS) before and after density gradient centrifugation are shown in **Table 1** and **Figure 1**. The TSC ($\times 10^6$, mean ± s.d.) recovered in F3 and F4 was similar (32.51 ± 25.78 and 34.32 ± 32.53, respectively) and significantly different ($P = 0.004665$) from the TSC in the prewash sample (79.98 ± 84.58) (**Figure 1a**). There was a significant decrease ($P < 0.000001$) in the number of round cells (median, 25th and 75th percentiles) from the prewash sample (1.90 [1.07, 3.75]) to that of F3 (0 [0, 0]) and F4 (0 [0, 0]) (**Figure 1b**). Spermatozoa recovered from F4 displayed the highest average motility (63.0 [53.5, 83.1]) compared with F1, F2, and F3 followed by F3 (36.6 [32.3, 43.2]) compared with F1, F2, and F4 ($P = 0.000001$) (**Figure 1c**). A higher recovery of TMS ($\times 10^6$ spermatozoa) was observed in F4 (23.46 ± 23.39) than F3 (12.68 ± 11.90) (**Figure 1d**).

Overall protein abundance

The relative abundance of the identified proteins in the sperm samples was quantified. For higher accuracy of proteomic results, the protein abundance should be similar between the analyzed samples. The range of total SC was 31188–51131 which showed that the abundance of proteins in these samples was similar.

Distribution of proteins in different fractions

All fractions, *i.e.* F1, F2, and F3 were compared with mature fraction F4. A total of 1585 proteins were identified in the four fractions together. Among those, 1202, 1140, 1129, and 890 proteins were detected in F1, F2, F3, and F4, respectively. By comparing the F1 and F4 proteomes, 136 proteins were overexpressed in F1, 177 were underexpressed in F1, 158 were unique to F1, and 51 were unique to F4. When F2 was compared with F4, 113 proteins were overexpressed in F2, 111 were underexpressed in F2, 114 were unique to F2, and 24 were unique to F4. Comparison of F3 versus F4 showed 89 proteins overexpressed in F3, 53 underexpressed in F3, 38 unique to F3, and 8 unique to F4. All the four fractions were compared to identify the differentially expressed proteins (DEPs). Of the 656 DEPs, 75 proteins showed an increasing trend from fraction F1 to F4 (**Supplementary Table 1**), while 279 showed a decreasing trend (**Supplementary Table 2**).

Identification of functional proteins

The Reactome pathway analysis showed that the majority of DEPs were involved in metabolism, particularly protein metabolism, disease processes, gene expression, and signal transduction (**Figure 2a**). The percentage of proteins involved in various pathways that were either underexpressed or showed a decreasing trend or were overexpressed and showed an increasing trend from F1 to F4 is shown in **Figure 2b**.

The functional annotations from use of the DAVID tool and their enrichment analysis with the number of key proteins involved are presented in **Supplementary Table 3**. F1 was characterized by proteins involved in translation, elongation, protein transport, oxidoreductase activity, reproductive processes, spermatid development/differentiation, and regulation of apoptotic pathways. The proteins identified in F2 are involved in cell differentiation, protein–protein interactions, protein transport and localization, oxidoreductase activity, spermatogenesis,

Table 1: Semen parameters before and after separation on a 3-layer density gradient

Parameter	*PW*	*F1*	*F2*	*F3*	*F4*	*P*
Motility (%)	45.40±18.84	7.71±6.82	21.21±7.35	35.78±13.48	64.46±18.12	0.000001[a]
	47.0 (29, 61)	6.6 (3.5, 7.8)	21.4 (17.1, 24.4)	36.6 (32.3, 43.2)	63.0 (53.5, 83.1)	
	F1, F2, F4	F2, F3, F4, (PW)	F1, F3, F4, (PW)	F1, F2, F4	F1, F2, F3	<0.00001[b]
Total sperm count ($\times 10^6$)	79.98±84.58	12.96±18.35	20.14±17.70	34.32±32.53	32.51±25.78	0.004665[a]
	61.80 (30.20, 81.60)	8.85 (1.78, 15.40)	17.00 (7.75, 22.80)	26.50 (13.30, 35.20)	27.45 (11.60, 52.20)	
	F1, F2	F3, F4, (PW)	(PW)	F1	F1	0.00007[b]
Total motile sperm ($\times 10^6$)	102.87±92.06	0.79±1.17	3.98±3.75	12.68±11.90	23.46±23.39	0.000002[a]
	65.93 (33.65, 154.87)	0.199 (0.16, 1.00)	2.79 (1.59, 4.20)	10.79 (4.00, 18.18)	10.90 (7.54, 35.96)	
	F1, F2, F3, F4	F2, F3, F4, (PW)	F1, F3, F4, (PW)	F1, F2, (PW)	F1, F2, (PW)	<0.00001[b]
Round cells ($\times 10^6$ ml^{-1})	1.90 (1.07, 3.75)	1.35 (0.60, 3.20)	0.30 (0.10, 0.70)	0 (0, 0)	0 (0, 0)	<0.000001[a]
	F2, F3, F4	F2, F3, F4	F1, F3, F4	F1, F2, (PW)	F1, F2, (PW)	0.28496[b]

Values are mean±s.d. and median (25th and 75th percentiles). [a]Comparison between PW and different fractions was done by Kruskal–Wallis test; [b]post hoc analysis for differences was done by Jonckheere–Terpstra trend test. $P<0.05$ was considered statistically significant. F1: least mature stage having the lowest density; F2, F3: intermediate stages; F4: includes the most dense and morphologically mature motile spermatozoa. s.d.: standard deviation; PW: prewash

gonad development, and proteolytic pathways. The key processes of F3 proteins include generation of precursor metabolites and integration of energy metabolism, oxidative phosphorylation, protein catabolic process, and protein ubiquitination. Finally, the proteins identified in F4 were largely involved in reproductive functions (**Table 2**). The proteins involved in intracellular transport, oxidation–reduction, cellular amino acid catabolic process, and alternative splicing showed an increasing trend from F1 to F4 (**Supplementary Table 2**). On the other hand, the proteins involved in spermatogenesis, protein metabolism, cell cycle, integration of energy metabolism, regulation of apoptosis, cell redox homeostasis, translational elongation, and response to protein folding showed a decreasing trend from F1 to F4 (**Supplementary Table 3**).

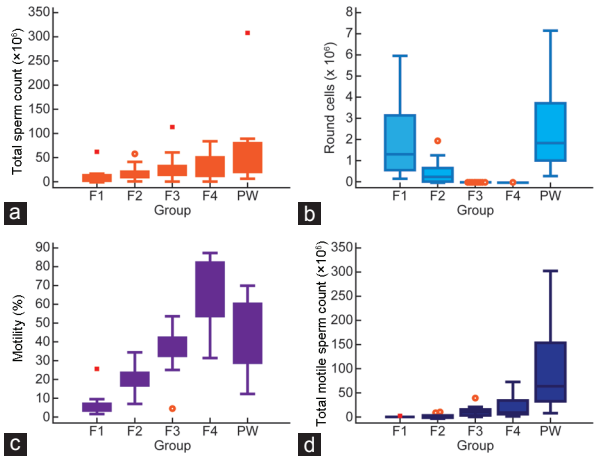

Figure 1: Semen parameters in four fractions compared to PW sample. (**a**) Total sperm count (×10⁶); (**b**) round cell count (×10⁶); (**c**) motility (%); and (**d**) total motile count (×10⁶) in F1–F4. PW: prewash; F1: least mature stage having the lowest density; F2, F3: intermediate stages; F4: includes the most dense and morphologically mature motile spermatozoa.

DEPs involved in various networks

The DEPs were further subjected to network analysis by using IPA. A total of 161 pathways were linked to 279 proteins showing a decreasing trend from F1 to F4. Among these proteins, 58 were involved in cell death and survival, posttranslational modification, and protein folding (**Figure 2c**). In addition, there were 58 proteins, mostly molecular chaperones, which were involved in developmental disorder, posttranslational modification, and protein folding (**Figure 2d**). A key network was identified with 58 focal proteins involved in molecular transport, protein trafficking, and cell cycle (**Figure 2e**), in which 18 proteins were involved in cell signaling, cancer, and cellular development processes. The proteins showing an increasing trend from immature to

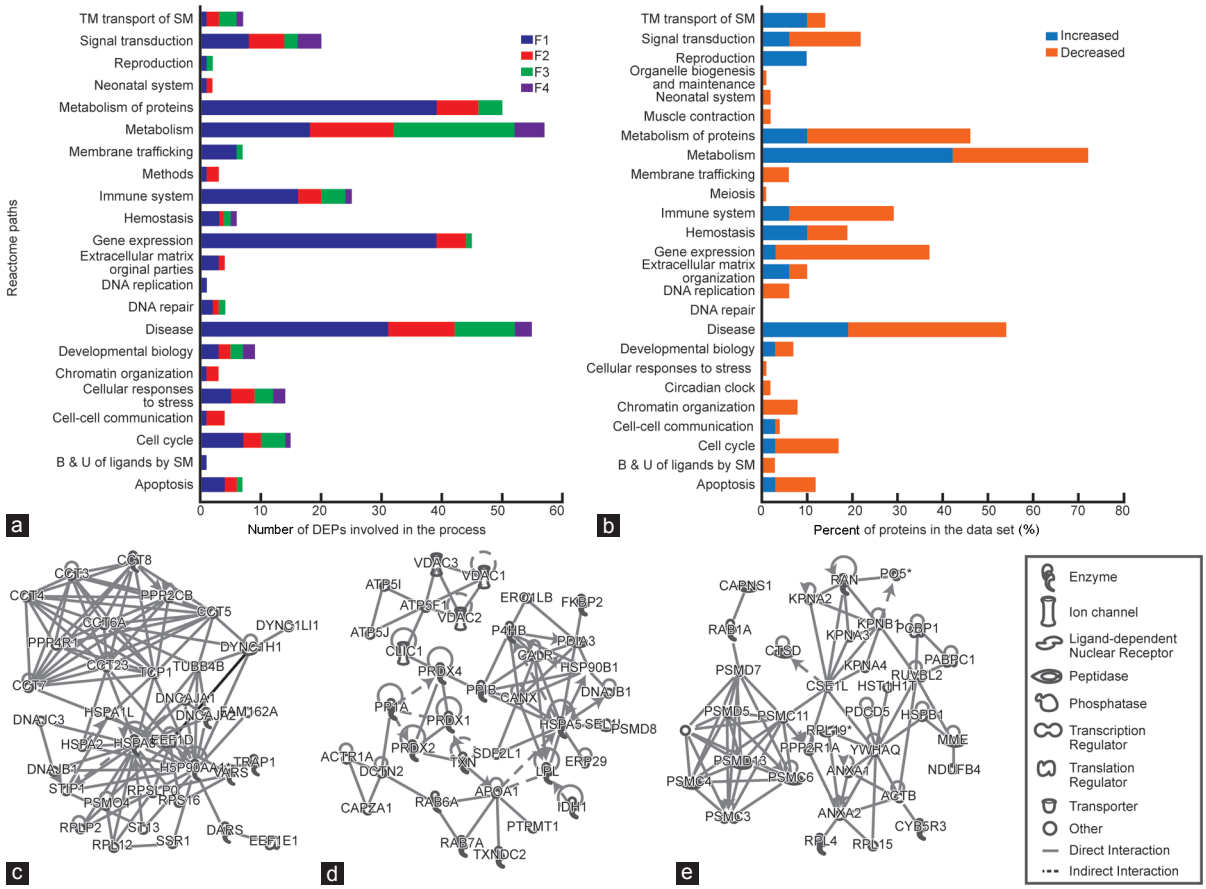

Figure 2: DEPs in fraction 1 (F1) through fraction 4 (F4). (**a**) Reactome pathway showing the number of DEPs involved in various pathways in the four fractions. F1 being the most immature and F4 the most mature. (**b**) Percentage of proteins in the data set that were increasing or decreasing for the different pathways. (**c**) IPA network showing molecular chaperons that were involved in developmental disorder, posttranslational modifications, and protein folding. (**d**) IPA network showing proteins involved in molecular transport, protein trafficking, and cell cycle and (**e**) IPA network showing molecular chaperons involved in cell death and survival, posttranslational modifications, and protein folding. IPA: ingenuity pathway analysis; DEPs: differentially expressed proteins; F1: least mature stage having the lowest density; F2, F3: intermediate stages; F4: includes the most dense and morphologically mature motile spermatozoa. Full names of abbreviated proteins are presented in **Supplementary Information**.

mature fraction (F1 to F4) were linked to 66 networks. The top networks were lipid metabolism, small molecule biochemistry, and molecular transport that included 25 proteins. Of these, 22 proteins participated in cell-to-cell signaling and interaction, cellular function and maintenance, and inflammatory response and 12 proteins were involved in cell death and survival, cellular compromise, and cancer (**Table 2**).

Proteins involved in top pathways and upstream regulators

On the basis of the common genes, 15 top pathways showed an overlap with each other and their translated protein product was shown to be gradually underexpressed from immature to mature fractions, while 8 connecting signaling pathways were predicted (**Figure 3**). The proteins identified in our data set were analyzed using MetaCore™ to predict the upstream regulator(s) and the pathways involved (**Figure 4a–4c**). The androgen receptor was predicted as one of the transcription regulators and the pathways affected were chaperone-mediated stress response, proteosomal pathway, and sperm function (**Figure 4d**).

DISCUSSION

We have previously reported the proteome profile changes in these four subsets of immature and mature spermatozoa from fertile men separated by three layer density gradient. Our results demonstrated that proteins critical for sperm maturation are altered in stage specific manner in fertile men. Of these, 4 proteins, namely, heat shock protein HSP701A, clusterin, tektin 2, and tektin 3 were validated by western blot to corroborate the proteomic findings.[9] In this study, the goal was to compare the same four subsets of immature and mature spermatozoa in infertile men to understand why spermatozoa obtained in the mature fraction (F4), which separate the good quality spermatozoa in terms of motility, TMC, and even TSC, are still dysfunctional in achieving pregnancy. Therefore, in this study, a high-throughput shotgun proteomic approach followed by pathway analysis was used to understand the impaired molecular mechanism involved in sperm dysfunction.

Despite the absence of effective transcription and translation, spermatozoa undergo important functional transformation during epididymal transit. These modifications are essential to produce functionally efficient spermatozoa and depend on loss, modification, or remodeling of existing sperm proteins in response to the signals conveyed by the male reproductive tract. These signals are carefully regulated by an array of gene products in a stage-specific manner in which molecular chaperones play a key role.[22,23] Molecular chaperones are structurally diverse proteins that are expressed nearly in all cells in response to cellular stress.[24] In the present study, we observed that two of the top three networks identified by IPA were associated with underexpression of molecular chaperones. These networks were related to posttranslational modification, protein folding, and molecular transport. Furthermore, these networks were involved in protein trafficking related to cellular processes such as cell cycle, cell survival, cell death, and developmental disorder.

Proposed pathway 1

In the first network, eight proteins of the chaperonin-containing T-complex/TCP1-ring complex (CCT2, CCT3, CCT4, CCT5, CCT6A, CCT7, CCT8, and TCP1), four HSP 70 family chaperones (HSPA2, HSPA1L, HSP90AA1, and HSP8), and three members of the co-chaperonin DNAJ (DNAJA1, DNAJB1, and DNAJC3) were identified as key proteins (**Figure 2c**). Vesicular trafficking is conducted by gene products of Bardet–Biedl syndrome (BBS) and McKusick–Kaufman syndrome (MKKS).[25,26] Disruptions of BBS or MKKS genes result in male infertility owing to the failure of flagellum formation in spermatozoa.[26–29] The chaperonin-containing T-complex/TCP1-ring complex as part of the BBS/CCT complex may play a role in the assembly of BBSome, a complex involved in ciliogenesis regulating transport vesicles to the cilia. These proteins also have a role in the folding of actin and tubulin *in vitro*, and TUBB4B protein was identified in this network. Our liquid chromatography (LC)-MS/MS data also showed a decrease in tubulin (TUBB4B) as well as in both dynein heavy and light chains (DYNC1H1 and DYNC1LI1), further corroborating the concept of flagellar disassembly. Owing to the defects in flagellar disassembly, we did not find any change in Tektin (structural components of outer doublet microtubules) proteins as reported by us in spermatozoa of fertile men.[9]

The ribosomal subunits that include two translation elongation factors (EEF1E1 and EEF1D) and two tRNA ligases (DARS and VARS) identified in the network were also underexpressed. These are responsible for linking aspartate and valine to their cognate tRNAs, thus inhibiting the incorporation of these amino acids into the protein. In fact, Talevi *et al*.[30] have reported that treatment of spermatozoa with aspartate *in vitro* enhances sperm motility in oligozoospermic samples along with inhibition of DNA damage and membrane lipid peroxidation.

The HSP family proteins identified in the network were inducible by stress (*e.g.* HSPA1), constitutively expressed, or both (*e.g.* HSPH1, HSPA8, and HSP90AA1). Expression of some HSPs is developmentally regulated or restricted to specific cells.[31] HSPA2 is expressed on the sperm surface, and it is essential for sperm membrane remodeling during sperm–oocyte interaction. It may be used as a biochemical marker of human sperm function and male fertility.[32] Levels of human HSPA2 expression have been correlated with sperm maturity and male fertility.[32,33] Therefore, a decline in HSPA2 expression may be responsible for improper maturation, which was not so in our previous study in fertile men.[9]

We found HSPA1A to be downregulated from F1 to F4 in this study in infertile men as well as in our previous study in fertile donors. HSPA1A plays a pivotal role in the protein quality control system, ensuring the correct folding of proteins, re-folding of misfolded proteins, and controlling the targeting of proteins for subsequent degradation. It maintains protein homeostasis during cellular stress through two opposing mechanisms: protein refolding and degradation.[34]

The other HSPs and DNAJ family proteins are reported to affect spermatogenesis or sperm function.[12] Proteins responsible for mitochondria-mediated cell death (FAM162) and the negative regulator of mitochondrial respiration (TRAP1) were also downregulated (**Figure 2c**). Therefore, this network indicates that during the process of sperm production and maturation in infertile men, there is an impairment of the stress response protein that is demonstrated by the underexpression of chaperones in the network. This may lead to defects in the formation of the flagella and abortive apoptosis. Therefore, though the spermatozoa obtained in the most mature fraction (F4) appear normal, they are incapable of proper functioning after ejaculation.

The above findings were corroborated by the second predicted network (**Figure 2d**) showing deregulated posttranslational modification and protein folding, leading to developmental disorder.

Proposed pathway 2

The most important proteins were thioredoxins and peroxiredoxins which are reported to be abundantly expressed in developing spermatocytes.[35] Thioredoxin domain-containing protein 2 (TXNDC2) is a spermatid-specific thioredoxin-1 (also known as Sptrx-1) which is found exclusively in the tail of elongating spermatids and spermatozoa.[36]

Table 2: Key proteins showing an increasing trend of expression from fraction 1 (F1: most immature) through faction 4 (F4: most mature) involved in structural assembly of spermatozoa, spermatogenesis, reproduction, sperm motility, energy metabolism, and oxidation–reduction processes

Processes	*Proteins*	*Protein Name*	*Normalized Spectral Count*				*NASF ratio*		
Reproduction and spermatogenesis	ADAM29	Disintegrin and metalloproteinase domain-containing protein 29	0	1.3	7.0	17.3	0	0.08	0.40
	ADAM30	Disintegrin and metalloproteinase domain-containing protein 30	23.0	25.0	31.3	42.0	0.55	0.59	0.75
	AK7	Adenylate kinase 7	7.7	17.3	24.0	38.3	0.20	0.45	0.63
	ATP1A4	Sodium/potassium-transporting ATPase subunit alpha-4	46.0	93.0	105.3	131.3	0.35	0.71	0.80
	BSPH 1	Binder of sperm protein homolog 1	0.7	4.0	7.3	9.7	0.069	0.417	0.76
	CABYR	Calcium-binding tyrosine phosphorylation-regulated protein	0	104.7	138.0	175.3	0	0.60	0.79
	CCIN	Calicin	0	0	33.7	50.7	0	0	0.66
	GAPDHS	Glyceraldehyde-3-phosphate dehydrogenase, testis-specific	128.7	139.0	207.7	254.7	0.50	0.55	0.81
	IDH1	Isocitrate dehydrogenase [NADP] cytoplasmic	0	5.7	2.0	5.3	0	1.06	0.37
	IZUMO1	Izumo sperm-egg fusion protein 1	2.7	19.7	23.0	50.0	0.05	0.39	0.46
	KCNU1	Potassium channel subfamily U member 1	0	0	4.0	5.7	0	0	0.71
	RSPH 1	Radial spoke head 1 homolog	27.0	36.3	49.0	49.0	0.55	0.74	1.00
	SPACA1	Sperm acrosome membrane-associated protein 1	0	6.0	6.3	6.3	0	0.95	0
	SUN5	SUN domain-containing protein 5	0	0	3.7	4.0	0	0	0.92
Sperm motility	ATP1A4	Sodium/potassium-transporting ATPase subunit alpha-4	46.0	93.0	105.3	131.3	0.35	0.71	0.80
	CCDC39	Coiled-coil domain-containing protein 39	0	3.7	4.0	5.7	0	0.65	0.71
	DNAH1	Dynein heavy chain 1, axonemal	0	18.3	18.3	18.7	0	0.98	0.98
	GAPDHS	Glyceraldehyde-3-phosphate dehydrogenase, testis-specific	128.7	139.0	207.7	254.7	0.50	0.55	0.81
Cilium morphogenesis, axoneme assembly and cilium movement	AK7	Adenylate kinase 7	7.7	17.3	24.0	38.3	0.2	0.45	0.63
	CABYR	Calcium-binding tyrosine phosphorylation-regulated protein	0	104.7	138.0	175.3	0	0.60	0.79
	CCDC39	Coiled-coil domain-containing protein 39	0	3.7	4.0	5.7	0	0.65	0.71
	DNAH1	Dynein heavy chain 1, axonemal	0	18.3	18.3	18.7	0	0.98	0.98
	DNAI2	Dynein intermediate chain 2, axonemal	0	8.7	12.0	15.7	0	0.55	0.77
	RSPH 1	Radial spoke head 1 homolog	27.0	36.3	49.0	49.0	0.55	0.74	1
	SEPT7	Septin-7	0	0	5.7	6.0	0	0	0.94
Cell recognition	ADAM30	Disintegrin and metalloproteinase domain-containing protein 30	23.0	25.0	31.3	42.0	0.55	0.59	0.75
	CRTAC1	Radial spoke head 1 homolog	0	0	5.0	10.7	0	0	0.47
	DYNLL2	Dynein light chain 2, cytoplasmic	17.7	31.7	33.0	54.7	0.32	0.58	0.60
	IZUMO1	Izumo sperm-egg fusion protein 1	2.7	19.7	23.0	50.0	0.05	0.39	0.46
	KCNU1	Potassium channel subfamily U member 1	0	0	4.0	5.7	0	0	0.71
Carboxylic acid metabolic process	AASS	Alpha-aminoadipic semialdehyde synthase, mitochondrial	2.3	1.0	3.7	15.7	0.15	0.06	0.23
	ACAD8	Isobutyryl-CoA dehydrogenase, mitochondrial	2.0	7.0	12.7	15.7	0.13	0.45	0.81
	ALDH9A1	4-trimethylaminobutyraldehyde dehydrogenase	61.0	74.7	113.0	115.0	0.53	0.65	0.98
	CRAT	Carnitine O-acetyltransferase	48.3	51.3	66.7	112.3	0.43	0.46	0.59
	DPEP1	Dipeptidase 1	6.0	9.3	11.3	17.0	0.35	0.55	0.67
	GAPDHS	Glyceraldehyde-3-phosphate dehydrogenase, testis-specific	128.7	139.0	207.7	254.7	0.50	0.55	0.81
	GSTO2	Glutathione S-transferase omega-2	7.3	15.3	38.7	43.7	0.17	0.35	0.88
	HIBCH	3-hydroxyisobutyryl-CoA hydrolase, mitochondrial	2.0	7.3	10.0	13.0	0.15	0.56	0.77
	IDH1	Isocitrate dehydrogenase [NADP] cytoplasmic	0	5.7	2.0	5.3	0	1.06	0.37
	MUT	Methylmalonyl-CoA mutase, mitochondrial	3.3	10.3	11.7	16.0	0.21	0.64	0.73

Contd....

Table 2: Contnd...

Processes	*Proteins*	*Protein Name*	*Normalized Spectral Count*				*NASF ratio*		
Oxidation-reduction process	AASS	Alpha-aminoadipic semialdehyde synthase, mitochondrial	2.3	1.0	3.7	15.7	0.15	0.06	0.23
	ACAD8	Isobutyryl-CoA dehydrogenase, mitochondrial	2.0	7.0	12.7	15.7	0.13	0.45	0.81
	ALDH9A1	4-trimethylaminobutyraldehyde dehydrogenase	61.0	74.7	113.0	115.0	0.53	0.65	0.98
	CHDH	Choline dehydrogenase, mitochondrial	3.7	13.3	15.0	24.7	0.15	0.54	0.61
	CRAT	Carnitine O-acetyltransferase	48.3	51.3	66.7	112.3	0.43	0.46	0.59
	GAPDHS	Glyceraldehyde-3-phosphate dehydrogenase, testis-specific	128.7	139.0	207.7	254.7	0.50	0.55	0.81
	GSTO2	Glutathione S-transferase omega-2	7.3	15.3	38.7	43.7	0.17	0.35	0.88
	HTATIP2	Oxidoreductase HTATIP2	7.3	9.3	10.3	5.0	1.47	1.87	2.07
	IDH1	Isocitrate dehydrogenase [NADP] cytoplasmic	0	5.7	2.0	5.3	0	1.06	0.37
	MUT	Methylmalonyl-CoA mutase, mitochondrial	3.3	10.3	11.7	16.0	0.21	0.64	0.73
	SLC25A4	ADP/ATP translocase 1	1.0	2.0	4.7	6.7	0.15	0.30	0.70

NASF: normalized spectral abundance factor.

TXNDC2 transiently associates with the longitudinal columns of the fibrous sheath during sperm tail assembly but does not remain as a permanent component of the fibrous sheath in the mature sperm cell.[37] Although the expression of this protein showed a gradual decline from F1 to F4, still its presence in the most mature fraction implies defects in sperm flagella, which corroborates the findings in pathway 1 and our previous study regarding the expression of tektins.[9] Similarly, peroxiredoxin 4 (PRDX4) (**Figure 2d**) is present as a membrane-bound form only in testes, and the unprocessed form may be involved in acrosome formation during spermiogenesis.[38] Peroxiredoxins are involved in hydrogen peroxide-mediated signaling in spermatozoa[39] and prevent oxidative stress during capacitation.[40] PRDX4 knockout mice showed elevated spermatogenic cell death via oxidative stress.[41] We observed a decrease in the expression of voltage-dependent mitochondrial outer membrane ion channels (VDAC1, VDAC2, and VDAC3) along with mitochondrial ATP synthase complex subunits (ATP5I, ATP5F1, and ATP5J). These findings suggest mitochondrial dysfunction, which may lead to oxidative stress and cell death as mentioned above. In addition, underexpression of isocitrate dehydrogenase (IHD1), lipoprotein lipase (LPL), and apolipoprotein A1 (APOA1) may be responsible for impaired lipid metabolism. This may result in altered lipid profile of the sperm membrane that in turn may affect fertilization.[42] Furthermore, underexpression of dynactin (DCTN2), contractin (actin-related protein 1 homolog A [ACTR1A]), and F-actin capping protein (capping protein Z line alpha [CAPZA]) observed in our study may lead to defects in microtubule packaging and centrosome formation. Since the spermatozoon donates its centriole during fertilization and is essential for cleavage, these spermatozoa from infertile patients may not be able to initiate cleavage in embryos after fertilization. One of the major groups from the network that we identified were the endoplasmic reticulum-specific protein suppressor enhancer lin-12 1 like (SEL1L) and endoplasmic reticulum protein 29 (ERP29) along with the Ras-related proteins RBA6A and RAB7A. The presence of these proteins suggests impaired membrane traffic from the Golgi apparatus toward the endoplasmic reticulum. In addition, proteins involved in posttranslational modification and folding of proteins such as chaperones (HSP90B1, HSPA5, calnexin [CANX], calreticulin [CALR]), protein disulfide isomerase (prolyl 4-hydroxylase, beta polypeptide [P4HB], and protein disulfide isomerase-associated 3 [PDIA3]), peptidyl prolyl-*cis-trans* isomerase (FK506-binding protein [FKBP]), and serine-threonine phosphatase (peptidylpropyl isomerase B [PPIB]) were also identified in this network. Thus, an endoplasmic reticulum stress may be responsible for improper protein formation leading to developmental disorders in the proteins obtained in F4 fraction in infertile men.

Proposed pathway 3

The third network demonstrated a deregulated molecular transport, protein trafficking, and cell cycle. In the present study, we identified nuclear transport receptor IOP5 or importin-5 and alpha karyopherins such as KPNA2, KPNA3, KPNA4, and KNPNB. These karyopherins are expressed in specific stages during spermatogenesis in the spermatogonium (KNNA1, KPNA2, and KPNA3), spermatocyte (KPNA2, KPNA3, and KPNA4), round spermatid (KPNA3 and KPNA4), and elongating spermatids (KPNA2 and IPO5).[43] These importins were all underexpressed in ejaculated sperm in our study. This finding suggests that together with RNA GTPase they may be responsible for impaired import of testis specific H1 histone (HIST1H1T). HIST1H1T is required for less compaction of nuclear DNA during meiosis and is subsequently replaced by protamines. Thus, these importins may play a role in nuclear abnormalities that may eventually render the spermatozoa incapable of fertilization or successful completion of postfertilization steps.

The other two important proteins involved in exosome-mediated cargo delivery were ANXA1 and ANXA2.[44] These were underexpressed in this network. It is hypothesized that sperm maturation in the epididymis and regulation of sperm function are mediated by regulatory RNAs that are delivered by exosomes.[45,46] In fact, Yang *et al*.[46] reported the presence of annexins in the seminal exosomes. Thus, an improper delivery of regulatory factors to the spermatozoa via exosomes might result in sperm dysfunction.

Proteasomes present in mammalian sperm play a pivotal role during fertilization. The enzymatic activity of the proteasome is modulated by protein kinase A and is involved in the progesterone-induced acrosome reaction. Furthermore, after capacitation, the acrosomal proteasomes facilitate the degradation of zona pellucida glycoproteins leading up to fertilization.[47] In this network, we identified eight subunits of 26S proteasome (PSMC3, PSMC4, PSMC6, PSMD5, PSMD7, PSMD8, PSMD11, and PSMD13). One of the nodal proteins, matrix-metalloproteinase (MME) or Neprilysin, was also identified in the network. Pinto *et al*.[48] reported that tachykinins present in human spermatozoa participate in the regulation of sperm motility, and their activity is regulated by Neprilysin. Another nodal protein involved in this network was 14-3-3 protein theta (YWHAQ) adapter protein which has been implicated in the regulation of a large spectrum of both general and specialized signaling pathways,[49] thus implying a signaling

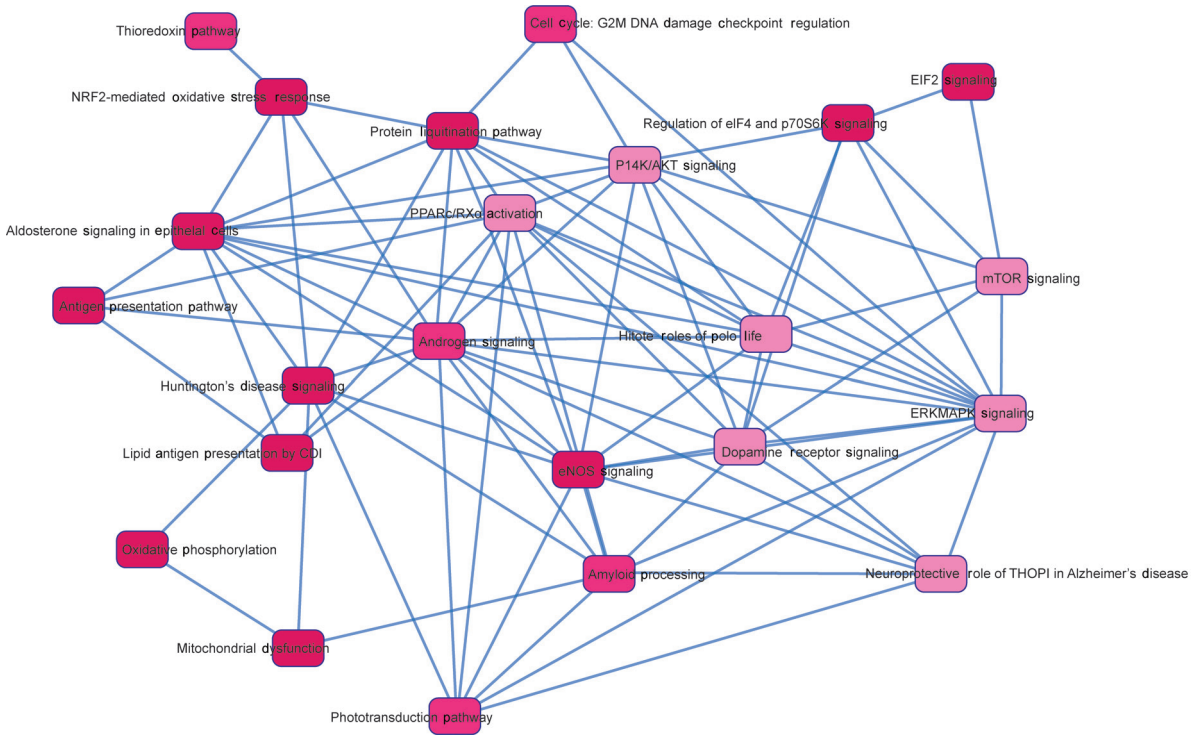

Figure 3: Fifteen overlapping pathways and eight connecting signaling pathways were predicted based on the identified common genes. Their translated protein products were underexpressed starting from F1 to F4. F1: least mature stage having the lowest density; F2, F3: intermediate stages; F4: includes the most dense and morphologically mature motile spermatozoa. Full names of abbreviated proteins are presented in **Supplementary Information**.

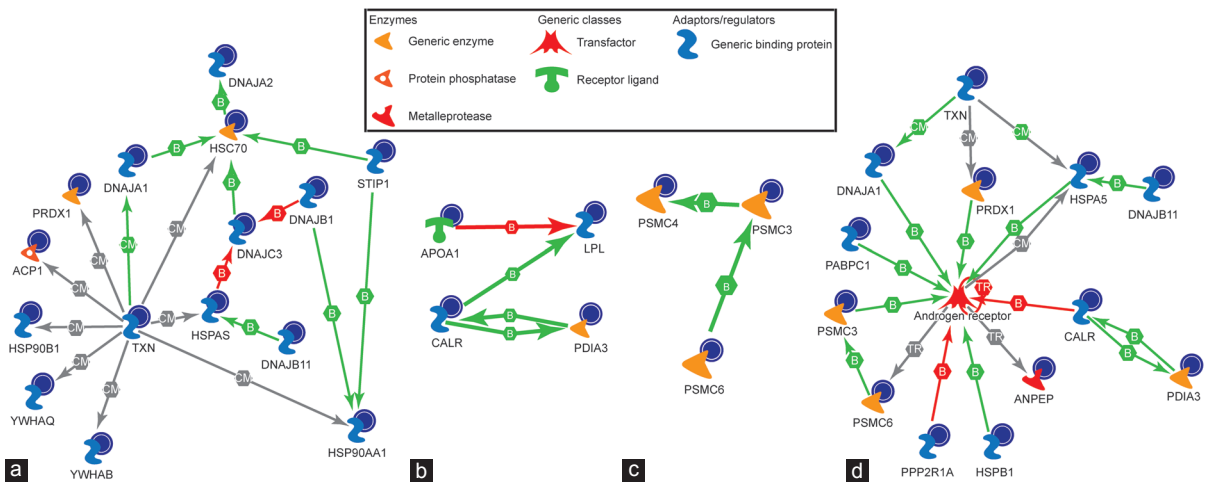

Figure 4: The biochemical process networks of the differentially expressed proteins generated by MetaCore™ software using a transcription regulation algorithm. The biochemical process regulation networks were ranked by a value and interpreted in terms of GO. Gene network illustrates proteins and interactions of the differentially expressed proteins. MetaCore™ pathway analysis predicting androgen receptor as the major regulator and pathways involved were; (**a**) chaperone-mediated stress response; (**b**) sperm function; (**c**) proteosomal pathway; and (**d**) predicted impaired androgen receptor signaling in spermatozoa of infertile patients. GO: gene ontology; F1: least mature stage having the lowest density; F2, F3: intermediate stages; F4: includes the most dense and morphologically mature motile spermatozoa. Full names of abbreviated proteins are presented in **Supplementary Information**.

failure in F4 fraction in our study. We identified 15 overlapping networks on the basis of mutual genes that were affected during the process of maturation in these infertile patients. These networks may be responsible for poor sperm quality.

MetaCore™ pathway analysis

While IPA reports protein–protein interactions, MetaCore™ is an integrated software suite for functional analysis of experimental data of human protein–protein interactions, protein–DNA interactions, transcription factors, and signaling and metabolic pathways including disease and toxicity. The top pathway affected during maturation was identified to be mediated through TXN, PRDX1, signaling adaptor molecules (YWHAQ and YWHAB), and the molecular chaperons (HSP90AA1, HSP90B1, HSC70, HSPA5, DNAJA1, DNAJ2, DNAJB1, DNAJB11, DNAJC3, and STIP1). Taken together, all the networks and pathways described above suggest that the spermatozoa

at different stages of maturation suffer from endoplasmic reticulum stress (ER stress) due to accumulation of unfolded or misfolded proteins as a result of underexpressed chaperone and proteasome activities, which will lead to induction of oxidative stress and abortive apoptosis. This is corroborated by the fact that the Clusterin is not differentially expressed in the spermatozoa of infertile men like that reported for fertile men.[9] Clusterin isoforms are responsible for differential regulation of apoptosis where the nuclear form promotes the process and the mitochondrial form opposes the process (Uniprot). In fertile men, its proper expression prevented apoptosis of mature spermatozoa where the mitochondria are intact and eliminated the spermatozoa with nuclear defects. Therefore, it is suggested that the process was impaired in infertile men leading to abortive apoptosis. Taken together, the cells that may appear morphologically mature may carry structural anomalies in the cytoskeleton, microtubules, and flagellum, thereby affecting the function of the spermatozoa. Androgen receptor is a transcription regulator. The presence of the AR in human spermatozoa has been demonstrated by Western blot and immunofluorescence assay.[50] AR is also localized in the head region.[51] In addition to stimulating cell growth, androgens or the AR plays an important role in apoptosis involving both intrinsic and extrinsic pathways. Short exposure of ejaculated spermatozoa to androgens produces an increase in AR phosphorylation, especially on the 110 kDa band which is the less expressed isoform in sperm cells.[51]

Taken together, all the networks and pathways described in the present study show that the spermatozoa at different stages of maturation suffer from ER stress due to accumulation of unfolded or misfolded proteins as a result of underexpressed chaperone and proteasome activities. This may lead to induction of oxidative stress and abortive apoptosis.[52] Thus, although the spermatozoa in F4 fraction may appear morphologically mature and exhibit good motility and morphology following separation on density gradient as is the case during intrauterine insemination or *in vitro* fertilization, these sperm can potentially carry structural anomalies in the cytoskeleton, microtubule, and flagellum, thereby affecting the functional capabilities of the spermatozoa. A limitation of our study was that we could not perform the western blot validation of the MS-spectrometry data as the quantity of protein from the fractions, particularly F3 and F4, was insufficient.

In conclusion, the results of the present study show that a defective signaling cascade is responsible for the defective sperm function observed in infertile patients, particularly a decline in mitochondrial function and oxidative phosphorylation in the most mature fraction (F4), implying a state of energy deprivation. Furthermore, a dysregulated protein turnover and protein folding may lead to accumulation of defective proteins or proteins that otherwise would have been eliminated during the process of maturation and thus result in the impairment of sperm function. The present study provides mounting evidence that aberrant chaperone expression may be a major contributing factor to the defective sperm function seen in many cases of male infertility. Our study also identified the involvement of AR as the core player in the downregulation of the signaling leading to production of defective spermatozoa incapable of fertilization. Future studies conducted in larger study population are necessary to examine the distribution of proteins in infertile men with different clinical diagnosis such as varicocele and validate the important proteins to further help in understanding the underlying pathology of male infertility.

AUTHOR CONTRIBUTIONS

AA and RS planned the experiments. ZC selected the samples according to the clinical histories. ZC and LS analyzed the collected data. LS was involved in the manuscript preparation. RS, AA, and LS critically reviewed the manuscript. All authors read and approved the final manuscript.

COMPETING INTERESTS

All authors declared no competing interests.

ACKNOWLEDGMENTS

The authors are grateful to the Andrology Center technologists for scheduling the study subjects; Cleveland Clinic Arts Department for creating the four figures; Belinda Willard, Director, Proteomic Core Lab, Lerner Research Institute for providing assistance with proteomic analysis and Banu Gopalan for the bioinformatics analysis of the results. The Orbitrap Elite mass spectrometer used in this study was purchased with funds from an NIH shared instrument grant 1S10RR031537-01 to Belinda Willard. Financial support was provided by the American Center for Reproductive Medicine, Cleveland Clinic. Dr. Zhihong Cui's visit was supported by a fellowship from the Chinese Government and the Institute of Toxicology, College of Preventive Medicine, the Third Military Medical University, Chongqing, China.

Supplementary information is linked to the online version of the paper on the *Asian Journal of Andrology* website.

REFERENCES

1 Jones RE, Kristen L. The male reproductive system. In: Jones R, Kristen L, editors. Human Reproductive Biology. Cambridge: Academic Press, Elsevier; 2013.
2 Sharma R, Agarwal A. Spermatogenesis: An overview. In: Armand Zini A, editor. Sperm chromatin: biological and clinical applications in male infertility and assisted reproduction. New York: Springer; 2011. p 19–44.
3 Kishigami S, Wakayama S, Nguyen VT, Wakayama T. Similar time restriction for intracytoplasmic sperm injection and round spermatid injection into activated oocytes for efficient offspring production. *Biol Reprod* 2004; 70: 1863–9.
4 Chemes HE. Phenotypes of sperm pathology: genetic and acquired forms in infertile men. *J Androl* 2000; 21: 799–808.
5 Gil-Guzman E, Ollero M, Lopez MC, Sharma RK, Alvarez JG, *et al*. Differential production of reactive oxygen species by subsets of human spermatozoa at different stages of maturation. *Hum Reprod* 2001; 16: 1922–30.
6 Gaucher J, Reynoird N, Montellier E, Boussouar F, Rousseaux S, *et al*. From meiosis to postmeiotic events: the secrets of histone disappearance. *FEBS J* 2010; 277: 599–604.
7 Guo W, Qu F, Xia L, Guo Q, Ying X, *et al*. Identification and characterization of ERp29 in rat spermatozoa during epididymal transit. *Reproduction* 2007; 133: 575–84.
8 Govin J, Gaucher J, Ferro M, Debernardi A, Garin J, *et al*. Proteomic strategy for the identification of critical actors in reorganization of the post-meiotic male genome. *Mol Hum Reprod* 2012; 18: 1–13.
9 Cui Z, Sharma R, Agarwal A. Proteomic analysis of mature and immature ejaculated spermatozoa from fertile men. *Asian J Androl* 2016; 18: 735–46.
10 Sakkas D, Mariethoz E, St John JC. Abnormal sperm parameters in humans are indicative of an abortive apoptotic mechanism linked to the Fas-mediated pathway. *Exp Cell Res* 1999; 251: 350–5.
11 Sakkas D, Mariethoz E, Manicardi G, Bizzaro D, Bianchi PG, *et al*. Origin of DNA damage in ejaculated human spermatozoa. *Rev Reprod* 1999; 4: 31–7.
12 Dun MD, Aitken RJ, Nixon B. The role of molecular chaperones in spermatogenesis and the post-testicular maturation of mammalian spermatozoa. *Hum Reprod Update* 2012; 18: 420–35.
13 Lin C, Tholen E, Jennen D, Ponsuksili S, Schellander K, *et al*. Evidence for effects of testis and epididymis expressed genes on sperm quality and boar fertility traits. *Reprod Domest Anim* 2006; 41: 538–43.
14 Ollero M, Gil-Guzman E, Lopez MC, Sharma RK, Agarwal A, *et al*. Characterization of subsets of human spermatozoa at different stages of maturation: implications in the diagnosis and treatment of male infertility. *Hum Reprod* 2001; 16: 1912–21.
15 Brahem S, Mehdi M, Elghezal H, Saad A. Semen processing by density gradient centrifugation is useful in selecting sperm with higher double-strand DNA integrity. *Andrologia* 2011; 43: 196–202.
16 Castillo J, Simon L, de Mateo S, Lewis S, Oliva R. Protamine/DNA ratios and DNA damage in native and density gradient centrifuged sperm from infertile patients. *J Androl* 2011; 32: 324–32.
17 Esteves SC, Roque M, Bradley CK, Garrido N. Reproductive outcomes of testicular versus ejaculated sperm for intracytoplasmic sperm injection among men with high levels of DNA fragmentation in semen: systematic review and meta-analysis. *Fertil Steril* 2017; 108: 456–67.e1.
18 World Health Organization. WHO laboratory manual for the examination and

processing of human semen. 5th ed. Geneva: WHO Press; 2010.
19 Keller A, Nesvizhskii AI, Kolker E, Aebersold R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* 2002; 74: 5383–92.
20 Zybailov B, Coleman MK, Florens L, Washburn MP. Correlation of relative abundance ratios derived from peptide ion chromatograms and spectrum counting for quantitative proteomic analysis using stable isotope labeling. *Anal Chem* 2005; 77: 6218–24.
21 Boyle EI, Weng S, Gollub J, Jin H, Botstein D, *et al*. GO: termfinder – open source software for accessing gene ontology information and finding significantly enriched gene ontology terms associated with a list of genes. *Bioinformatics* 2004; 20: 3710–5.
22 Hermo L, Pelletier RM, Cyr DG, Smith CE. Surfing the wave, cycle, life history, and genes/proteins expressed by testicular germ cells. Part 1: background to spermatogenesis, spermatogonia, and spermatocytes. *Microsc Res Tech* 2010; 73: 241–78.
23 Hermo L, Pelletier RM, Cyr DG, Smith CE. Surfing the wave, cycle, life history, and genes/proteins expressed by testicular germ cells. Part 5: intercellular junctions and contacts between germs cells and sertoli cells and their regulatory interactions, testicular cholesterol, and genes/proteins associated with more than one germ cell generation. *Microsc Res Tech* 2010; 73: 409–94.
24 Ellis RJ. Discovery of molecular chaperones. *Cell Stress Chaperones* 1996; 1: 155–60.
25 Nachury MV, Loktev AV, Zhang Q, Westlake CJ, Peranen J, *et al*. A core complex of BBS proteins cooperates with the GTPase Rab8 to promote ciliary membrane biogenesis. *Cell* 2007; 129: 1201–13.
26 Lechtreck KF, Johnson EC, Sakai T, Cochran D, Ballif BA, *et al*. The chlamydomonas reinhardtii BBSome is an IFT cargo required for export of specific signaling proteins from flagella. *J Cell Biol* 2009; 187: 1117–32.
27 Mykytyn K, Mullins RF, Andrews M, Chiang AP, Swiderski RE, *et al*. Bardet-Biedl syndrome type 4 (BBS4)-null mice implicate Bbs4 in flagella formation but not global cilia assembly. *Proc Natl Acad Sci U S A* 2004; 101: 8664–9.
28 Nishimura DY, Swiderski RE, Searby CC, Berg EM, Ferguson AL, *et al*. Comparative genomics and gene expression analysis identifies BBS9, a new Bardet-Biedl syndrome gene. *Am J Hum Genet* 2005; 77: 1021–33.
29 Davis RE, Swiderski RE, Rahmouni K, Nishimura DY, Mullins RF, *et al*. A knockin mouse model of the Bardet-Biedl syndrome 1 M390R mutation has cilia defects, ventriculomegaly, retinopathy, and obesity. *Proc Natl Acad Sci U S A* 2007; 104: 19422–7.
30 Talevi R, Barbato V, Fiorentino I, Braun S, Longobardi S, *et al*. Protective effects of *in vitro* treatment with zinc, d-aspartate and coenzyme q10 on human sperm motility, lipid peroxidation and DNA fragmentation. *Rep Biol Endocrinol* 2013; 11: 81.
31 Rupik W, Jasik K, Bembenek J, Widlak W. The expression patterns of heat shock genes and proteins and their role during vertebrate's development. *Comp Biochem Physiol A Mol Integr Physiol* 2011; 159: 349–66.
32 Naaby-Hansen S, Diekman A, Shetty J, Flickinger CJ, Westbrook A, *et al*. Identification of calcium-binding proteins associated with the human sperm plasma membrane. *Reprod Biol Endocrinol* 2010; 8: 6.
33 Mamelak D, Lingwood C. Expression and sulfogalactolipid binding specificity of the recombinant testis-specific cognate heat shock protein 70. *Glycoconj J* 1997; 14: 715–22.
34 Seo JH, Park JH, Lee EJ, Vo TT, Choi H, *et al*. ARD1-mediated Hsp70 acetylation balances stress-induced protein refolding and degradation. *Nat Commun* 2016; 7: 12882.
35 Hanschmann EM, Godoy JR, Berndt C, Hudemann C, Lillig CH. Thioredoxins, glutaredoxins, and peroxiredoxins – molecular mechanisms and health significance: from cofactors to antioxidants to redox signaling. *Antioxid Redox Signal* 2013; 19: 1539–605.
36 Yu Y, Oko R, Miranda-Vizuete A. Developmental expression of spermatid-specific thioredoxin-1 protein: transient association to the longitudinal columns of the fibrous sheath during sperm tail formation. *Biol Reprod* 2002; 67: 1546–54.
37 Ranney MK, Ahmed IS, Potts KR, Craven RJ. Multiple pathways regulating the anti-apoptotic protein clusterin in breast cancer. *Biochim Biophys Acta* 2007; 1772: 1103–11.
38 Matsuki S, Sasagawa I, Iuchi Y, Fujii J. Impaired expression of peroxiredoxin 4 in damaged testes by artificial cryptorchidism. *Redox Report* 2002; 7: 276–8.
39 O'Flaherty C, de Souza AR. Hydrogen peroxide modifies human sperm peroxiredoxins in a dose-dependent manner. *Biol Reprod* 2011; 84: 238–47.
40 Lee D, Moawad AR, Morielli T, Fernandez MC, O'Flaherty C. Peroxiredoxins prevent oxidative stress during human sperm capacitation. *Mol Hum Reprod* 2017; 23: 106–15.
41 Iuchi Y, Okada F, Tsunoda S, Kibe N, Shirasawa N, *et al*. Peroxiredoxin 4 knockout results in elevated spermatogenic cell death via oxidative stress. *Biochem J* 2009; 419: 149–58.
42 Koh HJ, Lee SM, Son BG, Lee SH, Ryoo ZY, *et al*. Cytosolic $NADP^+$-dependent isocitrate dehydrogenase plays a key role in lipid metabolism. *J Biol Chem* 2004; 279: 39968–74.
43 Major AT, Whiley PA, Loveland KL. Expression of nucleocytoplasmic transport machinery: clues to regulation of spermatogenic development. *Biochim Biophys Acta* 2011; 1813: 1668–88.
44 Leoni G, Neumann PA, Kamaly N, Quiros M, Nishio H, *et al*. Annexin A1-containing extracellular vesicles and polymeric nanoparticles promote epithelial wound repair. *J Clin Invest* 2015; 125: 1215–27.
45 Vojtech L, Woo S, Hughes S, Levy C, Ballweber L, *et al*. Exosomes in human semen carry a distinctive repertoire of small non-coding RNAs with potential regulatory functions. *Nucleic Acids Res* 2014; 42: 7290–304.
46 Yang C, Guo WB, Zhang WS, Bian J, Yang JK, *et al*. Comprehensive proteomics analysis of exosomes derived from human seminal plasma. *Andrology* 2017; 5: 1007–15.
47 Kerns K, Morales P, Sutovsky P. Regulation of sperm capacitation by the 26S proteasome: an emerging new paradigm in spermatology. *Biol Reprod* 2016; 94: 117.
48 Pinto FM, Ravina CG, Subiran N, Cejudo-Roman A, Fernandez-Sanchez M, *et al*. Autocrine regulation of human sperm motility by tachykinins. *Reprod Biol Endocrinol* 2010; 8: 104.
49 Olsen JV, Blagoev B, Gnad F, Macek B, Kumar C, *et al*. Global, *in vivo*, and site-specific phosphorylation dynamics in signaling networks. *Cell* 2006; 127: 635–48.
50 Solakidi S, Psarra AM, Nikolaropoulos S, Sekeris CE. Estrogen receptors alpha and beta (ERalpha and ERbeta) and androgen receptor (AR) in human sperm: localization of ERbeta and AR in mitochondria of the midpiece. *Hum Reprod* 2005; 20: 3481–7.
51 Aquila S, Middea E, Catalano S, Marsico S, Lanzino M, *et al*. Human sperm express a functional androgen receptor: effects on PI3K/AKT pathway. *Hum Reprod* 2007; 10: 2594–605.
52 Nakagawa T, Zhu H, Morishima N, Li E, Xu J, *et al*. Caspase-12 mediates endoplasmic-reticulum-specific apoptosis and cytotoxicity by amyloid-beta. *Nature* 2000; 403: 98–103.

Supplementary Table 1: Proteins showing an increasing trend during sperm maturation

Protein	*Uniprot number*	*MW*	*F1*	*F2*	*F3*	*F4*
		kDa	*Average SC*	*Average SC*	*Average SC*	*Average SC*
Actin-like protein 9	Q8TC94	46	0.0	5.0	8.3	50.3
AMY-1-associating protein expressed in testis 1	Q7Z4T9	90	0.0	0.0	3.0	4.3
Calcium-binding tyrosine phosphorylation-regulated protein isoform C	O75952	41	0.0	104.7	138.0	175.3
Calicin	Q13939	67	0.0	0.0	33.7	50.7
Cartilage acidic protein 1 isoform B precursor	Q9NQ79	70	0.0	0.0	5.0	10.7
Coiled-coil domain-containing protein 39	Q9UFE4	110	0.0	3.7	4.0	5.7
Disintegrin and metalloproteinase domain-containing protein 29 preproprotein	Q9UKF5	93	0.0	1.3	7.0	17.3
Dynein heavy chain 1, axonemal	Q9P2D7	488	0.0	18.3	18.3	18.7
Dynein heavy chain 12, axonemal isoform 1	Q6ZR08	357	0.0	8.7	12.0	15.7
Fibronectin type III domain-containing protein 8	Q8TC99	36	0.0	3.0	15.0	26.7
Fibrous sheath-interacting protein 2	J3QTJ6	790	0.0	0.0	2.3	13.0
GLIPR1-like protein 2	Q4G1C9	29	0.0	15.3	16.7	20.0
Isochorismatase domain-containing protein 2, mitochondrial isoform 1	P50213	22	0.0	56.3	60.7	60.7
Isocitrate dehydrogenase [NAD] subunit gamma, mitochondrial isoform A precursor	O75874	43	0.0	5.7	2.0	5.3
Leucine-rich repeat and IQ domain-containing protein 4	Q53EV4	64	0.0	3.0	6.7	13.3
Leucine-rich repeat-containing protein 23 isoform A	A6NMS7	40	0.0	11.0	16.3	18.0
leucine-rich repeat-containing protein 48 isoform A	Q08722	61	0.0	15.7	30.7	52.7
Mitochondrial thiamine pyrophosphate carrier	Q10713	36	0.0	2.3	8.7	9.0
Poly(ADP-ribose) glycohydrolase ARH3	Q15365	39	0.0	2.7	2.3	9.0
Potassium channel subfamily U member 1	A8MYU2	130	0.0	0.0	4.0	5.7
Protein FAM205A	Q6ZU69	148	0.0	0.0	11.7	39.3
Septin-7 isoform 1	Q16181	51	0.0	0.0	5.7	6.0
Sodium/potassium-transporting ATPase subunit beta-1	P05026	35	0.0	0.0	2.7	2.7
Speriolin isoform 1	Q9HBV2	62	0.0	6.0	6.3	6.3
SUN domain-containing protein 5	Q8TC36	43	0.0	0.0	3.7	4.0
Tripartite motif-containing protein 42	Q8IWZ5	83	0.0	0.0	1.3	10.7
Uncharacterized protein LOC730159 precursor		21	0.0	0.0	5.7	6.7
WD repeat-containing protein 52 isoform 1	Q9GZS3	214	0.0	15.7	17.0	24.0
Bovine seminal plasma protein homolog 1 precursor	Q075Z2	16	0.7	4.0	7.3	9.7
Probable Xaa-Pro aminopeptidase 3 isoform 1	Q9NQH7	57	0.7	6.3	16.0	16.7
ADP/ATP translocase 1	P12235	33	1.0	2.0	4.7	6.7
3-hydroxyisobutyryl-CoA hydrolase, mitochondrial isoform 1 precursor	Q6NVY1	43	2.0	7.3	10.0	13.0
Actin-like protein 7A	Q9Y615	49	2.0	8.0	15.3	42.0
Isobutyryl-CoA dehydrogenase, mitochondrial	Q9UKU7	45	2.0	7.0	12.7	15.7
Alpha-aminoadipic semialdehyde synthase, mitochondrial	Q9UDR5	102	2.3	1.0	3.7	15.7
Izumo sperm-egg fusion protein 1 precursor	Q8IYV9	39	2.7	19.7	23.0	50.0
Mitochondrial-processing peptidase subunit alpha precursor	Q10713	58	2.7	4.7	5.7	16.7
Radial spoke head protein 3 homolog	Q86UC2	64	2.7	20.0	24.7	31.3
Adenylate kinase domain-containing protein 1 isoform 1	Q5TCS8	221	3.3	39.0	58.3	58.7
Methylmalonyl-CoA mutase, mitochondrial precursor	P22033	83	3.3	10.3	11.7	16.0
Sorting and assembly machinery component 50 homolog	Q9Y512	52	3.3	5.7	10.7	18.3
Choline dehydrogenase, mitochondrial	Q8NE62	65	3.7	13.3	15.0	24.7
Cytochrome b5 domain-containing protein 1	Q6P9G0	27	4.0	9.7	10.7	11.7
Dynein intermediate chain 2, axonemal isoform 1	Q9GZS0	69	4.0	17.3	20.0	21.7
Peptidyl-prolyl cis-trans isomerase-like 6 isoform 1	Q8IXY8	35	4.0	13.7	18.3	22.7
Putative transferase CAF17, mitochondrial precursor	Q5T440	38	5.3	11.7	14.7	17.7
Dipeptidase 1 precursor	P16444	46	6.0	9.3	11.3	17.0
Glutathione S-transferase omega-2 isoform 1	Q9H4Y5	28	7.3	15.3	38.7	43.7
Integrin alpha-M isoform 1 precursor	P11215	127	7.3	19.0	0.0	0.0
Oxidoreductase HTATIP2 isoform A precursor	Q9BUP3	30	7.3	9.3	10.3	5.0
Adenylate kinase 7	Q96M32	83	7.7	17.3	24.0	38.3
Actin-related protein M1	Q9BYD9	41	9.0	19.7	44.7	70.0
Sodium/potassium-transporting ATPase subunit alpha-1 isoform A	P05023	113	9.0	12.3	16.3	19.0
SPRY domain-containing protein 7 isoform 1	Q5W111	22	9.3	10.3	13.0	20.7
EF-hand domain-containing protein KIAA0494	O75071	55	9.7	28.0	34.7	41.7

Contd...

Supplementary Table 1: Contd...

Protein	*Uniprot number*	*MW*	*F1*	*F2*	*F3*	*F4*
		kDa	*Average SC*	*Average SC*	*Average SC*	*Average SC*
Uncharacterized protein KIAA1683 isoform A	Q9H0B3	147	12.0	49.3	62.0	81.0
Serine/threonine-protein phosphatase with EF-hands 1 isoform 1	O14829	76	13.0	31.7	32.7	61.0
Casein kinase II subunit beta	P67870	25	16.7	20.0	26.0	34.3
Deoxyguanosine kinase, mitochondrial isoform A precursor	Q16854	32	17.0	20.0	32.0	40.3
Fibrinogen-like protein 1 precursor	Q08830	36	17.0	19.7	30.0	39.7
Dynein light chain 2, cytoplasmic	Q96FJ2	10	17.7	31.7	33.0	54.7
Disintegrin and metalloproteinase domain-containing protein 30 preproprotein	Q9UKF2	89	23.0	25.0	31.3	42.0
Radial spoke head 1 homolog	Q8WYR4	35	27.0	36.3	49.0	49.0
Izumo sperm-egg fusion protein 2 precursor	Q6UXV1	25	29.7	29.7	36.7	46.3
Coiled-coil domain-containing protein 147	Q5T655	103	32.7	53.3	60.0	76.7
Actin-related protein T2	Q8TDY3	42	33.0	59.0	64.7	87.0
Protein FAM71B	Q8TC56	65	41.3	60.0	81.0	103.3
Sodium/potassium-transporting ATPase subunit alpha-4 isoform 1	Q13733	114	46.0	93.0	105.3	131.3
Beta-galactosidase-1-like protein precursor	Q6UWU2	74	48.3	63.7	71.3	105.3
Carnitine O-acetyltransferase precursor	P43155	71	48.3	51.3	66.7	112.3
4-trimethylaminobutyraldehyde dehydrogenase	P49189	56	61.0	74.7	113.0	115.0
Glycerol kinase 2	Q14410	61	118.3	148.3	151.0	171.7
Stress-70 protein, mitochondrial precursor	P38646	74	121.7	130.3	140.7	206.7
Glyceraldehyde-3-phosphate dehydrogenase, testis-specific	O14556	45	128.7	139.0	207.7	254.7

F1: fraction 1; F2: fraction 2; F3: fraction 3; F4: fraction 4; SC: spectral count; MW: molecular weight

Supplementary Table 2: Proteins showing a decreasing trend during sperm maturation

Protein	*Uniprot number*	*MW*	*F1*	*F2*	*F3*	*F4*
		kDa	*Average SC*	*Average SC*	*Average SC*	*Average SC*
Lactotransferrin isoform 1 precursor	P02788	78	2260.0	1810.3	1407.3	299.3
Endoplasmin precursor	P14625	92	1002.0	609.0	436.7	246.0
Tubulin beta-4B chain	P68371	50	927.3	707.3	457.3	406.3
78 glucose-regulated protein precursor	P11021	72	875.7	528.0	506.7	368.7
Heat shock protein HSP 90-alpha isoform 1	Q86SX1	98	777.7	680.7	531.0	385.7
Semenogelin-2 precursor	P07900	65	712.0	397.7	170.0	70.0
Heat shock-related 70 protein 2	P54652	70	707.7	596.0	326.7	276.0
Protein disulfide-isomerase A3 precursor	P30101	57	622.3	389.3	367.7	212.0
Tubulin alpha-3C/D chain	Q13748	50	598.0	474.3	284.7	264.7
Hypoxia up-regulated protein 1 precursor	Q9Y4L1	111	509.3	331.7	173.7	80.3
60 HSP, mitochondrial	P10809	61	482.3	283.3	66.3	57.0
Aminopeptidase N precursor	P15144	110	363.7	258.3	95.0	54.0
Importin-5	B3KWG6	126	348.3	172.7	61.0	48.7
Semenogelin-1 preproprotein	O00410	52	336.7	212.3	166.0	87.7
Calreticulin precursor	P04279	48	291.7	195.0	192.0	110.3
RuvB-like 2	P27797	51	287.0	208.3	110.3	71.7
Actin, cytoplasmic 1	Q9Y230	42	285.3	164.0	84.0	61.3
Uncharacterized protein C1orf56 precursor	P60709	37	263.7	103.3	30.7	29.7
Heat shock 70 protein 1-like	Q9BUN1	70	258.3	236.3	183.3	130.0
Prostatic acid phosphatase isoform TM-PAP precursor	P34931	48	249.7	151.7	69.7	33.3
Elongation factor 1-gamma	P15309	50	247.3	167.0	128.0	80.7
Phosphoglycerate kinase 2	P26641	45	240.7	199.0	176.7	145.0
Calmegin precursor	P07205	70	229.7	94.7	14.0	0.0
Calnexin precursor	O14967	68	227.3	130.7	64.7	39.3
Protein disulfide-isomerase precursor	P27824	57	222.7	131.0	79.7	35.3
Importin subunit beta-1	P07237	97	217.0	78.3	47.7	28.7
T-complex protein 1 subunit beta isoform 1	Q14974	57	200.3	139.7	46.0	27.0
T-complex protein 1 subunit gamma isoform A	P78371	61	199.0	106.7	27.0	24.7
T-complex protein 1 subunit eta isoform A	P49368	59	193.0	144.0	58.3	35.0
Vesicular integral-membrane protein VIP36 precursor	Q99832	40	192.7	118.3	82.0	59.7

Contd...

Supplementary Table 2: Contd...

Protein	*Uniprot number*	*MW*	*F1*	*F2*	*F3*	*F4*
		kDa	*Average SC*	*Average SC*	*Average SC*	*Average SC*
T-complex protein 1 subunit theta	Q12907	60	189.3	126.7	48.0	33.0
Valyl-tRNA synthetase	P50990	140	176.3	95.7	91.0	20.0
Cullin-associated NEDD8-dissociated protein 1	P26640	136	168.0	138.0	82.3	53.7
T-complex protein 1 subunit delta	Q86VP6	58	162.0	98.7	44.3	30.3
Heat shock 70 protein 1A/1B	P50991	70	147.7	93.7	39.7	11.7
Heme oxygenase 2	P08107	36	145.3	30.3	0.0	0.0
Neprilysin	P30519	86	141.7	107.0	37.0	23.0
Histone H2B type 1-A	P08473	14	137.3	73.7	35.3	0.0
Cytoplasmic dynein 1 heavy chain 1	Q96A08	532	136.7	135.3	2.3	0.0
60S acidic ribosomal protein P0	Q14204	34	136.3	61.0	33.0	30.3
Heat shock cognate 71 protein isoform 1	P05388	71	122.7	72.7	47.7	41.0
Transmembrane emp24 domain-containing protein 10 precursor	P11142	25	122.7	72.0	71.0	54.0
Protein disulfide-isomerase A6 precursor	P49755	48	119.0	84.0	43.7	23.7
Peroxiredoxin-4 precursor	Q15084	31	118.0	85.7	83.3	56.0
Calcium-binding tyrosine phosphorylation-regulated protein isoform A	Q13162	53	117.0	26.7	22.0	21.7
Fatty-acid amide hydrolase 1	O75952	63	114.0	83.3	77.7	73.3
Peroxiredoxin-6	O00519	25	109.7	77.3	14.7	2.0
Stomatin-like protein 2	P30041	39	105.3	72.7	28.0	25.7
T-complex protein 1 subunit epsilon	Q9UJZ1	60	104.0	58.7	31.3	15.3
T-complex protein 1 subunit zeta isoform A	P48643	58	103.7	74.3	23.0	11.7
cAMP-dependent protein kinase type II-alpha regulatory subunit	P40227	46	100.0	82.7	59.0	42.3
ATP synthase subunit b, mitochondrial precursor	P13861	29	99.7	86.0	35.3	37.3
Ras-related protein Rab-14	P24539	24	95.0	46.3	30.0	12.7
Dehydrogenase/reductase SDR family member 7 precursor	P61106	38	93.3	47.7	10.0	0.0
Prostate-specific antigen isoform 1 preproprotein	Q9Y394	29	91.7	23.0	3.0	0.0
Voltage-dependent anion-selective channel protein 2 isoform 2	P07288	32	91.0	19.7	11.3	8.3
Arachidonate 15-lipoxygenase B isoform D	P45880	76	90.7	89.3	27.3	7.7
Heat shock 70 protein 4L	O15296	95	88.7	43.3	28.3	17.0
Bifunctional aminoacyl-tRNA synthetase	O95757	171	87.3	65.0	45.3	19.7
Histone H1t	P07814	22	86.0	37.3	8.7	3.7
Peroxiredoxin-1	P22492	22	85.0	42.7	34.3	17.0
T-complex protein 1 subunit alpha isoform A	Q06830	60	83.7	77.3	37.7	15.7
Transmembrane emp24 domain-containing protein 9 precursor	P17987	27	82.0	51.7	23.0	19.0
Endoplasmic reticulum resident protein 44 precursor	Q9BVK6	47	81.3	50.7	41.7	26.7
Importin subunit alpha-2	Q9BS26	58	80.3	56.7	45.0	35.0
Nuclear pore complex protein Nup93 isoform 1	P52292	93	78.7	65.7	31.7	22.0
General vesicular transport factor p115	Q8N1F7	108	75.0	14.3	4.7	1.0
Polyadenylate-binding protein 1	P11940	71	72.0	5.3	0.0	0.0
Creatine kinase B-type	P12277	43	71.0	51.0	6.0	0.0
Elongation factor 1-delta isoform 1	P29692	71	69.7	26.3	6.3	5.7
Mesencephalic astrocyte-derived neurotrophic factor precursor	P55145	21	69.7	34.0	26.0	8.3
Reticulocalbin-2 precursor	Q14257	37	69.7	26.0	0.0	0.0
Peroxisomal membrane protein 11B isoform 1	O96011	28	67.7	23.0	10.7	4.3
Calmodulin	P62158	17	66.7	52.7	41.3	38.3
Heat shock protein beta-1	P04792	23	66.0	41.7	4.3	1.0
14-3-3 protein theta	P27348	28	64.3	55.7	18.0	5.3
26S proteasome non-ATPase regulatory subunit 13 isoform 1	Q9UNM6	43	64.3	48.3	19.7	5.0
Peptidyl-prolyl cis-trans isomerase B precursor	P23284	24	63.0	28.3	20.3	9.0
Annexin A4	P09525	36	62.0	33.3	25.0	15.3
Protein ERGIC-53 precursor	P49257	58	61.7	29.0	7.0	1.3
60S ribosomal protein L12	P30050	18	60.3	23.3	13.0	11.3
Exportin-2	P55060	110	60.0	39.7	16.0	11.3
cAMP-dependent protein kinase type I-alpha regulatory subunit	P10644	43	58.3	37.7	20.7	8.0
Ras-related protein Rab-11B	Q15907	24	58.3	26.0	9.7	0.0
Testis-expressed protein 101 isoform 1	Q9BY14	29	58.3	43.0	28.3	20.7

Contd...

Supplementary Table 2: Contd...

Protein	*Uniprot number*	*MW*	*F1*	*F2*	*F3*	*F4*
		kDa	*Average SC*	*Average SC*	*Average SC*	*Average SC*
Ecto-ADP-ribosyltransferase 3 isoform A precursor	Q13508	44	58.0	48.7	46.0	12.3
Ras-related protein Rab-6A isoform A	P20340	24	56.3	18.0	1.3	0.0
T-complex protein 1 subunit zeta-2 isoform 1	Q92526	58	56.3	32.0	22.7	11.3
Voltage-dependent anion-selective channel protein 3 isoform 1	Q9Y277	31	56.0	10.0	5.7	0.0
CD177 antigen precursor	Q8N6Q3	46	55.7	35.3	5.3	0.0
Uncharacterized protein KIAA2013 precursor	Q8IYS2	69	54.3	34.0	33.3	23.7
Glutamate carboxypeptidase 2 isoform 1	Q04609	84	53.0	39.0	1.3	1.0
Sperm surface protein Sp17	Q15506	17	53.0	35.3	19.7	14.0
Endoplasmic reticulum resident protein 29 isoform 1 precursor	P30040	29	52.3	33.0	19.3	5.7
NADH-cytochrome b5 reductase 3 isoform 2	P00387	32	51.7	34.0	31.3	20.7
Protein sel-1 homolog 1 isoform 1 precursor	Q9UBV2	89	51.3	23.3	9.3	0.0
26S proteasome non-ATPase regulatory subunit 11	O00231	47	49.7	36.0	8.0	7.0
Annexin A2 isoform 2	P07355	39	48.3	24.3	13.3	2.0
DnaJ homolog subfamily B member 11 precursor	Q9UBS4	41	48.0	29.3	27.7	7.7
26S proteasome non-ATPase regulatory subunit 7	P51665	37	47.3	36.0	13.0	2.0
Phosphoglycerate mutase 2	P15259	29	47.3	29.3	24.7	20.0
Histone H2A type 1-A	Q96QV6	14	47.0	28.0	19.0	11.3
Receptor expression-enhancing protein 6	Q96HR9	21	45.3	31.0	20.7	12.0
Alpha-actinin-4	O43707	105	44.7	29.3	21.7	14.7
Melanoma inhibitory activity protein 3 precursor	Q5JRA6	214	44.3	10.7	7.0	1.3
Axonemal dynein light intermediate polypeptide 1	O14645	32	43.3	35.0	29.7	0.0
Chitinase domain-containing protein 1 isoform A	Q9BWS9	45	43.3	25.3	8.0	6.0
Ras-related protein Rab-1A isoform 1	P62820	23	42.7	12.0	5.3	3.3
Cytochrome c oxidase subunit 5B, mitochondrial precursor	P10606	14	42.3	22.7	20.0	15.7
60S ribosomal protein L4	P36578	48	41.3	4.0	0.0	0.0
Peroxiredoxin-2 isoform A	P32119	22	40.3	20.0	7.0	3.7
Annexin A1	P04083	39	39.3	27.3	19.7	14.0
Ribonuclease inhibitor	P13489	50	39.3	12.7	12.3	3.3
DnaJ homolog subfamily A member 2	O60884	46	38.3	29.7	17.7	13.0
60S acidic ribosomal protein P2	P05387	12	38.0	12.0	9.7	8.7
Prostate and testis expressed protein 1 precursor	Q8WXA2	14	37.7	29.7	24.7	22.7
60S ribosomal protein L6	Q02878	33	37.0	18.0	4.0	1.0
Lon protease homolog, mitochondrial	P36776	106	36.7	24.3	11.3	17.0
Stromal cell-derived factor 2-like protein 1 precursor	Q9HCN8	24	36.3	27.0	23.7	13.0
Proteasome subunit beta type-1	P20618	26	36.0	31.0	24.7	12.3
Myosin light polypeptide 6 isoform 1	P60660	17	35.0	21.3	0.0	0.0
Histone H2A-Bbd type 2/3	P0C5Z0	13	34.7	20.3	0.0	0.0
Voltage-dependent anion-selective channel protein 1	P21796	31	34.3	21.0	7.7	0.0
Serine/threonine-protein phosphatase 2A 65 regulatory subunit A alpha isoform	P30153	65	34.0	28.0	5.7	4.7
Alpha-centractin	P61163	43	33.7	26.3	10.0	0.0
40S ribosomal protein S9	P46781	23	33.3	12.3	4.0	0.0
Signal peptidase complex subunit 3	P61009	20	33.0	21.3	18.3	14.0
Importin subunit alpha-3	O00505	58	32.7	22.0	18.3	10.7
Calpain small subunit 1	P04632	28	32.3	20.3	1.0	0.0
26S proteasome non-ATPase regulatory subunit 8	P48556	40	31.7	24.7	5.7	4.3
Nucleobindin-2 precursor	P80303	50	31.3	21.0	18.7	0.0
Synaptogyrin-2	O43760	25	31.3	17.0	9.0	9.0
Cytoskeleton-associated protein 4	Q07065	66	30.7	17.7	0.0	0.0
Myosin regulatory light chain 12A	P19105	20	30.7	21.3	2.0	0.0
Ras-related protein Rab-7a	P51149	23	30.7	13.7	11.7	6.7
Syntaxin-12	Q86Y82	32	30.7	10.7	0.0	0.0
DnaJ homolog subfamily B member 1	P25685	38	30.0	12.3	10.3	10.0
Leucine zipper transcription factor-like protein 1	Q9NQ48	35	30.0	14.0	2.3	0.0
Peroxisomal membrane protein 11C	Q96HA9	27	30.0	19.3	2.7	0.0
Nicastrin precursor	Q92542	78	29.7	23.0	18.3	11.0

Contd...

Supplementary Table 2: Contd...

Protein	*Uniprot number*	*MW*	*F1*	*F2*	*F3*	*F4*
		kDa	*Average SC*	*Average SC*	*Average SC*	*Average SC*
Cytosolic nonspecific dipeptidase isoform 1	Q96KP4	53	29.3	19.3	5.0	0.0
40S ribosomal protein S8	P62241	24	29.0	4.3	2.0	0.0
Protein FAM162A	Q96A26	17	28.7	18.0	16.0	12.3
Ras-related protein Rab-18	Q9NP72	23	28.0	7.0	0.7	0.0
Erlin-1	O75477	39	27.7	21.7	11.7	8.7
26S protease regulatory subunit 10B	P62333	46	27.3	20.3	8.3	5.7
Eukaryotic translation initiation factor 3 subunit A	Q14152	167	27.0	15.0	2.0	0.0
Peptidyl-prolyl cis-trans isomerase A	P62937	18	27.0	4.7	0.0	0.0
Nucleophosmin isoform 1	P06748	33	26.7	10.0	0.0	0.0
Hsc70-interacting protein	P50502	41	26.3	24.7	16.3	7.3
Heat shock protein 75 , mitochondrial precursor	Q12931	80	26.0	17.3	0.0	0.0
Platelet-activating factor acetylhydrolase precursor	Q13093	50	25.0	16.7	0.0	0.0
Galectin-3-binding protein precursor	Q08380	65	24.7	12.7	11.7	5.0
Hemoglobin subunit beta	D9YZU5	16	24.7	13.0	8.3	7.3
40S ribosomal protein SA	P08865	33	24.0	2.7	0.0	0.0
Poly(rC)-binding protein 1	Q15365	37	24.0	3.7	0.0	0.0
Serine/threonine-protein phosphatase 2A catalytic subunit beta isoform	P62714	36	24.0	15.0	7.7	0.0
Probable inactive serine protease 37 isoform 1 precursor	A4D1T9	26	23.7	13.3	1.7	0.0
40S ribosomal protein S16	P62249	16	23.0	5.7	1.3	0.0
40S ribosomal protein S25	P62851	14	23.0	3.0	1.7	0.0
Cytoplasmic dynein 1 light intermediate chain 1	Q9Y6G9	57	23.0	17.0	0.0	0.0
Normal mucosa of esophagus-specific gene 1 protein	Q9C002	10	23.0	10.7	7.3	3.0
Sorbitol dehydrogenase	Q00796	38	23.0	17.0	12.7	0.0
26S protease regulatory subunit 6A	P17980	49	22.7	20.7	13.0	3.0
Alkyl dihydroxyacetone phosphate synthase, peroxisomal precursor	O00116	73	22.7	5.0	0.0	0.0
Mitochondrial import receptor subunit TOM22 homolog	Q9NS69	16	22.7	10.7	0.0	0.0
Anterior gradient protein 2 homolog precursor	O95994	20	22.3	6.7	0.7	0.0
Gamma-glutamyltranspeptidase 1 precursor	P19440	61	22.3	4.0	0.0	0.0
40S ribosomal protein S13	P62277	17	22.0	4.3	0.0	0.0
60S ribosomal protein L14	P50914	23	21.3	5.0	0.0	0.0
Dynactin subunit 2	Q13561	45	21.3	14.3	8.0	0.7
Prostate stem cell antigen preproprotein	D3DWI6	12	21.3	13.0	0.0	0.0
40S ribosomal protein S15a	P62244	15	21.0	10.3	6.7	6.0
60S acidic ribosomal protein P1 isoform 1	P84098	12	21.0	11.7	8.7	0.0
60S ribosomal protein L19	P84098	23	21.0	5.0	0.0	0.0
Arginyl-tRNA synthetase, cytoplasmic	P54136	75	21.0	9.7	4.3	0.7
Glutathione S-transferase P	P09211	23	21.0	11.3	6.3	0.0
GTP-binding nuclear protein Ran	P62826	24	21.0	6.3	4.3	0.0
60S ribosomal protein L13 isoform 1	P26373	24	20.7	2.3	0.0	0.0
Protein S100-A9	P06702	13	20.3	11.3	0.0	0.0
Talin-1	Q9Y490	270	20.3	19.3	0.0	0.0
Endoplasmic reticulum-Golgi intermediate compartment protein 1	Q969X5	33	20.0	6.7	0.0	0.0
F-actin-capping protein subunit alpha-1	P52907	33	20.0	11.3	5.7	0.0
60S ribosomal protein L15 isoform 1	P61313	24	19.7	3.7	0.0	0.0
Nucleoporin NUP53	Q8NFH5	35	19.3	6.0	0.0	0.0
26S protease regulatory subunit 6B isoform 1	P43686	47	19.0	13.7	10.3	3.7
Lipoprotein lipase precursor	P06858	53	19.0	18.3	0.0	0.0
Perilipin-3 isoform 1	O60664	47	19.0	4.3	0.0	0.0
Gastricsin isoform 1 preproprotein	P20142	42	18.7	22.3	11.0	7.7
ERO1-like protein beta precursor	Q86YB8	54	18.3	14.7	0.0	0.0
Transmembrane emp24 domain-containing protein 1 precursor	Q13445	25	18.3	11.7	5.3	3.0
Glutamate dehydrogenase 1, mitochondrial precursor	P00367	61	17.7	6.0	0.7	0.0
Importin subunit alpha-4	O00629	58	17.7	17.3	9.0	5.3
NADH dehydrogenase (ubiquinone) 1 alpha subcomplex subunit 10, mitochondrial precursor	O95299	41	17.3	9.0	1.3	1.3
Peptidyl-prolyl cis-trans isomerase FKBP2 precursor	P26885	16	17.3	10.0	0.0	0.0

Contd...

Supplementary Table 2: Contd...

Protein	*Uniprot number*	*MW*	*F1*	*F2*	*F3*	*F4*
		kDa	*Average SC*	*Average SC*	*Average SC*	*Average SC*
Protein S100-A8	P05109	11	17.3	9.7	0.0	0.0
Stress-induced-phosphoprotein 1	P31948	63	17.3	11.0	0.0	0.0
Transmembrane emp24 domain-containing protein 5 isoform 1 precursor	Q9Y3A6	26	16.7	10.3	0.0	0.0
60S ribosomal protein L27a	P46776	17	16.3	7.7	0.0	0.0
Isocitrate dehydrogenase [NADP] cytoplasmic	O75874	47	16.3	9.0	2.7	0.0
Lysosome membrane protein 2 isoform 1 precursor	Q14108	54	16.0	6.7	0.0	0.0
Mitochondria-eating protein	Q8TC71	61	16.0	10.7	7.3	6.0
NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 8, mitochondrial precursor	O95169	22	16.0	7.0	0.0	0.0
40S ribosomal protein S14	P62263	16	15.7	2.3	0.0	0.0
60S ribosomal protein L11 isoform 1	P62913	20	15.7	7.3	0.0	0.0
Protein canopy homolog 2 isoform 1 precursor	Q9Y2B0	21	15.3	7.3	4.0	0.0
Translocon-associated protein subunit alpha precursor	P43307	32	15.3	12.0	8.0	7.7
60S ribosomal protein L22 proprotein	P35268	15	15.0	8.3	4.0	1.3
Cysteinyl-tRNA synthetase, cytoplasmic isoform C	P49589	95	15.0	2.7	0.0	0.0
Abhydrolase domain-containing protein 16A isoform A	O95870	63	14.7	11.3	0.0	0.0
COP9 signalosome complex subunit 8 isoform 2	Q99627	18	14.7	5.0	0.0	0.0
Golgi apparatus protein 1 isoform 2 precursor	Q92896	136	14.7	14.0	7.7	2.3
S-phase kinase-associated protein 1 isoform B	P63208	19	14.7	7.3	0.0	0.0
Transcription elongation factor B polypeptide 1 isoform A	Q15369	12	14.7	5.7	3.7	1.3
26S proteasome non-ATPase regulatory subunit 5	Q16401	56	14.0	5.7	0.0	0.0
14-3-3 protein sigma	P31947	28	13.7	10.7	0.0	0.0
Coatomer subunit zeta-1	P61923	20	13.7	4.3	0.0	0.0
Eukaryotic translation initiation factor 3 subunit E	P60228	52	13.7	3.3	1.3	0.0
Neutrophil gelatinase-associated lipocalin precursor	P80188	23	13.7	7.0	0.0	0.0
Aspartyl-tRNA synthetase, cytoplasmic	P14868	57	13.3	12.0	5.3	1.0
Chloride intracellular channel protein 1	O00299	27	13.3	5.7	3.7	0.0
Programmed cell death protein 6	O75340	22	13.3	11.3	8.7	3.3
Signal recognition particle receptor subunit beta	Q9Y5M8	30	13.3	5.0	0.0	0.0
14-3-3 protein beta/alpha	P31946	28	13.0	6.3	0.0	0.0
Cathepsin D preproprotein	P07339	45	13.0	6.0	0.0	0.0
Phosphoglycerate kinase 1	P00558	45	13.0	12.0	8.7	0.0
Protein-tyrosine phosphatase mitochondrial 1 isoform 1	Q8WUK0	23	12.7	7.7	0.0	0.0
Ras-related protein Ral-A precursor	P11233	24	12.7	8.3	3.7	0.0
Low molecular weight phosphotyrosine protein phosphatase isoform C	P24666	18	12.3	3.3	0.0	0.0
NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 4 isoform 1	O95168	15	12.3	9.0	0.0	0.0
ATP synthase-coupling factor 6, mitochondrial isoform A precursor	P18859	13	12.0	7.0	7.0	3.3
CDGSH iron-sulfur domain-containing protein 2	Q8N5K1	15	12.0	6.0	0.0	0.0
HD domain-containing protein 2	Q7Z4H3	23	12.0	6.7	0.0	0.0
Ubiquitin carboxyl-terminal hydrolase isozyme L3	P15374	26	12.0	3.3	2.3	0.0
15 selenoprotein isoform 1 precursor	O60613	18	11.7	7.7	0.7	0.0
40S ribosomal protein S6	P62753	29	11.7	4.3	0.0	0.0
Coiled-coil-helix-coiled-coil-helix domain-containing protein 3, mitochondrial precursor	Q9NX63	26	11.7	6.0	5.3	4.3
Protein FAM166A	Q6J272	36	11.7	6.0	0.0	0.0
Serine/threonine-protein phosphatase PGAM5, mitochondrial isoform 1	Q96HS1	32	11.7	8.0	0.0	0.0
DnaJ homolog subfamily A member 1	P31689	45	11.3	4.3	0.0	0.0
Eukaryotic translation elongation factor 1 epsilon-1 isoform 1	O43324	20	11.3	8.0	2.7	2.7
Protein S100-A11	P31949	12	11.3	8.7	2.3	0.0
Histone H1.4	P10412	22	10.7	5.0	0.0	0.0
Peroxisomal biogenesis factor 16 isoform 1	Q9Y5Y5	39	10.7	1.7	0.0	0.0
Selenoprotein T precursor	P62341	22	10.7	4.7	0.0	1.0
ATP synthase subunit e, mitochondrial	P56385	8	10.3	8.3	3.0	0.0
Torsin-1A-interacting protein 1	Q5JTV8	66	10.0	5.7	0.0	0.0
DnaJ homolog subfamily C member 3 precursor	Q13217	58	9.7	8.0	0.7	0.0
Myoferlin isoform B	Q9NZM1	233	9.7	2.7	0.0	0.0
NADH dehydrogenase [ubiquinone] iron-sulfur protein 5	O43920	13	9.7	9.0	1.0	2.0

Contd...

Supplementary Table 2: Contd...

Protein	*Uniprot number*	*MW*	*F1*	*F2*	*F3*	*F4*
		kDa	*Average SC*	*Average SC*	*Average SC*	*Average SC*
Up-regulated during skeletal muscle growth protein 5	Q96IX5	6	9.7	6.0	6.0	4.7
Neutrophil defensin 1 preproprotein	P59665	10	9.3	6.7	3.0	0.0
Protein FAM3C precursor	Q92520	25	9.3	3.3	0.0	0.0
Cytochrome b-c1 complex subunit 9 isoform A	Q9UDW1	7	9.0	4.7	0.0	0.0
Eukaryotic translation initiation factor 3 subunit F	O00303	38	8.7	5.7	0.0	0.0
Sorcin isoform B	P30626	20	8.7	4.7	0.0	0.0
Thioredoxin isoform 1	P10599	12	8.7	4.7	0.0	0.0
Apolipoprotein A-I preproprotein	P02647	31	8.3	5.3	3.7	0.0
Calcium and integrin-binding protein 1	Q99828	22	8.3	2.7	0.0	0.0
Sperm-associated antigen 4 protein	Q9NPE6	48	8.3	7.3	5.7	3.3
D-3-phosphoglycerate dehydrogenase	O43175	57	8.0	3.7	0.0	0.0
Metalloproteinase inhibitor 1 precursor	P01033	23	8.0	3.0	0.0	0.0
Eukaryotic translation initiation factor 3 subunit M	Q7L2H7	43	7.7	6.3	0.0	0.0
Glutaredoxin-related protein 5, mitochondrial precursor	Q86SX6	17	7.7	3.0	5.0	2.0
Guanine nucleotide-binding protein G (I)/G (S)/G (T) subunit beta-1	P62873	37	7.7	3.3	0.0	0.0
UPF0733 protein C2orf88	Q9BSF0	11	7.7	3.0	0.0	0.0
Phosphoglycolate phosphatase	A6NDG6	34	7.3	2.7	0.0	0.0
Prostate- and testis-expressed protein 4	P0C8F1	11	7.3	4.7	0.0	0.0
Synaptogyrin-4	O95473	26	7.3	6.7	5.7	0.0
Macrophage migration inhibitory factor	P14174	12	7.0	4.0	1.3	0.0
LYR motif-containing protein 4 isoform 1	Q9HD34	11	6.7	3.0	0.0	0.0
Uncharacterized protein C13orf16	Q8N6K0	17	6.7	3.3	1.7	1.3
26S proteasome non-ATPase regulatory subunit 4	P55036	41	6.3	4.7	0.0	0.0
thioredoxin domain-containing protein 2 isoform 2	Q86VQ3	60	6.3	4.0	0.0	0.0
Serine/threonine-protein phosphatase 4 regulatory subunit 1 isoform A	Q8TF05	107	6.0	4.3	0.0	0.0
WD repeat-containing protein 61	Q9GZS3	34	5.7	5.0	0.0	0.0
DDB1- and CUL4-associated factor 7	P61962	39	5.3	3.0	0.0	0.0
Translin-associated protein X	Q99598	33	5.0	3.3	0.7	1.0
Protein SEC13 homolog isoform 2	P55735	34	3.7	2.3	1.3	0.0

F1: fraction 1; F2: fraction 2; F3: fraction 3; F4: fraction 4; SC: spectral count; HSP: heat shock protein; MW: molecular weight

Supplementary Table 3: Functional annotations using Database for Annotation, Visualization and Integrated Discovery for proteins expressed from immature through spermatozoa maturation process (F1 through F4)

Sample dataset	*Key pathways*	*Key processes*	*Enriched functional categories*	*Cellular location*	*Functional categories that are associated with majority of proteins*	*Top TFBS*	*Key functions*	*Summary highlights of key processes/ functions/pathways affected*
F1	Ribosome (21), protein export (3), SNARE interactions in vesicular transport (4), glutathione metabolism (4), influenza infection (27), UTR-mediated translational regulation (21), Met of proteins (27), GX (29), diabetes pathway (12), membrane trafficking (4)	Translational elongation (22), translation (27), intracellular transport (25), Golgi vesicle transport (10), cellular macromolecular complex assembly (15), protein localization (24), oxidation reduction (18), regulation of translation (8), cell death (19), regulation of apoptosis (17)	Translational elongation (22), ribosome (21), protein biosynthesis (21), RNA binding (28), structural molecule activity (26), ER (26), IC transport (925), Golgi vesicle transport (10), ER-Golgi transport (6), vesicle-mediated transport (15), macromolecular complex assembly (18), protein complex biogenesis (11), SRP (3), protein localization in organelle (8), NT binding (44), methylation (11), RasGTPase (6), lipoprotein (11), generation of precursor metabolites and energy (6) cell redox homeostasis (7), cell death (19), regulation of apoptosis (17), sexual reproduction (9), spermatogenesis (7), multicellular organism reproduction (8), spermatid development/differentiation (3) microtubule cytoskeleton (17), cytoskeletal part (21), mitochondrion (26)	Ribosomal subunit (20), ribonucleo-protein complex (34), cytosol (53), nonmembrane-bounded organelle (63), ER (33), organelle envelope (22), microtubule cytoskeleton (17), mitochondrion (26), Golgi apparatus (19), intracellular organelle lumen (33), nucleolus (15)	Acetylation (102), translational elongation (22), ribonucleo-protein (28), ribosome (21), protein biosynthesis (22); cytosol (53), translation (27), ER (26)	Pax4 (143), AML1 (140), YY1 (129)	Structural constituent of ribosome (21), RNA binding (28), GTP binding (17), GTPase activity (12), nucleotide binding (44), HSP binding (5), purine ribonucleotide binding (35)	Translation elongation, protein transport, oxidoreductase activity, reproductive process, spermatid development/ differentiation, regulation of apoptosis
F2	Telomere maintenance (3), systemic lupus erythematosus (4), pyruvate metabolism (3), val-leu-ilu degradation (3)	Epidermis development (8), ectoderm development (8), oxidation–reduction (10), epithelial cell differentiation (5), protein transport (9), protein localization (9), intermediate filament organization (2), sexual reproduction (6)	Intermediate filament (8), keratin (8), protein transport (9), protein localization (9), membrane-bounded vesicle (8), sexual reproduction (6), reproductive process in a multicellular organism (6), EF-Hand2 domain (4), nucleosome assembly (3), chromatin (3), gonad development (3), sex differentiation (3), actin filament binding (3), ER membrane (3), proteolysis involved in cellular protein catabolic process (3), regulation of apoptosis (3)	Keratin filament (7), IF (8), mitochondrion (16), cytoskeletal part (13), membrane-bounded vesicle (8), intracellular nonmembrane-bounded organelle (21), ER (10)	Acetylation (23), intracellular nonmembrane-bounded organelle (21), disease mutation (18), mitochondrion (16), cytoskeleton (15), coiled coil (15), structural molecular activity (12), transit peptide (10), oxidoreductase (10), ER (10), sexual reproduction (6)	SP1 (24), ROAZ (38)	Structural constituent of cytoskeleton (6), Structural molecular activity (12), coenzyme binding (5), actin filament binding (3), cofactor binding (5)	Epidermis and ectoderm development, cell differentiation, protein–protein interactions, protein transport and localization, oxidoreductase activity, gamete generation, gonad development, proteolysis

Contd...

Supplementary Table 3: Contd...

Sample dataset	*Key pathways*	*Key processes*	*Enriched functional categories*	*Cellular location*	*Functional categories that are associated with majority of proteins*	*Top TFBS*	*Key functions*	*Summary highlights of key processes/ functions/pathways affected*
F3	Huntington's disease (6), pyrimidine metabolism (4), oxidative phosphorylation (4), ubiquitin-mediated proteolysis (4), metabolism of protein import into nucleus (2), metabolism of vitamins and cofactors (3), integration of energy metabolism (5), metabolism of nucleotides (3)	Nucleoside metabolic process (4), generation of precursor metabolites and energy (7), oxidation–reduction (9), vitamin metabolic process (3)	Mitochondrial inner membrane (9), oxidoreductase activity (3), oxidative phosphorylation (3), protein ubiquitination (3), protein modification by small protein conjugation (3), phospholipid metabolic process (3), spermatogenesis (4), male gamete generation (4), nucleotide biosynthetic process (3), lysosome (3), lytic vacuole (3), ubl conjugation pathway (5), proteolysis involved in cellular protein catabolic process (5), GTP binding (3), ER membrane (3), mitotic cell cycle (3)	Organelle inner membrane (10), organelle envelope (13), mitochondrion (12), respiratory chain (3), nuclear pore (3)	Alternative splicing (44), acetylation (22), cytoplasm (22), transport (15), secreted (14), organelle envelope (13), mitochondrion (12), oxidoreductase (9), generation of precursor metabolites and energy (7)		Cytochrome-c oxidase activity (3), transmembrane transporter activity (4), nucleotidyl transferase activity (4)	Generation of precursor metabolites and integration of energy metabolism, oxidative phosphorylation, protein catabolic process, protein ubiquitination
F4	Hedgehog signaling pathway (2), metabolism of amino acids (3), nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2 (2)	Reproductive process in a multicellular organism (5), spermatogenesis (4), male gamete generation (4), sexual reproduction (4), phosphorus metabolic process (5), regulation of protein complex disassembly (2) heterocyclic biosynthetic process (2)	Multicellular organism reproduction (5), spermatogenesis (4), male gamete generation (4), sexual reproduction (4), transit peptide (5), mitochondrion (5), phosphate metabolic process (5), protein kinase activity (3), secreted (5), signal (8), cell surface (4), membrane (14), cation binding (8), cytoskeleton (3)	Cell surface (4)	Acetylation (11), hydrolase (8), multicellular organism reproduction (5), mitochondrion (5), phosphorus metabolic process (5), transit peptide (5), spermatogenesis (4), lipid catabolic process (3), hedgehog signaling pathway (2), regulation of protein complex	HAND1E47 (22), ATF (16), CDPCR3 (24); Domains - serine/ threonine protein kinases active site signatures (3)	Oxidoreductase activity, acting on sulfur group of donors (2)	Multicellular organism reproduction
Increasing trend	Metabolism of amino acids (4), Huntington disease (3), valine-leucine-isoleucine degradation (4), propionate metabolism (3), aldosterone-regulated sodium reabsorption (3), beta-alanine metabolism (2)	Carboxylic acid catabolic process (5), oxidation–reduction (8), sexual reproduction (6), spermatogenesis (4), male gamete generation (4), integrin-mediated signaling pathway (3), generation of precursor metabolites and energy (5), amine catabolic process (4), sperm motility (2), cell motility (4)	Dynein (4), microtubule motor activity (4), purine nucleotide metabolic process (4), cilium axoneme (3), sodium/ potassium transport (3), metallopeptidase activity (5), potassium ion binding (4), ion transport (4), cell motility (4), male gamete generation (4), spermatogenesis (4), multicellular organism reproduction (4), protein complex assembly (3)	Mitochondrion (17), dynein complex (4), cytoskeleton (12), cell projection (7), cilium axoneme (3), microtubule-based flagellum (3)	Alternative splicing (34), cytoplasm (19), mitochondrion (17), acetylation (15), mitochondrion (15), nucleotide binding (15), coiled coil (13), cytoskeleton (12), transit peptide (10), sexual reproduction (6), generation of precursor metabolites and energy (5)	YY1 (51), CREBP1CJUN (13), CDPCR1 (30), AML1 (54), HSF1 (24)	Sodium: potassium-exchanging ATPase activity (3), microtubule motor activity (4), nucleotide binding 915), purine nucleoside binding (11), metallo-endopeptidase activity (3), purine nucleotide binding (12), coenzyme binding (5), potassium ion binding (4), magnesium ion binding (6)	Intracellular transport, oxidation reduction, cellular amino acid catabolic process, alternative splicing

Contd...

Supplementary Table 3: Contd...

Sample dataset	*Key pathways*	*Key processes*	*Enriched functional categories*	*Cellular location*	*Functional categories that are associated with majority of proteins*	*Top TFBS*	*Key functions*	*Summary highlights of key processes/ functions/pathways affected*
Decreasing trend	Metabolism of proteins (38), UTR-mediated translational regulation (25), signaling by Wnt (13), apoptosis (13), integration of energy metabolism (16), cell cycle (19), metabolism of amino acids (12), diabetes pathways (22), regulation of activated PAK-2p34 by proteasome-mediated degradation (10), ubiquitin proteasome pathway (9), antigen processing and presentation (9), proteasome (9), ribosome (20), oxidative phosphorylation 99), role of Ran in mitotic spindle regulation (3), Prion pathway (3), mechanism of gene regulation by peroxisome proliferators via PPARa (4)	Protein folding (34), translational elongation (25), cell redox homeostasis (12), proteasomal protein catabolic process (13), protein transport (35), protein localization (36), regulation of ligase activity (11), proteolysis (28), regulation of apoptosis (27), homeostatic process (25), protein complex biogenesis (21), cell cycle (21), sexual reproduction (15)	Protein folding (34), translational elongation (25), protein biosynthesis (27), HSP70 (6), proteasome (10), stress response (12), response to unfolded protein (16), ER (38), mitochondrion (36), negative regulation of protein ubiquitination (10), regulation of ligase activity (11), oxidoreductase (17), intracellular protein transport (20), secreted (28), vesicle-mediated transport (15), cellular protein localization (20), EF-hand type domain (13), peroxidase activity (5), regulation of cell death (27), spermatogenesis (12), male gamete generation (12), sexual reproduction (15), spermatid development (3), spermatid differentiation (3), sperm cell development (3)	Cytosol (83), nonmembrane-bounded organelle (65), mitochondrion (36), melanosome (19), ribosomal subunit (19), ER (46), ribonucleo-protein complex (30), vesicle (33)	Phosphoprotein (151), acetylation (147), cytoplasm (97), cytosol (83), nonmembrane-bounded organelle (65), signal (60), nucleotide binding (53), ER (46), unfolded protein binding (31), protein biosynthesis (29), proteolysis (28), regulation of apoptosis (27), translational elongation (25), homeostatic process (25), protein transport (24), mitochondrion (23), protein complex assembly (21), cell cycle (21), oxidoreductase (17)	NFY (123)	Unfolded protein binding (31), structural constituent of ribosome (20), structural molecule activity (26), nucleotide binding (60), calcium ion binding (24), peptidase activity (15), ATPase activity (12), protein transporter activity (9), GTPase activity (11), peptide binding (8), antioxidant activity (6)	Spermatogenesis, protein metabolism, cell cycle, integration of energy metabolism, regulation of apoptosis, cell redox homeostasis, translational elongation, response to protein folding

The number of proteins is in the parenthesis. ER: endoplasmic reticulum; HSP: heat shock protein; TFBS: transcription factor binding sites