Towards the identification of reliable sperm biomarkers for male infertility: A sperm proteomic approach

P. Intasqui1,2 | A. Agarwal1 | R. Sharma1 | L. Samanta3 | R. P. Bertolla2

1American Center for Reproductive Medicine, Cleveland Clinic, Cleveland, OH, USA
2Department of Surgery, Division of Urology, Human Reproduction Section, São Paulo Federal University – São Paulo Hospital, São Paulo, Brazil
3Department of Zoology, Ravenshaw University, Cuttack, India

Correspondence
Ashok Agarwal, American Center for Reproductive Medicine, Cleveland Clinic, Cleveland, OH, USA.
Email: agarwaa@ccf.org

Funding information
This work was conducted with financial support from the American Center for Reproductive Medicine, Cleveland Clinic, USA, and the São Paulo Research Foundation, Brazil (FAPESP, grant# 2015/18867-6)

Summary
Male infertility evaluation is mainly based on semen analysis. Thus, identification of additional diagnostic methods is valuable. The aim of this study was to analyse the sperm proteome of infertile men to identify the underlying mechanisms and reliable diagnostic biomarkers. This cross-sectional study consisted of 16 infertile men and seven proven fertile men. An LC-MS/MS approach was performed in five pooled samples of each group (proven fertile men, primary infertility and secondary infertility). Differentially expressed proteins were used for functional enrichment analyses, and the most central proteins involved in altered functions in both infertile groups and the testis-specific proteins were validated using Western blotting and immunocytochemistry. In total, 1,305 sperm proteins were identified, of which 102 were underexpressed and 15 were overexpressed proteins in both infertile groups. Underexpressed proteins were mostly related to protein post-translational modification and folding, especially BAG6, HSPA2 and SPA17. Validation analysis revealed an underexpression of BAG6 in infertile men, whereas HSPA2 and SPA17 expressions did not differ between the groups. No differences were observed in the sperm localisation of these proteins. An overexpression of HIST1H2BA—a testis-specific protein—was observed in both proteomic approaches. Therefore, BAG6 and HIST1H2BA are potential candidates for male infertility biomarkers.

KEYWORDS
biomarkers, infertility, male, proteomics, spermatozoa, validation studies

1 | INTRODUCTION

Infertility affects approximately 15% of reproductive age couples, and male factors are responsible in approximately 50% of cases (Brugh & Lipshultz, 2004). Male infertility can be attributed to a myriad of causes ranging from varicocele (1992), obesity (Jensen et al., 2004), genetic factors (Shamsi, Kumar, & Dada, 2011), lifestyle factors (i.e., smoking; Hassan & Killick, 2004) and spinal cord injury (Monga, Bernie, & Rajasekaran, 1999). Currently, clinical evaluation of male infertility is mainly based on semen analysis, which, although being important to assess the male fertility potential, does not provide a direct prediction of reproductive success. Therefore, infertile men can present without any major alterations in semen variables (Abu, Franken, Hoffman, & Henkel, 2012; Shamsi et al., 2010).

Due to these limitations, additional methods are required to evaluate male fertility. Thus, efforts have been made towards understanding and identifying altered molecules that might be utilised during infertility management (Kovac, Pastuszak, & Lamb, 2013). In this context, the study of the sperm proteome is of great interest because sperm proteins play a key role in the maintenance of sperm morphology, motility patterns, acrosome formation and reaction, capacitation, and fertilisation (Ashrafzadeh, Karsani, & Nathan, 2013). Hence, altered sperm protein expression in male infertility may be further used as noninvasive diagnostic, prognostic and/or therapeutic tools of male
infertility (Kovac et al., 2013). Moreover, it may provide important insights into sperm physiology and its underlying mechanisms (Rahman, Lee, Kwon, & Pang, 2013).

It has been hypothesised that the sperm protein profile in infertile men is altered, and some researchers have assessed these profiles in several male infertility conditions including varicocele (Agarwal, Sharma, Samanta, Durairajanayagam, & Sabanegh, 2016; Agarwal, Sharma, Durairajanayagam, Ayaz, et al., 2015; Agarwal, Sharma, Durairajanayagam, et al., 2016; Hosseinifar et al., 2013), obesity (Kriegel et al., 2009; Liu et al., 2015), smoking (Chen et al., 2015), failed fertilisation after assisted reproduction techniques (ART; Frapsauce et al., 2014; Legare et al., 2014; Pixton et al., 2004), asthenozoospermia (Amaral et al., 2014; Hashemitabar, Sabbagh, Orazizadeh, Ghadiri, & Bahmanzadeh, 2015; Liu et al., 2015; Martinez-Heredia, de Mateo, Vidal-Taboada, Ballesca, & Oliva, 2008; Parte et al., 2012; Shen, Wang, Liang, & He, 2013; Siva et al., 2010; Zhao et al., 2007) and idiopathic infertility (Legare et al., 2014; McMenemy et al., 2014; Xu et al., 2012). Sperm protein profiles have also been evaluated in men with semen oxidative stress (Ayaz et al., 2015; Sharma et al., 2013) and sperm DNA fragmentation (Intasqui et al., 2013; de Mateo et al., 2007), which are common causes of male infertility. Additionally, the sperm proteome was also compared in samples from infertile men with those of men with proven fertility, although infertility was described only as oligozoospermia (Thacker et al., 2011).

To the best of our knowledge, no study has analysed the sperm proteome of men with primary or secondary infertility, regardless of the cause of infertility, and compared the results with those of proven fertile men. Such an analysis could help us better understand the underlying molecular mechanisms of male infertility that are common to different causes of infertility as well as provide sperm biomarkers of infertility that would be applicable to most infertile men.

Therefore, the aim of this study was to compare the sperm proteome between men with primary or secondary infertility, irrespective of the cause, with that of proven fertile men to identify altered molecular pathways and potential sperm biomarkers of male infertility. For this, we utilised a shotgun proteomic approach to identify the differentially expressed sperm proteins, and we validated the results using Western blotting and immunocytochemistry in order to assess protein expression and localisation respectively.

2 | MATERIAL AND METHODS

2.1 | Ethical approval

This study was approved by the Institutional Review Board from the Cleveland Clinic. All included subjects provided written informed consent.

2.2 | Study design

A cross-sectional observational study was conducted using infertile patients, who were referred to the Andrology Laboratory of the American Center for Reproductive Medicine (ACRM), Cleveland Clinic, USA, and men with proven fertility, who were enrolled by the Center. Infertile men were divided into primary and secondary infertility groups for the shotgun proteomic study. For the proven fertile group, only normozoospermic men who had fathered a child in the past year were included. Infertile men were aged 20–50 years, presented with infertility of at least one-year duration, irrespective of the cause and with a round cell concentration lower than $1 \times 10^6$/ml. Exclusion factors were azoospermia, history of cancer and cancer treatment, previous surgery of the male reproductive tract, erectile dysfunction, cryptorchidism or use of exogenous testosterone. Men were also excluded if their female partner had been diagnosed with a condition associated with female infertility, such as endometriosis, advanced age or irregular menses.

No differences between the groups were observed regarding age, body mass index (BMI), semen volume, sperm concentration and motility, and round cells concentration. Infertility duration, described as median; inter-quartile range (1st quartile–3rd quartile), was 15; 27 (9–36) months for the primary infertility group, 18; 12 (12–24) months for the secondary infertility group ($p > .05$) and 12; 12 (12–24), if merging both groups. In the primary infertility group, 22.2% of men were smokers, 11.1% had varicocele, 44.4% were obese, and 77.8% had alteration in their semen analysis. In the secondary infertility group, 42.9% were smokers, 57.1% had obesity, and 71.4% presented with altered semen analysis.

Men were collected at the Andrology Laboratory by masturbation following 2–6 days of ejaculatory abstinence (World Health Organization, 2010). On average, samples in the fertile, primary infertility and secondary infertility groups had an ejaculate volume of 4.06, 4.95 and 3.21 ml respectively, which was then used for the shotgun proteomic analysis and the validation studies.

After liquefaction, a semen aliquot was utilised for conventional semen analysis, which was performed according to the WHO (2010) guidelines (World Health Organization, 2010). Another aliquot (200 μl) was centrifuged for 10 min at 500 g; the pellet was fixed in 4% (v/v) paraformaldehyde in phosphate-buffered saline (PBS, Irvine Scientific, CA, USA) for 20 min at room temperature, washed in PBS (3 × 5 min, 500 g), resuspended in 1 ml of PBS and kept at 4°C for protein localisation analysis by immunocytochemistry. The remaining semen volume was frozen without cryoprotectants and kept at ~80°C for shotgun proteomic analysis and Western blotting. All reagents used in this study were purchased from Thermo Fisher Scientific (MA, USA), Sigma-Aldrich (MO, USA) or Bio-Rad Laboratories (CA, USA), unless otherwise described.

Five pooled sperm samples from each group (control, primary infertility and secondary infertility) were subjected to shotgun proteomic analysis using one-dimensional electrophoresis followed by liquid chromatography coupled to tandem mass spectrometry (1D-LC-MS/MS). Results were used for bioinformatic analyses to identify the protein–protein interaction networks and the altered biological pathways. From the comparative proteomic analysis, key proteins of the top molecular pathway in both primary and secondary infertility groups were chosen for validation by Western blotting and immunocytochemistry, together with proteins responsible for altered sperm function. Because in a clinical setting, the identification of a general
biomarker of infertility—without the distinction of primary or secondary infertility—would be more valuable, and for increased sensitivity, the infertility groups were merged into one infertility group for the validation study.

2.3 | Shotgun proteomic analysis (1D-LC-MS/MS)

At the time of the analyses, whole ejaculate samples were thawed at 4°C and centrifuged at 500 g for 10 min at 4°C. After the supernatant seminal plasma was completely removed, the pellet was washed with PBS (3 × 10 min, 500 g, 4°C). For sperm protein extraction, radioimmunoprecipitation assay (RIPA) buffer (Sigma-Aldrich) supplemented with protease inhibitors (Complete, Mini Protease Inhibitor Cocktail Tablets, EDTA-free, Roche Applied Science, Mannheim, Penzberg, Germany) was then added to the washed pellet according to the sperm concentration. Samples were thoroughly homogenised and kept overnight at 4°C. After centrifugation at 13,000 g for 30 min at 4°C, an aliquot of the supernatant was utilised for protein estimation, and the remaining volume was stored at −20°C until proteomic and Western blotting analyses.

Protein estimation was performed using the bicinchoninic acid (BCA) protein assay kit (Thermo Fisher Scientific), using bovine serum albumin (BSA) as standard. Samples quantified with coefficients of variation over 5% were quantified again in another run.

For shotgun proteomic analysis, samples were pooled separately for each group (five samples per group) and each pool was run in triplicate (technical replicates) following the same protocol described in depth in (Agarwal, Sharma, Durairajanayagam, et al., 2016; Gokce, 2016; Agarwal, Sharma, Samanta, et al., 2016). Both spermatozoa and protein concentrations were normalised in each group prior to analysis, that is, equal quantity of protein was obtained by an equal number of spermatozoa. Briefly, each pool was mixed with sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) buffer, boiled for protein denaturation, and separated on one-dimensional (1D) 12.5% Tris–HCl gels (15 μg protein/lane) with a constant voltage of 150 V for 35 min. Prior to in-gel digestion, protein bands were reduced with dithiothreitol (DTT) and alkylated with iodoacetamide. Each gel lane was cut into six pieces and washed with milli-Q water, then dehydrated in acetonitrile.

Complete protein digestion was obtained with 5 μl of 10 ng/μl trypsin in 50 mM ammonium bicarbonate overnight at room temperature. Formed peptides were extracted from the gel and analysed by LC-MS/MS in a LTQ-Orbitrap Elite Hybrid Mass Spectrometer system (Thermo Fisher Scientific). The HPLC system was a Dionex with an Acclaim Pepmap C18 2 μm and 100 Å reversed-phase capillary chromatography column. Tandem MS data were obtained in a data-dependent method. Briefly, a full-profile MS scan at 60,000 resolution (at 400 m/z) between 300 and 2,000 m/z was performed on the Orbitrap FT instrument using a precursor isolation window of 3 m/z at 35% relative collision energy. The charge state screening was not enabled, and the dynamic exclusion option was enabled after three repeated acquisitions within a 20-s duration with each exclusion lasting 90 s. Mass spectra were then extracted by the Proteome Discoverer (version 1.4.1.288, Thermo Fisher Scientific). Charge state deconvolution and de-isotoping were not performed.

For protein identification, all MS/MS data were analysed using Mascot (version 2.3.02, Matrix Science, London, UK), SEQUEST (version 1.4.0.288, Thermo Fisher Scientific) and X! Tandem (version CYCLONE, The GPM) search programs. Mascot, Sequest and X! Tandem were set up to search the human reference (33,292 entries) assuming trypsin as the digestion enzyme. These searches were performed with a fragment ion mass tolerance of 0.8 Da and a parent ion tolerance of 10 parts per million (PPM). Carbamidomethylation of cysteine was specified as a fixed modification, and oxidation of methionine was specified as a variable modification.

For validation of MS/MS-based peptide and protein identifications, these search results were then uploaded into the Scaffold 4.0.6.1 program (Proteome Software Inc., OR, USA). Only a >99% probability to achieve a false detection rate (FDR) < 1.0% was accepted as well as at least two identified peptides. Proteins were annotated with gene ontology (GO) terms from the National Center for Biotechnology Information (NCBI, downloaded 21 October 2013).

Protein quantification was performed by comparing the normalised spectral counts for these samples. Appropriate filters were used to identify the differentially expressed proteins (DEPs) between each comparison (primary infertility × proven fertile; secondary infertility × proven fertile; primary infertility × secondary infertility). DEPs were dependent on the overall abundance of the proteins among the three replicate runs (Agarwal, Sharma, Samanta, et al., 2016; Gokce, Shuford, Franck, Dean, & Muddiman, 2011).

2.4 | Bioinformatic analysis and selection of proteins for validation

Protein–protein interaction networks, functional annotation and functional enrichment analysis of DEPs were performed using publicly available bioinformatic annotation tools and databases such as GO Term Finder; GO Term Mapper; UniProt; Software for Researching Annotations of Proteins (STRAP); Database for Annotation, Visualization, and Integrated Discovery (DAVID); and proprietary software packages such as Ingenuity Pathway Analysis (IPA) and Search Tool for the Retrieval of Interacting Genes/Proteins (STRING).

Thereafter, the major enriched functions related to the underexpressed and overexpressed proteins in the primary and secondary infertility groups were identified. The proteins related to these functions that were differentially expressed in both infertility groups and with the highest number of interactions in the protein–protein network were then detected and, through an in-depth literature search of these proteins, the proteins more commonly related to infertility were selected for validation: large proline-rich protein BAG6 (BAG6), heat shock-related 70-kDa protein 2 (HSPA2) and sperm surface protein Sp17 (SPA17). Moreover, testis-specific proteins differentially expressed in all infertility groups were also validated, given that these proteins may reflect sperm maturation: SPA17 and histone H2B type 1-A (HIST1H2BA). These selected proteins may play the most important role in the observed infertility.
2.5 Protein validation

To identify potential biomarkers of male infertility that could be applied to most of the infertile men, irrespective of the cause or type of the infertility, primary and secondary infertility groups were merged into one infertility group for the validation study. Decision to merge both primary and secondary infertility groups as one group was due to (i) a relatively small variation in protein expression between both infertility groups, (ii) the possibility of increasing the number of individuals, thus decreasing the likelihood of incurring in type II error and (iii) the common basis of the clinical phenotype—both groups were infertile (with diverse underlying causes for infertility between different patients in the same group). Although we understand this increases the chance of an alpha error, we maintained stringency at the % cut-off for the alpha error while decreasing the odds of incurring in a beta error. Therefore, for protein expression and localisation analyses, proven fertile men were compared to infertile men. With this approach, sensitivity is increased, albeit at the cost of reducing specificity.

2.5.1 Protein expression analysis

Western blotting was performed in individual and pooled samples depending on the protein concentration (samples with low-protein concentration were pooled). For the proven fertile group, seven individual samples (n = 7) were utilised, whereas for the infertile group, seven individual samples were utilised and nine samples were pooled into four different pools (final n = 11), assuring biological variability between the pools (pool 1 with two different samples, pool 2 with three samples and pools 3 and 4 with two samples each).

Western blotting was carried out as previously described by Agarwal, Sharma, Samanta et al. (2016). Briefly, 20 µg of proteins of each pool or individual sample was added to a final volume of 30 µl of RIPA buffer and Laemmli buffer (4× Laemmli sample buffer—Bio-Rad with 10% [v/v] 2-mercaptoethanol—Sigma-Aldrich). Samples were boiled for 5 min and loaded into 10%-20% Tris-Tricine polyacrylamide gels (Mini-PROTEAN, Bio-Rad). SDS–PAGE was performed in the Bio-Rad system in a running buffer (Bio-Rad) consisting of tris-buffered saline Tween-20, Thermo Fisher Scientific). After washing three times with TBST (3 × 10 min), the membranes were incubated with transfer buffer with 25 mM Tris–HCl (Sigma-Aldrich), 192 mM glycine (Sigma-Aldrich) and 20% (v/v) methanol (Sigma-Aldrich). Blocking was performed in SuperBlock Blocking Buffer in PBS (Thermo Fisher Scientific), for 1 hr at room temperature. Primary antibodies were incubated overnight at 4°C, diluted in 1% (w/v) BSA (Sigma-Aldrich) in tris-buffered saline Tween-20 (TBST) buffer (25 mM Tris, 0.15 mM NaCl, 0.05% Tween-20, Thermo Fisher Scientific). After washing three times with TBST (3 × 10 min), the membranes were incubated with horseradish peroxidase-conjugated secondary antibody diluted in 1% (w/v) BSA in TBST, for 1 hr at room temperature. Antibody dilutions for each protein were as follows: (i) BAG6—0.5 µg/ml anti-BAG6 (ab88292, Abcam, Cambridge, UK) and 1:2,000 (v/v) anti mouse (sc-2005, Santa Cruz Biotechnology, TX, USA); (ii) HSPA2—1:5,000 (v/v) anti-HSPA2 (ab154374, Abcam) and 1:10,000 (v/v) anti-rabbit (ab97051, Abcam); (iii) 1:1,000 (v/v) anti-SPA17 (ab172626, Abcam) and 1:20,000 (v/v) anti-rabbit; and (iv) HIST1H2BA—1:1,000 (v/v) anti-HIST1H2BA (ab178426, Abcam) and 1:20,000 (v/v) anti-rabbit. Secondary antibodies were washed three times for 40 min with TBST at room temperature, and peroxidase activity was revealed using enhanced chemiluminescence (Pierce ECL Western blotting substrate, Thermo Fisher Scientific) as a substrate. Chemiluminescence was detected in the ChemiDoc Touch Imaging System (Bio-Rad). Densitometric analysis and quantification of the bands were done using the Image Lab 5.2.1 Software (Bio-Rad), and each band intensity was normalised by the housekeeping protein (α-Tubulin), in order to facilitate a better comparison between different membranes and decrease intra-assay variability. Data are shown as the densitometric value of the protein divided by the value of α-Tubulin (arbitrary units).

2.5.2 Immunocytochemistry analysis

For the immunocytochemistry, a new and independent cohort of samples was collected. A concentration of 2 × 10⁶ fixed spermatozoa/ml from three proven fertile men and two infertile men was placed on the coverslips (Deckglaser cover glasses, 12 mm, Marienfeld, Germany) previously coated with 0.01% (v/v) poly-L-lysine (Sigma-Aldrich) and allowed to dry at room temperature overnight. Cells were permeabilised with 0.2% (v/v in PBS) Triton X-100 (Bio-Rad) for 20 min, washed four times with 0.05 mM glycine in PBS and blocked with 1% (w/v) BSA in PBS for 1 hr, both at room temperature. Samples were then incubated with primary antibodies at 4°C overnight, in the following dilutions: (i) BAG6—10 µg/ml; (ii) HSPA2 (ab9130, Abcam)—20 µg/ml; and (iii) SPA17—1:100 (v/v). Coverslips were then washed four times and incubated with Alexa Fluor 488-conjugated secondary antibodies for 1.5 hr at 37°C. For BAG6 and HSPA2, an anti-mouse antibody was used (A-11001, Thermo Fisher Scientific), whereas for SPA17, an anti-rabbit antibody (ab150077, Abcam) was used, both in a 1:1,000 dilution. Coverslips were then washed again and incubated with 1 µg/ml 7-aminoactinomycin D (7-AAD, Thermo Fisher Scientific) in DMSO/PBS for 30 min at 37°C. After washing three times, coverslips were added to clean slides and mounted with mounting medium containing 0.01 µg/ml Hoechst 33258 (Sigma-Aldrich) in water and equal volumes of PBS and glycerol (Sigma-Aldrich). Cells were visualised by confocal microscopy using a Leica DM6000 CFS upright microscope and LAS AF software v2.7.3 (Leica Microsystems, GmbH, Wetzlar, Germany). Approximately 150 cells per group were analysed, and the percentage of cells showing protein localisation in the head or tail was counted for each protein.
2.6 Statistical analyses

Statistical analyses were performed in the PASW (SPSS) 18.0 software for Windows. For the quantitative data from shotgun proteomic analysis, the numerical values used in the quantitation correspond to the normalised spectral counts (nSC, SC/ΣSC). In order for the proteins to be considered significant, the nSC ratio should be greater than 2.0 or less than 0.5 with a $p$-value < .05 using $t$-test. Each set of data was tested for normality (Shapiro–Wilk test) and homogeneity of variance (Levene’s test). Data were compared between the groups using a Student’s $t$ test to identify the DEPs.

Semen analysis and Western blotting (protein/$\alpha$-tubulin) data were compared between controls and infertile men by the nonparametric Mann–Whitney test. In this case, data are presented as median; interquartile range (Q1–Q3). Furthermore, proteins presenting a statistically significant difference in the Western blotting analysis were used in a logistic regression analysis. These results were used to construct a receiver operating characteristics (ROC) curve, in order to assess specificity, sensitivity and the area under the curve.

For immunocytochemistry data, a Mann–Whitney test was also performed to compare the percentages of each protein localisation between the groups, and the results are presented as a boxplot. In all cases, a $p$ value < .05 was considered statistically significant.

3 RESULTS

Using a shotgun proteomic approach, we were able to identify a total of 1,305 sperm proteins among the primary and secondary infertility and the proven fertile groups. Venn diagrams with the DEPs in three different comparisons (proven fertile × primary infertility; proven fertile × secondary infertility; proven fertile × primary and secondary infertility), as well as the protein–protein interaction network of the DEPs related to altered biological functions, are shown in Figure 1. A total of 117 DEPs were differentially expressed in the infertile men (both primary and secondary infertility taken together) in comparison with fertile counterparts. Furthermore, 356 proteins were differentially expressed when comparing primary infertility to secondary infertility (data not shown), whereas 1,188 were common to all three groups. DEPs in each comparison are further described in Table S1.

Of all the identified networks by Ingenuity Pathway Analysis (IPA), the one involved in protein folding and post-translational modifications (PTM) was the most prominent one as all the 35 focus proteins were underexpressed in the infertile groups of our data set (Figure 1). From this network, the testis-enriched chaperone HSPA2 and its co-chaperone BAG6, important hubs in the protein–protein interaction network, were chosen for validation based on previous reports on their involvement in sperm–oocyte recognition (Bromfield, Aitken, & Nixon, 2015). Furthermore, important proteins implicated in sperm dysfunction, that is, SPA17 and HIST1H2BA, were also validated, the latter being overexpressed in these groups in the shotgun proteomic analysis.

Protein validation data (expression and localisation) demonstrated that, as expected, BAG6 was significantly underexpressed in sperm samples from infertile men (3.9-fold decrease, compared to proven fertile men, $p = .003$, Figure 2a), whereas HIST1H2BA was overexpressed in these men (1.8-fold increase, $p = .001$, Figure 3c).

SPA17 was identified with two different molecular masses—17 and 20 kDa (Lea, Richardson, Widgren, & O’Rand, 1996; Figure 4a). However, the expression of both SPA17 forms and also of HSPA2 did not differ between the groups ($p > .05$, Figures 3a and 4a). BAG6 and HSPA2 were expressed throughout the spermatozoon, both in the head and in tail (Figures 2b and 3b). SPA17, although also expressed along the spermatozoa, had its highest expression observed in the sperm tail, particularly in the tail principal piece (Figure 4b).
No differences were seen in the sperm localisation of these proteins ($p > .05$, Figures 2–4).

Finally, two logistic regression models were constructed, one for BAG6 and one for HIST1H2BA. For BAG6, a significant model was achieved ($p = .0002$), with a positive predictive value of 100%, a negative predictive value of 85.7% and a total predictive value of 93.8%. For HIST1H2BA, a significant model was achieved ($p = .0003$), with a positive predictive value (for detecting the “Infertile” group) of 90.9%, a negative predictive value of 100% and a total predictive value of 94.4%. A ROC curve was constructed for each model, with an
area under the curve value of 92.1% ($p = .005$) for BAG6 and of 93.5% ($p = .004$) for HIST1H2BA (Figure 5). Sensitivity and specificity values were, respectively, 100% and 85.7% for BAG6 and 88.9% and 100% for HIST1H2BA.

4 | DISCUSSION

Despite the high prevalence of infertility, and, particularly, of male infertility, little is known about its underlying molecular mechanisms. Furthermore, male infertility evaluation is based upon semen analysis, which is unable to provide a reliable diagnosis of infertility. Therefore, the aim of this study was to compare the sperm proteome between proven fertile men and infertile men (both with primary or secondary infertility), focusing on detecting altered post-genomic pathways and identifying potential biomarkers that could be used during male infertility management. Using a shotgun proteomic approach, we were able to identify a total of 1,305 sperm proteins. Of these, 117 proteins were differentially expressed in both infertile groups, which were mostly involved with protein PTM and folding functions, according to our bioinformatic analyses.

Post-translational modifications are especially important in mature spermatozoa, in which the nucleus is transcriptionally and translationally inert. Therefore, protein PTM or the acquisition of new proteins during epididymal transit is essential for regulation of sperm maturation and function (Baker, 2016; Baker et al., 2012). The most studied sperm PTMs so far are protein phosphorylation, glycosylation, methylation, acetylation and ubiquitination (Baker, 2016). Protein phosphorylation, especially tyrosine phosphorylation, for instance, is important to the process of sperm capacitation (Bajpai & Doncel, 2003; Battistone et al., 2014; Ficarro et al., 2003; Visconti et al., 1995). Similarly, a correct folding of the tertiary protein structure is essential for protein activity and to avoid protein degradation and the accumulation of misfolded proteins, which are involved with several diseases (Valastyan & Lindquist, 2014). Thus, the overall underexpression of sperm proteins related to PTM and folding in men with primary and secondary infertility may indicate an alteration in these processes in male infertility, but more studies are needed to confirm this finding. Similar results were also observed in infertile men with varicocele (Agarwal, Sharma, Durairajanayagam, Ayaz, et al., 2015; Agarwal, Sharma, Durairajanayagam, Cui, et al., 2015), asthenozoospermia.
in human, HSPA2 and BAG6 have been implicated in the alteration related to male infertility. Given the interaction of this protein with HSPA2, we also suggest that this underexpression may be involved with chromatin alterations leading to male infertility (Maselli et al., 2012).

On the other hand, SPA17 expression (both of 17 and 20 kDa) and localisation were not statistically different between the groups.

SPA17 is a low molecular weight, mannose-binding protein that binds to zona pellucida carbohydrates and is therefore an important protein for a number of processes including sperm capacitation, acrosome reaction, recognition and binding to the oocyte zona pellucida, and fertilisation (Chiriva-Internati et al., 2009; O’Rand, Widgren, & Fisher, 1988; Richardson, Yamasaki, & O’Rand, 1994; Wen, Richardson, & O’Rand, 1999). In the human testis, a weak expression of this protein was observed in the cytoplasm of spermatozoa, whereas a strong expression was detected in early and late spermatids, which suggests that SPA17 plays a role in the differentiation process. Furthermore, 95% of ejaculated spermatozoa also expressed SPA17, which was observed using light microscopy in the tail principal piece (Chiriva-Internati et al., 2009; Grizzi et al., 2003). Under electron microscopy, it was demonstrated that this expression was restricted to the sperm fibrous sheath (Chiriva-Internati et al., 2009). Another study, however, detected SPA17 expression in the sperm head and tail, mostly in the principal piece (Lea, Widgren, & O’Rand, 2004). In agreement with this finding, we also observed the presence of SPA17 in the principal piece of the sperm tail and in the head of about 5% of spermatozoa. SPA17 localisation in the sperm principal piece appears to be important for cell signalling and motility, as this protein may regulate protein kinase A anchoring protein (AKAP) activity (Carr et al., 2001; Chiriva-Internati et al., 2009; Colledge & Scott, 1999). Therefore, we can suggest that infertile men, both with primary or with secondary infertility, do not present with an alteration in SPA17 expression and localisation in spermatozoa.

Using BAG6 and HIST1H2BA, the proteins presenting a statistical difference between infertile and fertile groups, in separate logistic regression models demonstrated that these proteins present high specificities and sensitivities as sperm biomarkers of male infertility. Therefore, these proteins should be further studied to evaluate their role in male infertility, as well as to assess their real potential as diagnostic methods using targeted proteomic approaches and confirmatory studies. If confirmed, the evaluation of these proteins expression could be added to semen analysis as an additional diagnostic method of male infertility.

The most important limitation of this study was the small sample size and the pooling of samples. In the Western blotting analysis,
because sperm concentration in some samples was low and, thus, samples had to be pooled, we could not detect all inter-individual differences within the groups. To address this issue, we ran individual samples as well. Additionally, we only had a limited number of samples from infertile men with varicocele. In the immunocytochemistry analysis, the small sample size may have impaired our ability to identify minor differences. Finally, in the shotgun proteomic analysis, several proteins were differentially expressed between the groups, and each one of them could have been utilised for validation. We chose the proposed proteins based on their involvement with protein turnover and sperm function, which suggests that these proteins could be more important for infertility than the others. However, it is highly likely that we may still have missed some other important proteins that could be validated, such as proteins involved in proteasomal and lysosomal pathways. The validated proteins, although with altered expression in the infertile groups, cannot be related to a specific defect because of the heterogeneity of the infertility groups.

5 | CONCLUSION

In conclusion, the sperm proteome of infertile men with primary or secondary infertility differs from that of proven fertile men. PTMs and protein folding are enriched by the underexpressed proteins in the primary and secondary infertility groups, suggesting that they may play a role in male infertility. After confirmatory studies, we suggest that BAG6 and HIST1H2BA proteins may be potential biomarkers of male infertility, irrespective of its cause.

ACKNOWLEDGEMENTS

The authors acknowledge Ms. Banu Gopalan and Vathsalya Senapathi for the bioinformatic annotation analysis on this dataset and Mr. John Peterson and Ms. Gauravi Deshpande for the confocal analysis.

CONFLICT OF INTEREST

None.

ORCID

P. Intasqui http://orcid.org/0000-0002-8494-6009
A. Agarwal http://orcid.org/0000-0003-0585-1026
R. Sharma http://orcid.org/0000-0002-9664-6978
L. Samanta http://orcid.org/0000-0002-2969-0071
R. P. Bertolla http://orcid.org/0000-0001-6892-9244

REFERENCES


SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.