

**Proteomic profile of maternal aged blastocoel fluid suggests a novel role for ubiquitin system in blastocyst quality**

Gabriella Tedeschi<sup>1,2,\*</sup> , Elena Albani<sup>3,\*</sup> , Elena Monica Borroni<sup>4</sup> , Valentina Parini<sup>3</sup> , Anna Maria Brucculeri<sup>3</sup> , Elisa Maffioli<sup>2</sup> , Armando Negri<sup>1</sup> , Simona Nonnis<sup>1</sup> , Mauro Maccarrone<sup>5</sup> , Paolo Emanuele Levi-Setti<sup>3</sup>

1 Dipartimento di Scienze Veterinarie e Sanità Pubblica, Università degli Studi di Milano, Milan, Italy.

2 Fondazione Filarete, 20139, Milan, Italy.

3 Humanitas Fertility Center, Department of Gynecology, Division of Gynecology and Reproductive Medicine, Humanitas Research Hospital, Via Manzoni 56, Rozzano, 20089 Milan, Italy.

4 Humanitas Research Hospital, Via Manzoni 56, Rozzano, 20089 Milan, Italy.

5 Department of Medicine, Campus Bio-Medico University of Rome, Via Alvaro del Portillo 21, 00128, Rome, Italy.

\* These authors equally contributed to the work

## **Abstract**

**Purpose** The etiology of maternal ageing, a common cause of female factor infertility and a rate-limiting step in vitro fertilization (IVF) success, remains still unclear. Proteomic changes responsible for the impaired successful pregnancy outcome after IVF with aged blastocysts have not been yet evaluated. The objective of this prospective study was to employ proteomic techniques and bioinformatic tools to enlight differences at the protein level in blastocoel fluid of aged and younger woman.

**Methods** Protein composition of human blastocoel fluid isolated by micromanipulation from 46 blastocysts of women aged <37 years (group A) and 29 of women aged  $\geq 37$  years (group B) have been identified by a shotgun proteomic approach based on high resolution nano-liquid chromatography electrospray-ionization-tandem mass spectrometry (nLC-ESI-MS/MS) using label free for the relative quantification of their expression levels.

**Results** The proteomic analysis leads to the identification and quantification of 148 proteins; 132 and 116 proteins were identified in group A and B, respectively. Interestingly, the identified proteins are mainly involved in processes aimed at fine tuning embryo implantation and development. Among the 100 proteins commonly expressed in both groups, 17 proteins are up-regulated and 44 down-regulated in group B compared to group A. Overall, the analysis identified 33 proteins, which were increased or present only in B while 76 were decreased in B or present only in A.

**Conclusions** Data revealed that maternal ageing mainly affects blastocyst survival and implantation through unbalancing the equilibrium of the ubiquitin system, known to play a crucial role in fine-tuning several aspects required to ensure successful pregnancy outcome.

## **Keywords**

Blastocoel fluid, ageing, embryo implantation, proteomics, shotgun

## **Capsule**

Maternal ageing mainly affects blastocyst survival and implantation through unbalancing the equilibrium of ubiquitin system-

## Introduction

Embryo implantation requires synchronous development of the blastocyst and the endometrium [1]. During the early stages of implantation, the blastocyst enters the uterine cavity apposes and adheres to the endometrial epithelium to initiate implantation. It is well established that abnormalities in the bidirectional embryo-endometrial dialogue are one of the major cause of infertility resulting in implantation failure [2, 3]. Noteworthy, advancing maternal age represents one of the major causes of the altered embryo-endometrial communication [4]. As expected, several large studies have reported age-specific outcomes of IVF, and maternal age has clearly been associated with a decline in cumulative live birth rates in patients older than 37 [5, 6]. Unfortunately, the molecular changes that occur in the endometrium during the “window of receptivity” and the blastocyst contribution have been poorly investigated, due to the difficulty and the ethical limitations in studying implantation in humans; thus, no clinically useful markers have been identified to assess embryo quality and positive pregnancy outcome [7, 8].

In the recent years, the development of a wealth of “omic” technologies has opened the way for identifying potential markers, although validation of these is still a major issue [9, 10]. In particular, proteomics is an emerging powerful technology in the identification of both biomarkers and disease targets, allowing for the detection and monitoring of differing conditions or treatments. Although the classic proteomics approach has had limited success in the field of embryology [11, 12], the new developments in mass spectrometry have been helpful and the sensitivity of this approach has enabled the development of new protocols that are capable of profiling the proteome of small groups of mouse embryos [13], and even of single human blastocysts [14-16]. Therefore, several studies have recently addressed the analysis of the proteome of the human preimplantation embryo, in particular the ability of the blastocyst to secrete and/or consume different proteins in surrounding medium (secretome), in order to communicate with the maternal endometrium and proceed with the implantation process [13-15, 17, 18]. Interestingly, degenerating embryos presented a different protein expression profile compared to that of developing blastocysts, with a significant alteration of numerous biomarkers. Besides surrounding medium, the cells of the blastocyst are exposed to the fluid of the blastocoel, an internal cavity that becomes the yolk sac of the developing embryo [19]. Thus, the blastocoel fluid is the natural environment of blastocyst growth and the assessment of its protein composition provides an opportunity to expand the knowledge of embryonic physiology, including the maternal-embryonic dialog at the time of implantation. From a clinical perspective, the analysis of blastocoel fluid is emerging as a crucial topic in the field of reproductive medicine aimed at identifying new molecular markers related to highest implantation potential. The ultimate goal is to improve blastocyst culture conditions and fertilization rate, as well as to develop an affordable and cost-effective non-invasive biochemical assay to diagnose blastocyst viability, thus facilitating the move to single embryo transfer. The

major problem concerning the analysis of the blastocoel fluid relies on its limited bioavailability (1-8 nl) and, despite the fact that recent developments in mass spectrometry allow the identification and characterization of thousands of proteins from low microgram levels of protein extracted from tissues or cell cultures [17, 16, 20, 21], the low amounts of proteins extractable from blastocoel fluid is still a limiting step in this study.

Here, by a shotgun proteomic approach and label free quantification, we report for the first time the blastocoel fluid proteins that are differentially expressed in 29 collected surplus aged and 46 non-aged blastocysts, taking 37 year old women as the cut off parameter to categorized human blastocysts. The differences in the protein profile of blastocoel fluid of aged versus non-aged blastocysts relies on the identification of 109 differentially expressed proteins, which are mainly involved in processes aimed at fine tuning embryo implantation and development.. Upon validation, the suggested key proteins may serve as potential biomarkers of embryo quality and may contribute towards the effectiveness of protocols for blastocyst culture management.

## **Materials and methods**

### *Patient population*

Written informed consent was obtained from each patient and the study was approved by our local Ethics Committee. All consenting patients undergoing fertility treatment at the Humanitas Fertility Center, Humanitas Research Hospital, Rozzano, Italy, with supernumerary blastocysts were included in the study irrespective of the woman's age (range years: 26-41), infertility diagnosis (female or male), ovarian stimulation protocol (use of GnRH agonist or antagonist protocols in combination with recombinant and/or highly purified urinary gonadotropins), or performance of standard IVF or intracytoplasmic sperm injection (ICSI). The study was part of a trial on blastocoel fluid collection in order to test the safety of this procedure on pregnancy outcome [22].

### *Blastocyst generation*

Oocytes, sperm and embryos were cultured in the SAGE sequential culture media system (SAGE, Pasadena, USA) at 37°C in 5% CO<sub>2</sub>, 5%O<sub>2</sub> in nitrogen. Oocytes retrieval was performed 36 h after the administration of recombinant human chorionic gonadotropin through transvaginal ultrasonography. Oocytes were inseminated using either standard IVF insemination or ICSI; after ICSI, individual oocytes were transferred to a single 25 µl droplets under oil in Cleavage Medium supplemented with 10% Serum Protein Substitute (SAGE, Pasadena, USA), while standard IVF-inseminated oocytes were fertilized in 35 ul Fertilization Medium droplets under oil and transferred on day 1 to the Cleavage Medium. Fertilization results were assessed at 16–20 hours post-sperm insemination. On day 3, and in some

cases on day 2, 2/3 embryos were selected for transfer to each patient in the study; the embryos in excess of the number appropriate for transfer were cultured in single 25 µl droplets under oil in the Blastocyst Medium up to day 5 or day 6, and were cryopreserved at the blastocyst stage. Human blastocysts were scored up 120 or 144 hours after insemination according to the expansion of the blastocoel cavity and number and integrity of both the inner cell mass (ICM) and trophoctoderm (TE).

#### *Isolation of the blastocoel fluid and sample collection*

Blastocoel fluids from 75 blastocysts obtained from 52 IVF patients were collected. The blastocoel fluid was isolated by micromanipulation using an Olympus X40 (200x) microscope with Narishige manipulators and pipettes (angle: 35°, inner diameter: 5.5 µm) from Cook (Brisbane, Australia). A droplet of 5-10 µl Blastocyst Medium and a droplet of 3 µl water were prepared in an ICSI dish (Falcon-Corning brand, NY, USA) and covered with oil. The blastocysts were transferred to the droplet with culture medium and the dish was placed in the micromanipulator. The microinjection needle was loaded with water before sampling. The blastocyst was immobilized by the holding pipette with the ICM placed toward the holding pipette. The microinjection needle was then gently led in to blastocyst cavity, the fluid was aspirated until the blastocyst collapsed and the blastocoel fluid was transferred to the water droplet. After micromanipulation the samples were transferred to a labeled PCR tube using a 10 µl pipette and immediately frozen at -80°C. To obtain adequate protein material for analysis, the samples were collected and analyzed in two pools according on the patient's age: group A (blastocoel fluid from 46 samples of women aged <37 years) and group B (29 samples of women aged ≥ to 37 years). Both groups were homogeneous for infertility factor, for dose of gonadotropin administrated for ovarian stimulation and for average of oocytes retrieved after induction.

#### *Proteomic analysis*

The proteins of the blastocoel fluid were lysed in a buffer containing 8 M urea, 0.1 M ammonium bicarbonate and the complete protease inhibitor cocktail (Roche, Basel, Switzerland). After reduction and derivatization, proteins were digested with sequence grade trypsin for 16 hours at 37°C using a protein:trypsin ratio of 1:20. The proteolytic digest were desalted using Zip-Tip C<sub>18</sub> (Millipore) before mass spectrometric (MS) analysis [23, 24]. LC-ESI-MS/MS analysis was performed on a Dionex UltiMate 3000 HPLC System with a PicoFrit ProteoPrep C18 column (200 mm, internal diameter of 75 µm) (New Objective, USA). Gradient: 1% ACN in 0.1 % formic acid for 10 min, 1-4 % ACN in 0.1% formic acid for 6 min, 4-30% ACN in 0.1% formic acid for 147 min and 30-50 % ACN in 0.1% formic for 3 min at a flow rate of 0.3 µl/min. The eluate was electrosprayed into an LTQ Orbitrap Velos (Thermo Fisher Scientific, Bremen,

Germany) through a Proxeon nanoelectrospray ion source. The LTQ-Orbitrap was operated in positive mode in data-dependent acquisition mode to automatically alternate between a full scan ( $m/z$  350-2000) in the Orbitrap (at resolution 60000, AGC target 1000000) and subsequent CID MS/MS in the linear ion trap of the 20 most intense peaks from full scan (normalized collision energy of 35%, 10 ms activation). Isolation window: 3 Da, unassigned charge states: rejected, charge state 1: rejected, charge states 2+, 3+, 4+: not rejected; dynamic exclusion enabled (60 s, exclusion list size: 200) [25]. Data acquisition was controlled by Xcalibur 2.0 and Tune 2.4 software (Thermo Fisher Scientific, MA USA). A sample containing only culture medium, kept as a control (CTR), was also treated and analysed.

### *Data analysis*

Mass spectra were analyzed using MaxQuant software (version 1.3.0.5). The initial maximum allowed mass deviation was set to 20 ppm for monoisotopic precursor ions and 0.5 Da for MS/MS peaks. Enzyme specificity was set to trypsin, defined as C-terminal to arginine and lysine excluding proline, and a maximum of two missed cleavages were allowed. Carbamidomethylcysteine was set as a fixed modification, N-terminal acetylation, methionine oxidation, and asparagine/glutamine deamidation as variable modifications. The spectra were searched by the Andromeda search engine against the human Uniprot sequence database (release 22.01.2014). The reversed sequences of the target database were used as decoy database. Protein identification required at least one unique or razor peptide per protein group. Quantification in MaxQuant was performed using the built in XIC-based label free quantification (LFQ) algorithm [26] using fast LFQ. False protein identifications (1% or 5%) [17] were estimated by searching MSMS spectra against the corresponding reversed-sequence (decoy) database. The minimum required peptide length was set to 9 amino acids. Statistical analyses were performed using the Perseus software (version 1.4.0.6). Five technical replicates were carried out for each group (A and B) and for the control. Only proteins present and quantified in at least 3 out of 5 technical repeats and not detected in the sample containing only the culture medium (CTR) were positively identified as blastocoel fluid components and used for statistical analyses. Many keratins known to be due to a contamination during sample handling, were excluded from the analysis. The intensity of each protein (measured in at least 3 replicates in group A and group B) in each replicate are used for the t-test. Proteins were considered differentially expressed if they were present only in women aged <37 years (group A) or women aged  $\geq$ 37 years (group B) or showed significant t-test difference (filtered for permutation-based False Discovery Rate of many independent hypotheses, cut-off at 5%) between the two groups. The proteins differently expressed in the two groups were clustered according to their functions using the DAVID platform [27] filtered for significant Gene Ontology terms: Biological Process (GOBP) and Molecular Function (GOMF), using a p value <0.05. Protein-protein interactions were analysed by

String [28]. Gene products previously experimentally detected at the protein level as secreted and/or in extracellular region and/or exosome and/or microvesicles were determined according to public available databases (www.neXtProt.org, www.exocarta.org, microvesicle.org, geneontology.org, www.proteinatlas.org), and dataset [29]. Gene products previously detected at the transcription level in blastocysts or embryonic tissue are listed Tables 1-4 and S2- according to microarray datasets and gene expression databases (<http://genomewidepdb.proteomix.org> or <https://www.ncbi.nlm.nih.gov/geoprofiles>) [30, 16]. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD003535

## Results

### *Proteomic profile of human blastocoel fluid isolated from women aged < and ≥ 37 years*

The aim of the study was to examine the protein composition of human blastocoel fluid isolated from blastocysts of women aged <37 years (group A) and women aged ≥ 37 years (group B) to enlighten differences, at the protein level, that can possibly explain blastocyst survival and implantation success known to be very different in the two groups. Blastocoel fluid was isolated by micromanipulation, and proteins were identified by a shotgun proteomic approach, using label free for the relative quantification of their expression levels. Proteins identified in the control (containing only culture medium) and keratins- likely due to a contamination during sample preparation-were discarded from the analysis (Supplementary Table 1). The proteomic analysis leads to the identification and quantification of 148 proteins (Fig. 1); 132 and 116 proteins were identified in group A (Supplementary Table 2, average CV 1.2%) and B (Supplementary Table 3, average CV 1.3%), respectively (Fig. 1a). Among the 100 proteins commonly expressed in both groups, 17 proteins are up-regulated and 44 down-regulated in group B compared to group A (Fig. 1b). These proteins are listed in Tables 1 and 2, respectively, while Tables 3 and 4 report the proteins expressed only in A (32) or in B (16), respectively. Overall, the analysis identified 33 proteins, which were increased or present only in B while 76 were decreased in B or present only in A (Fig. 1 As reported in Tables 1-4, Supplementary Table 2 and 3, 98% of the differentially regulated proteins and 91 % of the total proteins identified have been previously experimentally detected either at the protein level in blastocoel fluid, exosomes, microvesicles or extracellular region or at the gene expression level in microarray studies concerning blastocyst transcriptome and gene expression databases.

Up to now only two papers addressed the issue of the protein content in the blastocoel fluid by a proteomic analysis: Jansen et al. [17] described the protein content of the blastocoel together with the medium and the cells, while Poli et al. [16], very recently, reported on the proteome of the blastocoel fluid combining the results from shot gun

proteomics and gene expression analysis. The present study allows to identify 108 proteins never described before, adding unprecedented information on this important human body fluid.

#### *Classification of the differentially expressed proteins based on bioinformatic analysis*

The proteins differentially expressed in the two groups were classified based on Gene Ontology Molecular Function and Biological Processes, as shown in Fig. 2. As a further step in the bioinformatic analysis, the proteins were manually grouped according to their function taking into account, besides GOMF and GOBP, also keywords and known annotations from the literature. It was thus possible to ascribe each protein to 6 major groups: 1) fertilization and embryo implantation, 2) embryo development, 3) nucleic acid interaction, 4) signaling, 5) uncharacterized proteins, 6) proteins excluded from the groups listed above categorized as “other” (Fig. 3, Supplementary Table 4). Interestingly, comparing the analysis of proteins increased or only expressed in A (Fig. 3a) with the proteins increased or only expressed in B (Fig. 3b), the major differences can be observed in 2 categories: fertilization and embryo implantation (12% group A vs 2% group B) and nucleic acid interaction (4% group A vs 13% group B).

#### **Discussion**

According to the promising results obtained by Jensen and collaborators and by Poli and coworkers [16, 17], to better delineate possible mechanisms for the effect of maternal ageing on fertility and potential biomarkers for evaluating the blastocysts quality and improving pregnancy outcome rate in IVF patients, we examined the proteomic profile of blastocoel fluid in woman <37 year old (group A) or ≥37 year old (group B). Comparison between the two matched study groups demonstrated, as already known, that pregnancy rate was lower in group B. In this group 4/12 (33,3%) pregnancies were obtained after frozen- thawed blastocyst transfer cycles versus group A (13/30 pregnancies, 43,3%). The shotgun proteomic approach allowed identifying 148 proteins, out of which 108 are described for the first time in the blastocoel fluid, and 109 are differentially expressed between the two groups. These proteins were functionally profiled using a suite of bioinformatics annotation tools and pathway databases, to obtain a comprehensive view of the impacted processes, pathways, functions, and networks related to maternal ageing. The bioinformatic analysis showed that differential proteins participate in various biological processes, but the main perturbed networks refer to fertilization and embryo implantation, as well as to nucleic acid interactions (Fig. 3 and Table S4).

Interestingly, many of the proteins differentially expressed are related to processes involving ubiquitin, that is largely known to play a crucial role in controlling several aspects required to ensure successful pregnancy outcome. In particular, the major molecular functions of these proteins consist of histone turnover and transcription factors, crucially



required for cell cycle progression and proliferation. Therefore, it is presumed that alterations in their expression level could affect implantation potential by altering blastocyst homeostasis towards apoptosis and autophagy settings. The following paragraph details the known biological role in reproductive processes of ubiquitin-related proteins which have been identified as blastocoel fluid components in the present study and listed in Tables 1-4 and S2-S3. In addition, a network showing the known interactions amongst up- or down-regulated ubiquitin-related proteins is shown in Fig.4.

In the recent years, increasing observations have demonstrated the role of ubiquitin system in reproductive processes, such as gametogenesis, modulation of steroid receptor concentrations, placental development and endometrial modification at the beginning of pregnancy [31]. These widerange effects have led to extensive research and several studies recently demonstrated the increase of ubiquitylation in the uterus during embryo implantation; thus, the proper control and turnover of this key signaling protein activity suggest potential roles in controlling critical aspects of implantation [32-35]. This is also emphasized by the observation that ubiquitin is accumulated in trophoblast of expanding blastocysts [36], suggesting that dysregulation of ubiquitin system may be responsible for failure IVF outcome. In 2009 Katz-Jaffe and collaborators reviewed findings on potential protein biomarkers in the embryo secretome [18], identifying a crucial role for ubiquitin. In our study, a huge number of proteins related to ubiquitin system has been identified and differentially regulated among the two groups analyzed. The ubiquitinase UBC and Ubiquitin Carboxyl-terminal Hydrolase isozyme L1 (UCHL1) are down-regulated in aged blastocysts, as well as proteins containing a RING finger domain that plays a key role in the ubiquitination pathway during embryonic development [37, 38] such as RING Finger protein 223 (RNF223) and Kelch-like protein 17 (KLHL17). Instead, the expression of the 26S Proteasome non-ATPase regulatory subunit 4 (PSMD4), a component of the 26S proteasome complex responsible for ubiquitylated proteins degradation, and the AN1-type zinc finger protein 6 (ZFAND6), a polyubiquitin binding protein involved in regulating NF-kappaB signaling [39], are completely switched off. Interestingly, several differential expressed proteins are substrate of Small Ubiquitin MOdifier (SUMO), which is also involved in regulating embryonic viability in mammals [40]. In particular, aged blastocysts lack Ran GTPaseactivating protein 1 (RANGAP1), that regulates pathways mainly involved in differentiation, apoptosis and cell cycle by altering protein function through changes in activity or cellular localization or by protecting substrates from ubiquitination [41]. Instead, they show increased expression of BEN domain-containing protein 4 that belongs to a family of proteins involved in histone deubiquitination. Accumulated evidence in the literature has currently demonstrated that the monoubiquitylation of histone H2A at lysine 119 is crucially associated with transcriptionally repressed chromatin, as a result of the repression activity of Polycomb group proteins [42-46]. The latter proteins are crucial developmental regulators that maintain transcriptional repression of hundreds of genes involved in development from early

embryogenesis through birth to adulthood [47-50, 37], as well as in signaling or cancer development using chromatin-based epigenetic mechanisms [51, 52]. In particular, the Polycomb group RING finger protein 1 (PCGF1) acts as key cell growth regulator, that promotes cell cycle progression and proliferation by transcriptional repression of the CDK cyclin inhibitor p21Waf1/Cip1 [53]. Interestingly, our study reveals that Polycomb group RING finger protein 1 (PCGF1) has not been detected in aged blastocysts, providing a further confirmation of the key role of ubiquitylation for initiation of pregnancy. Noteworthy, the importance of the crosstalk between Polycomb group proteins and histones has been recently demonstrated by the identification of the Jumonji protein JARID2, a component of the histone methyltransferase complex that regulates binding of the Polycomb repressive complex 2 to target genes in embryonic stem cells [49], as the most represented proteins in the secretome of the preimplantation human embryo able to predict positive pregnancy outcome [54]. Moreover, according to a recent observation describing the important role of histone H2A variants (H2A.X, H2A.Z and macroH2A) in genome remodeling after fertilization [55] and preimplantation embryo development [56], our study reveals that blastocoel fluid of aged woman undergoes a global change in histone composition, as it contains soluble histones (H2A.X, H2B and H4) that are completely absent in younger women. Noteworthy, any excess of the positively charged histones can allow them to potentially associate non-specifically with negatively charged molecules such as DNA in the cell, resulting in deleterious effects on cell viability and genomic stability [57]. Indeed, recent studies have highlighted a link between increased soluble H2A.X and apoptosis induction [58, 59]. Excessive soluble H2A.X causes chromatin aggregation and inhibition of ongoing gene transcription, but it has also non-nucleosomal functions, specifically pro-apoptotic activities. Therefore, to avoid the problems associated with excess histone accumulation, most eukaryotic cells largely rely on the strict regulation of their histone protein levels and excess histones are degraded by phosphorylation and ubiquitylation-dependent proteolysis [60, 61]. Moreover, a very recent report suggests that histones may be targeted for degradation also via the chaperone-mediated autophagy pathway in mammalian cells [62]. In keeping with these observations, we identified in blastocoel fluid from aged blastocysts an increased amount of Ubiquilin-2 (UBQLN2), which is involved in regulation of different protein degradation mechanisms and pathways including ubiquitin-proteasome system and autophagy, a survival mechanism whose deregulation has been linked to non-apoptotic cell death through inducing cell cycle arrest [63, 64]. Noteworthy, increased autophagy of aged blastocysts is further confirmed by absence in the blastocoel fluid of Laminin subunit gamma-1 (LAMC1), which is crucially involved in protection from autophagy [65, 66].

Besides ubiquitin, our proteomic data reveals that also FAAH2, a key protein in the endocannabinoid system, is differentially expressed in young and aged women blastocoel fluid. Recently, comprehensive studies have indicated that endocannabinoids, a group of bioactive lipids including anandamide (AEA), are actively involved in the process of

reproduction and pregnancy, from gametogenesis, uterine receptivity, early embryo-blastocyst development, oviductal transport and implantation through to birth [67-69]. As a consequence of its potential physiological impact, endocannabinoid signaling is finely regulated at multiple levels through several metabolic routes, including the main degradative enzyme fatty acid amide hydrolase (FAAH), that cleaves and inactivates AEA [70]. Changes in the activity and/or expression of FAAH determine significant fluctuations in AEA levels, which in turn can lead to success or failure of pregnancy [71]. In line with this, it has been reported that women in an IVF program who have successful implantation, have low serum AEA levels associated with elevated levels of in their peripheral lymphocytes at 6 weeks gestation, whereas low FAAH [72, 73] and high AEA levels in blood are associated with spontaneous threatened miscarriage [74]. In keeping with these observations, our data revealed that lower levels of FAAH2, an isoform highly expressed in higher placental mammals [75] in the blastocoel fluid of aged woman. Given the potential impact of this observation, additional studies will be devoted to validate this preliminary result.

In summary, by applying a proteomic approach to compare blastocoel fluid from <37 years old and  $\geq 37$  year old women blastocysts, we suggest that increased rate of implantation failure in aged IVF patients relies on the pro-apoptotic signature of their blastocysts, as a consequence of the dramatic dysregulation of mechanisms involved in the regulation of ubiquitin system as well as, potentially, in the control of a correct AEA tone. Noteworthy, several additional proteins not related to the ubiquitin system but equally involved in apoptotic processes have been found differentially regulated among the two groups, such as the Isoform 2 of Protein BEX3 (NGFRAP1) [76], Dermicin (DCD) [77], the Isoform 3 of Proline-rich AKT1 substrate 1 (AKT1S1) [78], Calmodulin (CALM) [79] and the eukaryotic translation elongation factor 1A (EEF1A) [80, 81], providing a further confirmation of the pro-apoptotic signature of aged blastocysts.

In conclusion the results of our deep proteomic investigation provides valuable information on the molecular player involved in biological processes related to blastocyst features in aged women which, in turn, can be used as a starting point towards the identification of non-invasive tools to investigate and ameliorate blastocyst viability, and improving clinical outcome of IVF cycles.

## **Acknowledgments**

We thank Dr. Francesca Grassi Scalvini and Dr. Fabiana Santagata for their skillful technical assistance.

## **Funding**

None declared.

## **Authors' Roles**

G.T., E.A., V.P., A.B., P.E.L.S. designed research; G.T., E.A., V.P., A.B., E.M., A.N., S.N. performed research; G.T., E.A., V.P., A.B., E.M.B., E.M., A.N., S.N., M.M. analyzed data; G.T., E.A., V.P., A.B., E.M.B., M.M. wrote the paper.

## **Compliance with ethical standards**

## **Conflict of interest**

None declared.

## References List

1. Paiva P, Salamonsen LA, Manuelpillai U, Dimitriadis E. Interleukin 11 inhibits human trophoblast invasion indicating a likely role in the decidual restraint of trophoblast invasion during placentation. *Biology of reproduction*. 2009;80(2):302-10. doi:10.1095/biolreprod.108.071415.
2. Koot YE, Teklenburg G, Salker MS, Brosens JJ, Macklon NS. Molecular aspects of implantation failure. *Biochimica et biophysica acta*. 2012;1822(12):1943-50. doi:10.1016/j.bbadis.2012.05.017.
3. Dimitriadis E, White CA, Jones RL, Salamonsen LA. Cytokines, chemokines and growth factors in endometrium related to implantation. *Human reproduction update*. 2005;11(6):613-30. doi:10.1093/humupd/dmi023.
4. Humm KC, Dodge LE, Wu LH, Penzias AS, Malizia BA, Sakkas D et al. In vitro fertilization in women under 35: counseling should differ by age. *Journal of assisted reproduction and genetics*. 2015. doi:10.1007/s10815-015-0570-7.
5. Malizia BA, Hacker MR, Penzias AS. Cumulative live-birth rates after in vitro fertilization. *The New England journal of medicine*. 2009;360(3):236-43. doi:10.1056/NEJMoa0803072.
6. Luke B, Brown MB, Wantman E, Lederman A, Gibbons W, Schattman GL et al. Cumulative birth rates with linked assisted reproductive technology cycles. *The New England journal of medicine*. 2012;366(26):2483-91. doi:10.1056/NEJMoa1110238.
7. Aplin JD. Embryo implantation: the molecular mechanism remains elusive. *Reproductive biomedicine online*. 2006;13(6):833-9.
8. Garrido-Gomez T, Dominguez F, Simon C. Proteomics of embryonic implantation. *Handbook of experimental pharmacology*. 2010(198):67-78. doi:10.1007/978-3-642-02062-9\_5.
9. Aydiner F, Yetkin CE, Seli E. Perspectives on emerging biomarkers for non-invasive assessment of embryo viability in assisted reproduction. *Current molecular medicine*. 2010;10(2):206-15.
10. Bromer JG, Seli E. Assessment of embryo viability in assisted reproductive technology: shortcomings of current approaches and the emerging role of metabolomics. *Current opinion in obstetrics & gynecology*. 2008;20(3):234-41. doi:10.1097/GCO.0b013e3282fe723d.
11. Latham KE, Garrels JI, Chang C, Solter D. Analysis of embryonic mouse development: construction of a high-resolution, two-dimensional gel protein database. *Applied and theoretical electrophoresis : the official journal of the International Electrophoresis Society*. 1992;2(6):163-70.
12. Shi CZ, Collins HW, Garside WT, Buettger CW, Matschinsky FM, Heyner S. Protein databases for compacted eight-cell and blastocyst-stage mouse embryos. *Molecular reproduction and development*. 1994;37(1):34-47. doi:10.1002/mrd.1080370106.

13. Katz-Jaffe MG, Linck DW, Schoolcraft WB, Gardner DK. A proteomic analysis of mammalian preimplantation embryonic development. *Reproduction*. 2005;130(6):899-905. doi:10.1530/rep.1.00854.
14. Katz-Jaffe MG, Gardner DK, Schoolcraft WB. Proteomic analysis of individual human embryos to identify novel biomarkers of development and viability. *Fertility and sterility*. 2006;85(1):101-7. doi:10.1016/j.fertnstert.2005.09.011.
15. Katz-Jaffe MG, Schoolcraft WB, Gardner DK. Analysis of protein expression (secretome) by human and mouse preimplantation embryos. *Fertility and sterility*. 2006;86(3):678-85. doi:10.1016/j.fertnstert.2006.05.022.
16. Poli M, Ori A, Child T, Jaroudi S, Spath K, Beck M et al. Characterization and quantification of proteins secreted by single human embryos prior to implantation. *EMBO molecular medicine*. 2015;7(11):1465-79. doi:10.15252/emmm.201505344.
17. Jensen PL, Beck HC, Petersen J, Hreinsson J, Wanggren K, Laursen SB et al. Proteomic analysis of human blastocoel fluid and blastocyst cells. *Stem cells and development*. 2013;22(7):1126-35. doi:10.1089/scd.2012.0239.
18. Katz-Jaffe MG, McReynolds S, Gardner DK, Schoolcraft WB. The role of proteomics in defining the human embryonic secretome. *Molecular human reproduction*. 2009;15(5):271-7. doi:10.1093/molehr/gap012.
19. Gardner RL, Beddington RS. Multi-lineage 'stem' cells in the mammalian embryo. *Journal of cell science Supplement*. 1988;10:11-27.
20. Chambery A, Colucci-D'Amato L, Vissers JP, Scarpella S, Langridge JI, Parente A. Proteomic profiling of proliferating and differentiated neural mes-c-myc A1 cell line from mouse embryonic mesencephalon by LC-MS. *Journal of proteome research*. 2009;8(1):227-38. doi:10.1021/pr800454n.
21. Levin Y, Wang L, Ingudomnukul E, Schwarz E, Baron-Cohen S, Palotas A et al. Real-time evaluation of experimental variation in large-scale LC-MS/MS-based quantitative proteomics of complex samples. *Journal of chromatography B, Analytical technologies in the biomedical and life sciences*. 2009;877(13):1299-305. doi:10.1016/j.jchromb.2008.11.007.
22. Levi-Setti PE, Menduni F, Smeraldi A, Patrizio P, Morengi E, Albani E. Artificial shrinkage of blastocysts prior to vitrification improves pregnancy outcome: analysis of 1028 consecutive warming cycles. *Journal of assisted reproduction and genetics*. 2016;33(4):461-6. doi:10.1007/s10815-016-0655-y.
23. Vernocchi V, Morselli MG, Varesi S, Nonnis S, Maffioli E, Negri A et al. Sperm ubiquitination in epididymal feline semen. *Theriogenology*. 2014;82(4):636-42. doi:10.1016/j.theriogenology.2014.06.002.
24. Maffioli E AF, Nonnis S, Santagata F, Negri A, DeLano FA, Santamaria MH, Kistler EB, Schmid-Schönbein GW, Tedeschi G. Analysis of Rat Plasma Peptidome in Hemorrhagic Shock. *Shock* 2015;44 (Suppl 2:6): .

25. Tamplenizza M, Lenardi C, Maffioli E, Nonnis S, Negri A, Forti S et al. Nitric oxide synthase mediates PC12 differentiation induced by the surface topography of nanostructured TiO<sub>2</sub>. *Journal of nanobiotechnology*. 2013;11:35. doi:10.1186/1477-3155-11-35.
26. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology*. 2008;26(12):1367-72. doi:10.1038/nbt.1511.
27. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*. 2009;4(1):44-57. doi:10.1038/nprot.2008.211.
28. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic acids research*. 2015;43(Database issue):D447-52. doi:10.1093/nar/gku1003.
29. Fraser KB, Moehle MS, Daher JP, Webber PJ, Williams JY, Stewart CA et al. LRRK2 secretion in exosomes is regulated by 14-3-3. *Human molecular genetics*. 2013;22(24):4988-5000. doi:10.1093/hmg/ddt346.
30. Vassena R, Boue S, Gonzalez-Roca E, Aran B, Auer H, Veiga A et al. Waves of early transcriptional activation and pluripotency program initiation during human preimplantation development. *Development*. 2011;138(17):3699-709. doi:10.1242/dev.064741.
31. Bebington C, Doherty FJ, Fleming SD. The possible biological and reproductive functions of ubiquitin. *Human reproduction update*. 2001;7(1):102-11.
32. Bebington C, Bell SC, Doherty FJ, Fazleabas AT, Fleming SD. Localization of ubiquitin and ubiquitin cross-reactive protein in human and baboon endometrium and decidua during the menstrual cycle and early pregnancy. *Biology of reproduction*. 1999;60(4):920-8.
33. Bebington C, Doherty FJ, Fleming SD. Ubiquitin cross-reactive protein gene expression is increased in decidualized endometrial stromal cells at the initiation of pregnancy. *Molecular human reproduction*. 1999;5(10):966-72.
34. Bebington C, Doherty FJ, Ndukwe G, Fleming SD. The progesterone receptor and ubiquitin are differentially regulated within the endometrial glands of the natural and stimulated cycle. *Molecular human reproduction*. 2000;6(3):264-8.
35. Grzmil P, Altmann ME, Adham IM, Engel U, Jarry H, Schweyer S et al. Embryo implantation failure and other reproductive defects in Ube2q1-deficient female mice. *Reproduction*. 2013;145(1):45-56. doi:10.1530/REP-12-0054.
36. Sutovsky P, Motlik J, Neuber E, Pavlok A, Schatten G, Palecek J et al. Accumulation of the proteolytic marker peptide ubiquitin in the trophoblast of mammalian blastocysts. *Cloning and stem cells*. 2001;3(3):157-61. doi:10.1089/153623001753205115.

37. Voncken JW, Roelen BA, Roefs M, de Vries S, Verhoeven E, Marino S et al. Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proceedings of the National Academy of Sciences of the United States of America*. 2003;100(5):2468-73. doi:10.1073/pnas.0434312100.
38. Leeb M, Wutz A. Ring1B is crucial for the regulation of developmental control genes and PRC1 proteins but not X inactivation in embryonic cells. *The Journal of cell biology*. 2007;178(2):219-29. doi:10.1083/jcb.200612127.
39. Fenner BJ, Scannell M, Prehn JH. Identification of polyubiquitin binding proteins involved in NF-kappaB signaling using protein arrays. *Biochimica et biophysica acta*. 2009;1794(7):1010-6. doi:10.1016/j.bbapap.2009.02.013.
40. Nacerddine K, Lehembre F, Bhaumik M, Artus J, Cohen-Tannoudji M, Babinet C et al. The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Developmental cell*. 2005;9(6):769-79. doi:10.1016/j.devcel.2005.10.007.
41. Reverter D, Lima CD. Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. *Nature*. 2005;435(7042):687-92. doi:10.1038/nature03588.
42. de Napoles M, Mermoud JE, Wakao R, Tang YA, Endoh M, Appanah R et al. Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Developmental cell*. 2004;7(5):663-76. doi:10.1016/j.devcel.2004.10.005.
43. Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS et al. Role of histone H2A ubiquitination in Polycomb silencing. *Nature*. 2004;431(7010):873-8. doi:10.1038/nature02985.
44. Endoh M, Endo TA, Endoh T, Isono K, Sharif J, Ohara O et al. Histone H2A mono-ubiquitination is a crucial step to mediate PRC1-dependent repression of developmental genes to maintain ES cell identity. *PLoS genetics*. 2012;8(7):e1002774. doi:10.1371/journal.pgen.1002774.
45. Kalb R, Latwiel S, Baymaz HI, Jansen PW, Muller CW, Vermeulen M et al. Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nature structural & molecular biology*. 2014;21(6):569-71. doi:10.1038/nsmb.2833.
46. Luo M, Zhou J, Leu NA, Abreu CM, Wang J, Anguera MC et al. Polycomb protein SCML2 associates with USP7 and counteracts histone H2A ubiquitination in the XY chromatin during male meiosis. *PLoS genetics*. 2015;11(1):e1004954. doi:10.1371/journal.pgen.1004954.
47. Qin J, Whyte WA, Anderssen E, Apostolou E, Chen HH, Akbarian S et al. The polycomb group protein L3mbtl2 assembles an atypical PRC1-family complex that is essential in pluripotent stem cells and early development. *Cell stem cell*. 2012;11(3):319-32. doi:10.1016/j.stem.2012.06.002.



48. Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature*. 2006;441(7091):349-53. doi:10.1038/nature04733.
49. Pasini D, Cloos PA, Walfridsson J, Olsson L, Bukowski JP, Johansen JV et al. JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. *Nature*. 2010;464(7286):306-10. doi:10.1038/nature08788.
50. O'Carroll D, Erhardt S, Pagani M, Barton SC, Surani MA, Jenuwein T. The polycomb-group gene *Ezh2* is required for early mouse development. *Molecular and cellular biology*. 2001;21(13):4330-6. doi:10.1128/MCB.21.13.4330-4336.2001.
51. Kerppola TK. Polycomb group complexes--many combinations, many functions. *Trends in cell biology*. 2009;19(12):692-704. doi:10.1016/j.tcb.2009.10.001.
52. Simon JA, Kingston RE. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nature reviews Molecular cell biology*. 2009;10(10):697-708. doi:10.1038/nrm2763.
53. Gong Y, Yue J, Wu X, Wang X, Wen J, Lu L et al. NSPc1 is a cell growth regulator that acts as a transcriptional repressor of p21<sup>Waf1/Cip1</sup> via the RARE element. *Nucleic acids research*. 2006;34(21):6158-69. doi:10.1093/nar/gkl834.
54. Cortezzi SS, Garcia JS, Ferreira CR, Braga DP, Figueira RC, Iaconelli A, Jr. et al. Secretome of the preimplantation human embryo by bottom-up label-free proteomics. *Analytical and bioanalytical chemistry*. 2011;401(4):1331-9. doi:10.1007/s00216-011-5202-1.
55. Nashun B, Yukawa M, Liu H, Akiyama T, Aoki F. Changes in the nuclear deposition of histone H2A variants during pre-implantation development in mice. *Development*. 2010;137(22):3785-94. doi:10.1242/dev.051805.
56. Wu BJ, Dong FL, Ma XS, Wang XG, Lin F, Liu HL. Localization and expression of histone H2A variants during mouse oogenesis and preimplantation embryo development. *Genetics and molecular research : GMR*. 2014;13(3):5929-39. doi:10.4238/2014.August.7.8.
57. Singh RK, Liang D, Gajjalaiahvari UR, Kabbaj MH, Paik J, Gunjan A. Excess histone levels mediate cytotoxicity via multiple mechanisms. *Cell Cycle*. 2010;9(20):4236-44.
58. Liu Y, Tseng M, Perdreau SA, Rossi F, Antonescu C, Besmer P et al. Histone H2AX is a mediator of gastrointestinal stromal tumor cell apoptosis following treatment with imatinib mesylate. *Cancer research*. 2007;67(6):2685-92. doi:10.1158/0008-5472.CAN-06-3497.
59. Liu Y, Parry JA, Chin A, Duensing S, Duensing A. Soluble histone H2AX is induced by DNA replication stress and sensitizes cells to undergo apoptosis. *Molecular cancer*. 2008;7:61. doi:10.1186/1476-4598-7-61.

60. Singh RK, Kabbaj MH, Paik J, Gunjan A. Histone levels are regulated by phosphorylation and ubiquitylation-dependent proteolysis. *Nature cell biology*. 2009;11(8):925-33. doi:10.1038/ncb1903.
61. Singh RK, Paik J, Gunjan A. Generation and management of excess histones during the cell cycle. *Front Biosci (Landmark Ed)*. 2009;14:3145-58.
62. Cook AJ, Gurard-Levin ZA, Vassias I, Almouzni G. A specific function for the histone chaperone NASP to fine-tune a reservoir of soluble H3-H4 in the histone supply chain. *Molecular cell*. 2011;44(6):918-27. doi:10.1016/j.molcel.2011.11.021.
63. Li Y, Lu J, Prochownik EV. Dual role for SUMO E2 conjugase Ubc9 in modulating the transforming and growth-promoting properties of the HMGA1b architectural transcription factor. *The Journal of biological chemistry*. 2007;282(18):13363-71. doi:10.1074/jbc.M610919200.
64. Benbrook DM, Long A. Integration of autophagy, proteasomal degradation, unfolded protein response and apoptosis. *Experimental oncology*. 2012;34(3):286-97.
65. Murray P, Edgar D. Regulation of programmed cell death by basement membranes in embryonic development. *The Journal of cell biology*. 2000;150(5):1215-21.
66. Carmignac V, Svensson M, Korner Z, Elowsson L, Matsumura C, Gawlik KI et al. Autophagy is increased in laminin alpha2 chain-deficient muscle and its inhibition improves muscle morphology in a mouse model of MDC1A. *Human molecular genetics*. 2011;20(24):4891-902. doi:10.1093/hmg/ddr427.
67. Taylor AH, Amoako AA, Bambang K, Karasu T, Gebeh A, Lam PM et al. Endocannabinoids and pregnancy. *Clinica chimica acta; international journal of clinical chemistry*. 2010;411(13-14):921-30. doi:10.1016/j.cca.2010.03.012.
68. Maccarrone M. Endocannabinoids: friends and foes of reproduction. *Progress in lipid research*. 2009;48(6):344-54. doi:10.1016/j.plipres.2009.07.001.
69. Battista N, Bari M, Maccarrone M. Endocannabinoids and Reproductive Events in Health and Disease. *Handbook of experimental pharmacology*. 2015;231:341-65. doi:10.1007/978-3-319-20825-1\_12.
70. McKinney MK, Cravatt BF. Structure and function of fatty acid amide hydrolase. *Annual review of biochemistry*. 2005;74:411-32. doi:10.1146/annurev.biochem.74.082803.133450.
71. Rapino C, Battista N, Bari M, Maccarrone M. Endocannabinoids as biomarkers of human reproduction. *Human reproduction update*. 2014;20(4):501-16. doi:10.1093/humupd/dmu004.

72. Maccarrone M, Bisogno T, Valensise H, Lazzarin N, Fezza F, Manna C et al. Low fatty acid amide hydrolase and high anandamide levels are associated with failure to achieve an ongoing pregnancy after IVF and embryo transfer. *Molecular human reproduction*. 2002;8(2):188-95.
73. Maccarrone M, Valensise H, Bari M, Lazzarin N, Romanini C, Finazzi-Agro A. Relation between decreased anandamide hydrolase concentrations in human lymphocytes and miscarriage. *Lancet*. 2000;355(9212):1326-9. doi:10.1016/S0140-6736(00)02115-2.
74. Habayeb OM, Taylor AH, Finney M, Evans MD, Konje JC. Plasma anandamide concentration and pregnancy outcome in women with threatened miscarriage. *Jama*. 2008;299(10):1135-6. doi:10.1001/jama.299.10.1135.
75. Wei BQ, Mikkelsen TS, McKinney MK, Lander ES, Cravatt BF. A second fatty acid amide hydrolase with variable distribution among placental mammals. *The Journal of biological chemistry*. 2006;281(48):36569-78. doi:10.1074/jbc.M606646200.
76. Jiang JY, Xiong H, Cao M, Xia X, Sirard MA, Tsang BK. Mural granulosa cell gene expression associated with oocyte developmental competence. *Journal of ovarian research*. 2010;3:6. doi:10.1186/1757-2215-3-6.
77. Esposito G, Schiattarella GG, Perrino C, Cattaneo F, Pironti G, Franzone A et al. Dermcidin: a skeletal muscle myokine modulating cardiomyocyte survival and infarct size after coronary artery ligation. *Cardiovascular research*. 2015;107(4):431-41. doi:10.1093/cvr/cvv173.
78. Huang B, Porter G. Expression of proline-rich Akt-substrate PRAS40 in cell survival pathway and carcinogenesis. *Acta pharmacologica Sinica*. 2005;26(10):1253-8. doi:10.1111/j.1745-7254.2005.00184.x.
79. O'Neill C, Li Y, Jin XL. Survival signaling in the preimplantation embryo. *Theriogenology*. 2012;77(4):773-84. doi:10.1016/j.theriogenology.2011.12.016.
80. Abbas W, Kumar A, Herbein G. The eEF1A Proteins: At the Crossroads of Oncogenesis, Apoptosis, and Viral Infections. *Frontiers in oncology*. 2015;5:75. doi:10.3389/fonc.2015.00075.
81. Migliaccio N, Ruggiero I, Martucci NM, Sanges C, Arbucci S, Tate R et al. New insights on the interaction between the isoforms 1 and 2 of human translation elongation factor 1A. *Biochimie*. 2015. doi:10.1016/j.biochi.2015.07.021.

## Figure Legends

**Fig. 1 Proteomic profile of human blastocoel fluid isolated from women aged < (group A) and  $\geq$  (group B) 37 years.** **a** Venn diagram of the proteins identified in group A and B. **b** Vulcano plot showing the proteins common to both groups differentially expressed in group B compared to group A. Common proteins were considered differentially expressed if they showed significant t-test difference (cut-off at 5% permutation-based False Discovery Rate) between the two groups. Proteins up and down regulated are indicated in black and dark grey, respectively. Proteins which are present at levels not statistically different between groups A and B are indicated in light grey below the horizontal threshold line

**Fig. 2 Classification of the proteins differentially expressed based on GOMF and GOBP.** The proteins differentially expressed in A and B groups were classified based on molecular function (GOMF) and biological processes (GOBP) **a** Classification of proteins down regulated in B or present only in A. **b** Classification of proteins up regulated or present only in B.

**Fig. 3 Functional analysis of the proteins differentially expressed in A and B.** The proteins were manually grouped according to their function in biological processes taking into account, besides GOMF, and GOBP, also keywords and annotations known from literature (Supplementary Table 4). **a** Classification of proteins down regulated in B or present only in A. **b** Classification of proteins up regulated or present only in B.

**Fig. 4 Protein-protein interaction analysis of the proteins differentially expressed in A and B involved in the ubiquitin system.** The analysis was performed by STRING. Stronger associations are represented by thicker lines. Proteins increased or decreased in B are indicated by an arrow. PSMD8, PSMD4, PSMD2, PSMD14 are isoforms of PSMD4 which was found decreased in B.

**Table S1: List of proteins discarded from the analysis.** Proteins identified in the control (containing only culture medium) and keratins likely due to a contamination during sample preparation, were discarded from the analysis.

| Protein IDs | Medium/Keratin  |             |
|-------------|---|-------------|
|             | Protein names   | Gene names  |
| P04217      | Alpha-1B-glycoprotein   | A1BG        |
| P43652      | Afamin  | AFM         |
| C9JV77      | Alpha-2-HS-glycoprotein   | AHSG        |
| P02768      | Serum albumin   | ALB         |
| P25311      | Zinc-alpha-2-glycoprotein   | AZGP1       |
| Q6ISB3      | Grainyhead-like protein 2 homolog                                 | GRHL2       |
| P69905      | Hemoglobin subunit alpha  | HBA1        |
| P00738      | Haptoglobin   | HP          |
| P02790      | Hemopexin   | HPX         |
| P01876      | Ig alpha-1 chain C region   | IGHA1       |
| P04264      | Keratin, type II cytoskeletal 1                                   | KRT1        |
| P13645      | Keratin, type I cytoskeletal 10                                   | KRT10       |
| P02533      | Keratin, type I cytoskeletal 14                                   | KRT14       |
| P08779      | Keratin, type I cytoskeletal 16                                   | KRT16       |
| Q04695      | Keratin, type I cytoskeletal 17                                   | KRT17       |
| P35908      | Keratin, type II cytoskeletal 2 epidermal                         | KRT2        |
| P13647      | Keratin, type II cytoskeletal 5                                   | KRT5        |
| P02538      | Keratin, type II cytoskeletal 6A;Keratin, type II cytoskeletal 6C | KRT6A;KRT6C |
| Q7Z794      | Keratin, type II cytoskeletal 1b                                  | KRT77       |
| Q8N1N4      | Keratin, type II cytoskeletal 78                                  | KRT78       |
| Q9NSB4      | Keratin, type II cuticular Hb2                                    | KRT82       |
| P35527      | Keratin, type I cytoskeletal 9                                    | KRT9        |
| F8WCG0      | Myotubularin-related protein 14                                   | MTMR14      |
| P02763      | Alpha-1-acid glycoprotein 1                                       | ORM1        |
| A6NNF7      | Amiloride-sensitive sodium channel subunit delta                  | SCNN1D      |
| C9JNR7      | Septin-10   | SEPT10      |
| P02787      | Serotransferrin   | TF          |
| P02766      | Transthyretin   | TTR         |
| E7EQ64      | Trypsin-1;Alpha-trypsin chain 1;Alpha-trypsin chain 2             | PRSS1       |

**Table S2: List of proteins identified in group A by a shotgun approach.** Previous identification: in the Table the proteins previously identified in human blastocoel fluid (16, 17) are indicated with a capital X, while x indicates gene products previously experimentally detected at the protein level as secreted and/or present in extracellular region and/or in exosome and/or in microvesicles according to public available databases and dataset (www.neXtProt.org, www.exocarta.org, microvesicle.org, geneontology.org, www.proteinatlas.org, 29) or at the gene expression level from microarray data (16, 92) or at the transcription level in blastocystis or embryonic tissue according to <http://genomewidepdb.proteomix.org> or <https://www.ncbi.nlm.nih.gov/geoprofiles> databases. The table lists only the proteins present and quantified in at least 3 out of 5 technical repeats using a false positive rate of 0.05 and minimum peptide length 9 amino acids, reporting in a column the proteins identified with FDR 0.01.

| Protein IDs | Protein names   | Gene names     | PEP      | Razor + unique peptides | FDR 1% | Previous identification |
|-------------|---|----------------|----------|-------------------------|--------|-------------------------|
| P12273      | Prolactin-inducible protein                                       | PIP            | 5.24E-06 | 2                       | X      | X                       |
| Q8NBJ4      | Golgi membrane protein 1  | GOLM1          | 0.018711 | 1                       | X      | x                       |
| P01040      | Cystatin-A  | CSTA           | 1.22E-08 | 2                       | X      | X                       |
| P20382      | Pro-MCH;Neuropeptide-glycine-glutamic acid;                       | PMCH           | 0.002462 | 1                       | X      | x                       |
| Q8TE76      | MORC family CW-type zinc finger protein 4                         | MORC4          | 0.018241 | 1                       | X      | x                       |
| Q8TE56      | A disintegrin and metalloproteinase with thrombospondin motifs 17 | ADAMTS17       | 0.022011 | 1                       | X      | x                       |
| D6R956      | Ubiquitin carboxyl-terminal hydrolase isozyme L1                  | UCHL1          | 0.004585 | 1                       |        | X                       |
| O95202      | LETM1 and EF-hand domain-containing protein 1, mitochondrial      | LETM1          | 0.028673 | 1                       |        | x                       |
| Q9UGM3      | Deleted in malignant brain tumors 1 protein                       | DMBT1          | 0.006933 | 1                       | X      | X                       |
| P31025      | Lipocalin-1;Putative lipocalin 1-like protein 1                   | LCN1;LCN1P1    | 1.85E-55 | 2                       | X      | X                       |
| Q6UWP8      | Suprabasin  | SBSN           | 0.000208 | 1                       | X      | X                       |
| Q9BR77      | Coiled-coil domain-containing protein 77                          | CCDC77         | 0.009722 | 1                       | X      | x                       |
| Q96RW7      | Hemicentin-1  | HMCN1          | 0.024825 | 1                       | X      | x                       |
| P31151      | Protein S100-A7;Protein S100-A7A                                  | S100A7;S100A7A | 3.14E-51 | 4                       | X      | X                       |
| H0Y7A7      | Calmodulin  | CALM2;CALM1    | 0.005618 | 1                       | X      | x                       |
| Q6GMR7      | Fatty-acid amide hydrolase 2                                      | FAAH2          | 0.02707  | 1                       | X      | x                       |
| P19021      | Peptidyl-glycine alpha-amidating monooxygenase;                   | PAM            | 0.029696 | 1                       |        | x                       |
| O94788-3    | Retinal dehydrogenase 2   | ALDH1A2        | 0.026311 | 1                       |        | x                       |
| Q5VW26      | Nuclear factor 1;Nuclear factor 1 B-type                          | NFIB           | 0.01383  | 1                       |        | x                       |
| Q9BV73      | Centrosome-associated protein CEP250                              | CEP250         | 0.025154 | 1                       |        | X                       |
| Q9NZT1      | Calmodulin-like protein 5   | CALML5         | 8E-153   | 4                       | X      | X                       |
| P81605      | Dermcidin;Survival-promoting peptide;DCD-1                        | DCD            | 4.24E-08 | 2                       | X      | X                       |
| Q6PGQ7      | Protein aurora borealis   | BORA;FLJ22624  | 0.03061  | 1                       | X      | x                       |
| C9K0L2      | Maestro heat-like repeat-containing protein family member 7       | HEATR8         | 0.030121 | 1                       | X      | x                       |
| P61626      | Lysozyme C  | LYZ            | 2.78E-23 | 1                       | X      | X                       |
| Q8N8E3      | Centrosomal protein of 112 kDa                                    | CEP112         | 0.008359 | 1                       | X      | x                       |

|          |  |                |          |   |   |   |
|----------|--|----------------|----------|---|---|---|
| Q9BY89   | Uncharacterized protein KIAA1671                                     | KIAA1671       | 0.012704 | 1 |   | x |
| Q96F24   | Nuclear receptor-binding factor 2                                    | NRBF2          | 0.005732 | 1 | X | x |
| E7ERA6   | RING finger protein 223  | RNF223         | 0.028835 | 1 | X |   |
| F8W8U3   | Homeobox protein Meis1   | MEIS1          | 0.031584 | 1 |   | x |
| P0CG48   | Polyubiquitin-C;Ubiquitin;Polyubiquitin-B                            | UBC;UBB        | 0.001463 | 1 | X | X |
| Q6JVE6   | Epididymal-specific lipocalin-10                                     | LCN10          | 0.027122 | 1 |   | x |
| Q9HCJ0   | Trinucleotide repeat-containing gene 6C protein                      | TNRC6C         | 0.02214  | 1 | X | x |
| F8W878   | Brain-specific angiogenesis inhibitor 1-associated protein 2         | BAIAP2         | 0.018041 | 1 | X | x |
| P05109   | Protein S100-A8  | S100A8         | 3.51E-69 | 2 | X | X |
| Q13635-4 | Isoform S of Protein patched homolog 1                               | PTCH1          | 0.0054   | 1 | X | x |
| Q8NBP5   | Major facilitator superfamily domain-containing protein 9            | MFSD9          | 0.024966 | 1 | X | x |
| F8WCM5   | Insulin;Insulin B chain;Insulin A chain                              | INS-IGF2;INS   | 6.49E-07 | 1 | X | X |
| G8JLG2   | Corneodesmosin   | CDSN           | 0.018446 | 1 |   | X |
| P01566   | Interferon alpha-10;Interferon alpha-21;                             | IFNA10;IFNA21  | 0.03168  | 1 | X | x |
| P22732   | Solute carrier family 2, facilitated glucose transporter member 5    | SLC2A5         | 0.028098 | 1 |   | x |
| Q8TAL5   | Uncharacterized protein C9orf43                                      | C9orf43        | 0.028085 | 1 |   | x |
| J3KNN1   | Polypeptide N-acetylgalactosaminyltransferase 9                      | GALNT9         | 0.024813 | 1 | X |   |
| Q02413   | Desmoglein-1   | DSG1           | 2.58E-31 | 2 | X | X |
| P14923   | Junction plakoglobin   | JUP            | 5.14E-15 | 2 | X | X |
| Q8NI35-5 | Isoform 5 of InaD-like protein                                       | INADL          | 0.018277 | 1 |   | x |
| G3V314   | Ena/VASP-like protein  | EVL            | 0.03031  | 1 |   | x |
| Q12972-3 | Isoform Gamma of Nuclear inhibitor of protein phosphatase 1          | PPP1R8         | 0.031102 | 1 |   | x |
| Q9HCE0   | Ectopic P granules protein 5 homolog                                 | EPG5           | 0.030831 | 1 |   | x |
| Q76G19   | PDZ domain-containing protein 4                                      | PDZD4          | 5.95E-15 | 2 | X | x |
| Q5T749   | Keratinocyte proline-rich protein                                    | KPRP           | 5.50E-12 | 5 | X | X |
| H3BUN7   | SH2B adapter protein 1   | SH2B1          | 0.017553 | 1 | X | x |
| H0YNB0   | Unconventional myosin-Ie   | MYO1E          | 0.022468 | 1 |   | x |
| J3KMZ5   | Serotonin N-acetyltransferase  | AANAT          | 0.014783 | 1 |   | x |
| P28290   | Sperm-specific antigen 2   | SSFA2          | 0.015064 | 1 | X | x |
| Q6ZU67   | BEN domain-containing protein 4                                      | BEND4          | 0.029939 | 1 |   | x |
| Q9H1N7   | Adenosine 3-phospho 5-phosphosulfate transporter 2                   | SLC35B3        | 0.023363 | 1 |   | x |
| P05060   | Secretogranin-1;GAWK peptide;CCB peptide                             | CHGB           | 0.023689 | 1 | X | x |
| Q9NZN9-5 | Isoform 5 of Methionine adenosyltransferase 2 subunit beta           | MAT2B          | 0.02745  | 1 |   | x |
| Q9UHD9   | Ubiquilin-2  | UBQLN2         | 0.02898  | 1 | X | x |
| Q96NH3-4 | Protein broad-minded   | BROMI;C6orf170 | 0.011429 | 1 | X | x |
| P46060   | Ran GTPase-activating protein 1                                      | RANGAP1        | 0.020731 | 1 |   | x |
| Q9H1K1-2 | Isoform 2 of Iron-sulfur cluster assembly enzyme ISCU, mitochondrial | ISCU           | 0.029122 | 1 |   | x |
| P04745   | Alpha-amylase 1;Pancreatic alpha-amylase                             | AMY1A;AMY2A    | 0.016273 | 1 |   | X |
| B4DIP4   | Eukaryotic translation initiation factor 2-alpha kinase 1            | EIF2AK1        | 0.024445 | 1 |   | x |
| Q9NYL2-  | Isoform 2 of Mitogen-activated protein                               | ZAK            | 0.010561 | 1 |   | x |

|          |   |          |          |   |   |  |   |
|----------|---|----------|----------|---|---|--|---|
| 2        | kinase kinase kinase MLT  |          |          |   |   |  |   |
| B4DYB5   | cDNA FLJ53372   | SPAG16   | 0.025535 | 1 |   |  |   |
| B7Z8Z7   | RNA-binding protein 47  | RBM47    | 0.02829  | 1 |   |  | x |
| F5H1U9   | Multiple PDZ domain protein   | MPDZ     | 0.03135  | 1 | X |  | x |
| P06454   | Prothymosin alpha;Thymosin alpha-1  | PTMA     | 0.001343 | 1 | X |  | X |
| C9JH18   | PERQ amino acid-rich with GYF domain-containing protein 2                   | GIGYF2   | 0.000158 | 2 |   |  | x |
| C9JNL6   | Electrogenic sodium bicarbonate cotransporter 4                             | SLC4A5   | 0.014409 | 1 |   |  | x |
| C9JSC0   | POU domain, class 6, transcription factor 2                                 | POU6F2   | 0.029639 | 1 |   |  | x |
| Q96BQ5   | Coiled-coil domain-containing protein 127                                   | CCDC127  | 0.031674 | 1 | X |  | x |
| Q9NQ38   | Serine protease inhibitor Kazal-type 5                                      | SPINK5   | 0.001557 | 1 | X |  | x |
| E7ENK3   | Glutamate receptor, ionotropic kainate 1                                    | GRIK1    | 0.027858 | 1 |   |  | x |
| Q13201   | Multimerin-1;Platelet glycoprotein Ia*;155 kDa platelet multimerin          | MMRN1    | 0.024454 | 1 | X |  | x |
| P02788   | Lactotransferrin;Kaliocin-1;Lactoferroxin-A;Lactoferroxin-B;Lactoferroxin-C | LTF      | 0.004647 | 1 | X |  | X |
| P04406   | Glyceraldehyde-3-phosphate dehydrogenase                                    | GAPDH    | 2.45E-07 | 3 | X |  | X |
| P11532   | Dystrophin  | DMD      | 0.015725 | 1 |   |  | x |
| E9PQB1   | Serine/arginine repetitive matrix protein 5                                 | SRRM5    | 0.015627 | 1 | X |  | x |
| Q9UKR0   | Kallikrein-12   | KLK12    | 0.02885  | 1 |   |  | x |
| F8WAU4   | Elongation factor G, mitochondrial  | GFM1     | 0.014496 | 1 | X |  | x |
| F8WCP6   | Kinesin-like protein KIF21A   | KIF21A   | 0.029222 | 1 |   |  | x |
| Q9P0W5-4 | Schwannomin-interacting protein 1   | SCHIP1   | 0.028326 | 1 |   |  | x |
| H0Y3B9   | BTB/POZ domain-containing protein KCTD2                                     | KCTD2    | 0.027761 | 1 | X |  | x |
| H0Y7H7   | Dedicator of cytokinesis protein 4  | DOCK4    | 0.030214 | 1 |   |  | x |
| Q13535   | Serine/threonine-protein kinase ATR   | ATR      | 0.024667 | 1 |   |  | x |
| Q6FIF0   | AN1-type zinc finger protein 6  | ZFAND6   | 0.030174 | 1 |   |  | x |
| P14618   | Pyruvate kinase isozymes M1/M2;Pyruvate kinase                              | PKM2     | 0.000118 | 2 | X |  | X |
| J9JIC5   | Protein Njmu-R1   | C17orf75 | 0.031192 | 1 | X |  | x |
| Q96DH6   | RNA-binding protein Musashi homolog 2                                       | MSI2     | 0.007825 | 1 |   |  | x |
| K7EKM7   | Small integral membrane protein 24  | SMIM24   | 0.009651 | 1 | X |  | x |
| M0QXR7   | Leucine-rich repeat-containing protein 63                                   | LRRC63   | 0.020219 | 1 |   |  |   |
| O00555   | Voltage-dependent P/Q-type calcium channel subunit alpha-1A                 | CACNA1A  | 0.000283 | 2 | X |  | x |
| O60320-2 | Isoform 2 of Protein FAM189A1   | FAM189A1 | 0.012422 | 1 |   |  | x |
| O75556   | Mammaglobin-B   | SCGB2A1  | 0.030478 | 1 |   |  | x |
| O76042   | Putative uncharacterized protein C3orf51                                    | C3orf51  | 0.01043  | 1 | X |  |   |
| P05089   | Arginase-1  | ARG1     | 0.000272 | 2 | X |  | X |
| P06702   | Protein S100-A9   | S100A9   | 8.91E-09 | 3 | X |  | X |
| P06733   | Alpha-enolase   | ENO1     | 2.54E-06 | 2 |   |  | X |
| P11047   | Laminin subunit gamma-1   | LAMC1    | 0.029392 | 1 |   |  | x |
| P15924   | Desmoplakin   | DSP      | 3.07E-21 | 7 | X |  | X |
| P31944   | Caspase-14;Caspase-14 subunit p19;Caspase-14 subunit p10                    | CASP14   | 3.99E-26 | 4 | X |  | X |
| P51843   | Nuclear receptor subfamily 0 group B member 1                               | NR0B1    | 0.013215 | 1 |   |  | x |



|          |   |          |          |   |   |   |
|----------|---|----------|----------|---|---|---|
| P52740   | Zinc finger protein 132                                       | ZNF132   | 0.002584 | 1 | X | x |
| Q5VWC4   | 26S proteasome non-ATPase regulatory subunit 4                | PSMD4    | 0.031718 | 1 |   | x |
| Q00994-2 | Isoform 2 of Protein BEX3                                     | NGFRAP1  | 0.008688 | 1 |   | x |
| Q01469   | Fatty acid-binding protein, epidermal                         | FABP5    | 1.08E-24 | 2 | X | X |
| Q08188   | Protein-glutamine gamma-glutamyltransferase E;                | TGM3     | 0.011907 | 1 |   | X |
| Q08554   | Desmocollin-1   | DSC1     | 2.80E-06 | 3 | X | X |
| Q12778   | Forkhead box protein O1                                       | FOXO1    | 0.007693 | 1 |   | x |
| Q13164-4 | Mitogen-activated protein kinase 7                            | MAPK7    | 0.027772 | 1 | X | x |
| Q13422-4 | DNA-binding protein Ikaros                                    | IKZF1    | 0.021892 | 1 |   |   |
| Q13542   | Eukaryotic translation initiation factor 4E-binding protein 2 | EIF4EBP2 | 0.020583 | 1 | X | x |
| Q14764   | Major vault protein   | MVP      | 1.25E-07 | 4 | X | X |
| Q3SYA6   | MAGEC3 protein  | MAGEC3   | 0.027463 | 1 |   |   |
| Q5D862   | Filaggrin-2   | FLG2     | 0.000426 | 2 | X | X |
| Q5GH76   | XK-related protein 4  | XKR4     | 0.008783 | 1 | X |   |
| Q5HYA8   | Meckelin  | TMEM67   | 0.010896 | 1 |   | x |
| Q5HYK7   | SH3 domain-containing protein 19                              | SH3D19   | 0.012887 | 1 | X | x |
| Q5T2D2   | Trem-like transcript 2 protein                                | TREML2   | 0.025847 | 1 |   | x |
| Q5T750   | Skin-specific protein 32                                      | XP32     | 0.016572 | 1 |   | X |
| Q5TA76   | Late cornified envelope protein 3A                            | LCE3A    | 0.024214 | 1 |   |   |
| Q6TDP4   | Kelch-like protein 17   | KLHL17   | 0.024283 | 1 |   | x |
| Q9UJQ4   | Sal-like protein 4  | SALL4    | 0.024516 | 1 |   | x |
| Q86YZ3   | Hornerin  | HRNR     | 3.22E-10 | 2 | X | X |
| Q8N139   | ATP-binding cassette sub-family A member 6                    | ABCA6    | 0.014869 | 1 | X | x |
| Q96B36-3 | Isoform 3 of Proline-rich AKT1 substrate                      | AKT1S1   | 0.028634 | 1 |   | x |
| Q9BSM1   | Polycomb group RING finger protein 1                          | PCGF1    | 0.023367 | 1 | X | x |
| Q9BTT4   | Mediator of RNA polymerase II transcription subunit 10        | MED10    | 0.003212 | 1 | X | x |
| Q9H3U1   | Protein unc-45 homolog A                                      | UNC45A   | 0.011958 | 1 |   | x |

**Table S3: List of proteins identified in group B by a shotgun approach.** Previous identification: in the Table the proteins previously identified in human blastocoel fluid (16, 17) are indicated with a capital X, while x indicates gene products previously experimentally detected at the protein level as secreted and/or present in extracellular region and/or in exosome and/or in microvesicles according to public available databases and dataset (www.neXtProt.org, www.exocarta.org, microvesicle.org, geneontology.org, www.proteinatlas.org, 29) or at the gene expression level from microarray data (16, 92) or at the transcription level in blastocystis or embryonic tissue according to <http://genomewidepdb.proteomix.org> or <https://www.ncbi.nlm.nih.gov/geoprofiles> databases. The table lists only the proteins present and quantified in at least 3 out of 5 technical repeats using a false positive rate of 0.05 and minimum peptide length 9 amino acids, reporting in a column the proteins identified with FDR 0.01.

| Protein IDs | Protein names   | Gene names     | PEP      | Razor + unique peptides | FDR 1% | Previous identification |
|-------------|---|----------------|----------|-------------------------|--------|-------------------------|
| P12273      | Prolactin-inducible protein                                       | PIP            | 5.24E-06 | 2                       | X      | X                       |
| Q8NBJ4      | Golgi membrane protein 1  | GOLM1          | 0.018711 | 1                       | X      | x                       |
| P01040      | Cystatin-A  | CSTA           | 1.22E-08 | 2                       | X      | X                       |
| P20382      | Pro-MCH;Neuropeptide-glycine-glutamic acid                        | PMCH           | 0.002462 | 1                       | X      | x                       |
| Q8TE76      | MORC family CW-type zinc finger protein 4                         | MORC4          | 0.018241 | 1                       | X      | x                       |
| Q8TE56      | A disintegrin and metalloproteinase with thrombospondin motifs 17 | ADAMTS17       | 0.022011 | 1                       | X      | x                       |
| D6R956      | Ubiquitin carboxyl-terminal hydrolase isozyme L1                  | UCHL1          | 0.004585 | 1                       |        | X                       |
| O95202      | LETM1 and EF-hand domain-containing protein 1, mitochondrial      | LETM1          | 0.028673 | 1                       |        | x                       |
| Q9UGM3      | Deleted in malignant brain tumors 1 protein                       | DMBT1          | 0.006933 | 1                       | X      | X                       |
| P31025      | Lipocalin-1;Putative lipocalin 1-like protein 1                   | LCN1;LCN1P1    | 1.85E-55 | 2                       | X      | X                       |
| Q6UWP8      | Suprabasin  | SBSN           | 0.000208 | 1                       | X      | X                       |
| Q9BR77      | Coiled-coil domain-containing protein 77                          | CCDC77         | 0.009722 | 1                       | X      | x                       |
| Q96RW7      | Hemicentin-1  | HMCN1          | 0.024825 | 1                       | X      | x                       |
| P31151      | Protein S100-A7;Protein S100-A7A                                  | S100A7;S100A7A | 3.14E-51 | 4                       | X      | X                       |
| H0Y7A7      | Calmodulin  | CALM2;CALM1    | 0.005618 | 1                       | X      | x                       |
| Q6GMR7      | Fatty-acid amide hydrolase 2                                      | FAAH2          | 0.02707  | 1                       | X      | x                       |
| P19021      | Peptidyl-glycine alpha-amidating monooxygenase                    | PAM            | 0.029696 | 1                       |        | x                       |
| O94788-3    | Isoform 3 of Retinal dehydrogenase 2                              | ALDH1A2        | 0.026311 | 1                       | X      | x                       |
| Q5VW26      | Nuclear factor 1;Nuclear factor 1 B-type                          | NFIB           | 0.01383  | 1                       |        | x                       |
| Q9BV73      | Centrosome-associated protein CEP250                              | CEP250         | 0.025154 | 1                       |        | X                       |
| Q9NZT1      | Calmodulin-like protein 5   | CALML5         | 8E-153   | 4                       | X      | X                       |
| P81605      | Dermcidin;Survival-promoting                                      | DCD            | 4.24E-08 | 2                       | X      | X                       |

|          |   |               |          |   |   |   |
|----------|---|---------------|----------|---|---|---|
|          | peptide;DCD-1   |               |          |   |   |   |
| Q6PGQ7   | Protein aurora borealis   | BORA;FLJ22624 | 0.03061  | 1 | X | x |
| C9K0L2   | Maestro heat-like repeat-containing protein family member 7       | HEATR8        | 0.030121 | 1 | X |   |
| P61626   | Lysozyme C  | LYZ           | 2.78E-23 | 1 | X | X |
| Q8N8E3   | Centrosomal protein of 112 kDa                                    | CEP112        | 0.008359 | 1 | X |   |
| Q9BY89   | Uncharacterized protein   | KIAA1671      | 0.012704 | 1 |   | x |
| Q96F24-2 | Nuclear receptor-binding factor 2                                 | NRBF2         | 0.005732 | 1 | X | x |
| E7ERA6   | RING finger protein 223   | RNF223        | 0.028835 | 1 | X |   |
| F8W8U3   | Homeobox protein Meis1  | MEIS1         | 0.031584 | 1 |   | x |
| P0CG48   | Polyubiquitin-C;Ubiquitin;Polyubiquitin-B                         | UBC;UBB       | 0.001463 | 1 | X | X |
| Q6JVE6   | Epididymal-specific lipocalin-10                                  | LCN10         | 0.027122 | 1 |   | x |
| Q9HCJ0   | Trinucleotide repeat-containing gene 6C protein                   | TNRC6C        | 0.02214  | 1 | X | x |
| F8W878   | Brain-specific angiogenesis inhibitor 1-associated protein 2      | BAIAP2        | 0.018041 | 1 | X | x |
| P05109   | Protein S100-A8   | S100A8        | 3.51E-69 | 2 | X | X |
| Q13635-4 | Isoform S of Protein patched homolog 1                            | PTCH1         | 0.0054   | 1 | X | x |
| Q8NBP5   | Major facilitator superfamily domain-containing protein 9         | MFSD9         | 0.024966 | 1 | X | x |
| F8WCM5   | Insulin;Insulin B chain;Insulin A chain                           | INS-IGF2;INS  | 6.49E-07 | 1 | X | X |
| G8JLG2   | Corneodesmosin  | CDSN          | 0.018446 | 1 |   | X |
| P01566   | Interferon alpha-10;Interferon alpha-21                           | IFNA10;IFNA21 | 0.03168  | 1 | X | x |
| P22732   | Solute carrier family 2, facilitated glucose transporter member 5 | SLC2A5        | 0.028098 | 1 |   | x |
| Q8TAL5   | Uncharacterized protein C9orf43                                   | C9orf43       | 0.028085 | 1 |   | x |
| J3KNN1   | Polypeptide N-acetylgalactosaminyltransferase 9                   | GALNT9        | 0.024813 | 1 | X |   |
| Q02413   | Desmoglein-1  | DSG1          | 2.58E-31 | 2 | X | X |
| Q8NI35-5 | Isoform 5 of InaD-like protein                                    | INADL         | 0.018277 | 1 |   | x |
| G3V314   | Ena/VASP-like protein   | EVL           | 0.03031  | 1 |   | x |
| Q12972-3 | Isoform Gamma of Nuclear inhibitor of protein phosphatase 1       | PPP1R8        | 0.031102 | 1 |   | x |
| Q9HCE0   | Ectopic P granules protein 5 homolog                              | EPG5          | 0.030831 | 1 |   | x |
| Q76G19   | PDZ domain-containing protein 4                                   | PDZD4         | 5.95E-15 | 2 | X | x |
| Q5T749   | Keratinocyte proline-rich protein                                 | KPRP          | 5.50E-12 | 5 | X | X |
| H3BUN7   | SH2B adapter protein 1  | SH2B1         | 0.017553 | 1 | X | x |
| H0YNB0   | Unconventional myosin-Ie  | MYO1E         | 0.022468 | 1 |   | x |
| J3KMZ5   | Serotonin N-acetyltransferase                                     | AANAT         | 0.014783 | 1 |   | x |
| P28290   | Sperm-specific antigen 2  | SSFA2         | 0.015064 | 1 | X | x |
| Q6ZU67   | BEN domain-containing protein 4                                   | BEND4         | 0.029939 | 1 |   | x |
| Q9H1N7   | Adenosine 3-phospho 5-phosphosulfate transporter 2                | SLC35B3       | 0.023363 | 1 |   | x |
| P05060   | Secretogranin-1;GAWK peptide;CCB peptide                          | CHGB          | 0.023689 | 1 | X | x |

|          |  |                     |          |   |   |   |
|----------|--|---------------------|----------|---|---|---|
| Q9NZL9-5 | Isoform 5 of Methionine adenosyltransferase 2 subunit beta         | MAT2B               | 0.02745  | 1 |   | x |
| Q9UHD9   | Ubiquilin-2  | UBQLN2              | 0.02898  | 1 | X | x |
| Q96NH3   | Protein broad-minded   | BROMI;C6orf170      | 0.011429 | 1 | X | x |
| A6NKX1   | Putative uncharacterized protein CXorf52B                          | NCRNA00204          | 0.028229 | 1 | X |   |
| B4DR52   | Histone H2B;Histone H2B type 1-K                                   | HIST2H2BF;HIST1H2BK | 4.95E-25 | 1 | X | X |
| B4DYB5   | cDNA FLJ53372  | SPAG16              | 0.025535 | 1 |   |   |
| B7Z8Z7   | RNA-binding protein 47   | RBM47               | 0.02829  | 1 |   | x |
| P06454   | Prothymosin alpha;Thymosin alpha-1                                 | PTMA                | 0.001343 | 1 | X | X |
| Q9UJ42   | Probable G-protein coupled receptor 160                            | GPR160              | 0.009327 | 1 | X | x |
| C9JH18   | PERQ amino acid-rich with GYF domain-containing protein 2          | GIGYF2              | 0.000158 | 2 |   | x |
| C9JSC0   | POU domain, class 6, transcription factor 2                        | POU6F2              | 0.029639 | 1 |   | x |
| Q96BQ5   | Coiled-coil domain-containing protein 127                          | CCDC127             | 0.031674 | 1 | X | x |
| E7ENK3   | Glutamate receptor, ionotropic kainate 1                           | GRIK1               | 0.027858 | 1 |   | x |
| Q13201   | Multimerin-1;Platelet glycoprotein Ia*;155 kDa platelet multimerin | MMRN1               | 0.024454 | 1 | X | x |
| P04406   | Glyceraldehyde-3-phosphate dehydrogenase                           | GAPDH               | 2.45E-07 | 3 | X | X |
| P11532   | Dystrophin   | DMD                 | 0.015725 | 1 |   | x |
| E9PQB1   | Serine/arginine repetitive matrix protein 5                        | SRRM5               | 0.015627 | 1 | X | x |
| Q96DR8   | Mucin-like protein 1   | MUCL1               | 0.010421 | 1 | X | X |
| F8WAU4   | Elongation factor G, mitochondrial                                 | GFM1                | 0.014496 | 1 | X | x |
| Q9P0W5-4 | Schwannomin-interacting protein 1                                  | SCHIP1              | 0.028326 | 1 |   | x |
| H0Y3T1   | Protein CROC-4   | C1orf61             | 0.031572 | 1 |   | x |
| P14618   | Pyruvate kinase isozymes M1/M2;Pyruvate kinase                     | PKM2                | 0.000118 | 2 | X | X |
| Q96ST8   | Centrosomal protein of 89 kDa                                      | CEP89               | 0.03077  | 1 |   | x |
| K7EKM7   | Small integral membrane protein 24                                 | SMIM24              | 0.009651 | 1 | X | x |
| M0QXR7   | Leucine-rich repeat-containing protein 63                          | LRRC63              | 0.020219 | 1 |   |   |
| O00555   | Voltage-dependent P/Q-type calcium channel subunit alpha-1A        | CACNA1A             | 0.000283 | 2 | X |   |
| O75553-5 | Disabled homolog 1   | DAB1                | 0.029822 | 1 |   | x |
| O75665   | Oral-facial-digital syndrome 1 protein                             | OFD1                | 0.000409 | 2 |   | x |
| P16104   | Histone H2A.x;Histone H2A type 1-A                                 | H2AFX;HIST1H2AA     | 1.95E-29 | 2 | X | X |
| P05089   | Arginase-1   | ARG1                | 0.000272 | 2 | X | X |
| P06702   | Protein S100-A9  | S100A9              | 8.91E-09 | 3 | X | X |

|          |  |               |          |   |   |   |
|----------|--|---------------|----------|---|---|---|
| P06733   | Alpha-enolase  | ENO1          | 2.54E-06 | 2 | X | X |
| P14923   | Junction plakoglobin   | JUP           | 5.14E-15 | 2 | X | X |
| P15924   | Desmoplakin  | DSP           | 3.07E-21 | 7 | X | X |
| P31944   | Caspase-14;Caspase-14 subunit p19;Caspase-14 subunit p10         | CASP14        | 3.99E-26 | 4 | X | X |
| P48431   | Transcription factor SOX-2                                       | SOX2          | 0.027433 | 1 | X | x |
| P51654   | Glypican-3;Secreted glypican-3                                   | GPC3          | 0.031467 | 1 |   | x |
| P52740   | Zinc finger protein 132  | ZNF132        | 0.002584 | 1 | X | x |
| P62805   | Histone H4   | HIST1H4A      | 0.006819 | 1 | X | X |
| Q05639   | Elongation factor 1-alpha 2;Elongation factor 1-alpha 1;         | EEF1A2;EEF1A1 | 0.000444 | 1 | X | X |
| Q01469   | Fatty acid-binding protein, epidermal                            | FABP5         | 1.08E-24 | 2 | X | X |
| Q08188   | Protein-glutamine gamma-glutamyltransferase E;                   | TGM3          | 0.011907 | 1 |   | X |
| Q08554   | Desmocollin-1  | DSC1          | 2.80E-06 | 3 | X | X |
| Q12778   | Forkhead box protein O1  | FOXO1         | 0.007693 | 1 |   | x |
| Q13164-4 | Mitogen-activated protein kinase 7                               | MAPK7         | 0.027772 | 1 | X | x |
| Q13422-4 | DNA-binding protein Ikaros                                       | IKZF1         | 0.021892 | 1 |   |   |
| Q14764   | Major vault protein  | MVP           | 1.25E-07 | 4 | X | X |
| Q15147   | 1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta-4 | PLCB4         | 1.58E-19 | 2 |   | x |
| Q5D862   | Filaggrin-2  | FLG2          | 0.000426 | 2 |   | X |
| Q5HYA8   | Meckelin   | TMEM67        | 0.010896 | 1 |   | x |
| Q5HYK7   | SH3 domain-containing protein 19                                 | SH3D19        | 0.012887 | 1 | X | x |
| Q5J5C9   | Beta-defensin 121  | DEFB121       | 0.031279 | 1 |   | x |
| Q5T750   | Skin-specific protein 32   | XP32          | 0.016572 | 1 |   | X |
| Q5TA76   | Late cornified envelope protein 3A                               | LCE3A         | 0.024214 | 1 |   |   |
| Q9UJQ4   | Sal-like protein 4   | SALL4         | 0.024516 | 1 |   | x |
| Q86YZ3   | Hornerin   | HRNR          | 3.22E-10 | 2 | X | X |
| Q8N139   | ATP-binding cassette sub-family A member 6                       | ABCA6         | 0.014869 | 1 | X | x |
| Q9BTT4   | Mediator of RNA polymerase II transcription subunit 10           | MED10         | 0.003212 | 1 | X | x |
| Q9H6A9   | Pecanex-like protein 3   | PCNXL3        | 0.018665 | 1 |   | x |

**Table S4: Functional analysis of differentially expressed proteins.** The proteins were manually grouped according to their function in biological processes taking into account, besides GOMF, GOBP and GOCC, also keywords and annotations known from the literature. It was therefore possible to ascribe each protein to 6 major groups: 1) fertilization and embryo implantation, 2) embryo development, 3) nucleic acid interaction, 4) signaling, 5) uncharacterized proteins. Proteins excluded from the groups listed above were categorized as “other”. The table lists only the proteins present and quantified in at least 3 out of 5 technical repeats using a false positive rate of 0.05 and minimum peptide length 9 amino acids, reporting bold the proteins identified with FDR 0.01.

| Under expressed in B or present only in A                   |             |                       | Over expressed or present only in B   |           |                |  |
|---|-------------|-----------------------|---|-----------|----------------|--|
|   | Gene Name   | Protein name          |   | Gene name | Protein name   |  |
| <b>Fertilization, Embryo pre-implantation, Implantation</b> | only A      | C17orf75              |   | only B    | DEFB121        | Beta-defensin 121                      |
|   | only A      | KLK12                 | Kallikrein-12   |           |                |  |
|   | only A      | LAMC1                 | Laminin subunit gamma-1   |           |                |  |
|   | only A      | PSMD4                 | 26S proteasome non-ATPase regulatory subunit 4                              |           |                |  |
|   | only A      | RANGAP1               | Ran GTPase-activating protein 1   |           |                |  |
|   |             | <b>FAAH2</b>          | Fatty-acid amide hydrolase 2  |           |                |  |
|   |             | <b>CALM2;CALM1</b>    | Calmodulin  |           |                |  |
|   |             | UCHL1                 | Ubiquitin carboxyl-terminal hydrolase isozyme L1                            |           |                |  |
|   | <b>CSTA</b> | Cystatin-A            |   |           |                |  |
| <b>Embryo development</b>                                   | only A      | <b>AMY1A;AMY2A...</b> | Alpha-amylase 1;Pancreatic alpha-amylase;Alpha-amylase 2B                   | only B    | DAB1           | Isoform DAB553 of Disabled homolog 1   |
|   | only A      | ATR                   | Serine/threonine-protein kinase ATR   | only B    | OFD1           | Oral-facial-digital syndrome 1 protein |
|   | only A      | KIF21A                | Kinesin-like protein KIF21A   |           |                |  |
|   | only A      | <b>LTF</b>            | Lactotransferrin;Kaliocin-1;Lactoferroxin-A;Lactoferroxin-B;Lactoferroxin-C | only B    | <b>SOX2</b>    | Transcription factor SOX-2             |
|   | only A      | MSI2                  | RNA-binding protein Musashi homolog 2                                       | only B    | GPC3           | Glypican-3;Secreted glypican-3         |
|   | only A      | NR0B1                 | Nuclear receptor subfamily 0 group B member 1                               | only B    | <b>MUCL1</b>   | Mucin-like protein 1                   |
|   | only A      | SLC4A5                | Electrogenic sodium bicarbonate cotransporter 4                             | only B    | CEP89          | Centrosomal protein of 89 kDa          |
|   | only A      | UNC45A                | Protein unc-45 homolog A  |           | BROM1;C6orf170 | Protein broad-minded                   |
|   |             | <b>DSG1</b>           | Desmoglein-1  |           | <b>UBQLN2</b>  | Ubiquilin-2                            |
|   |             | <b>GALNT9</b>         | Polypeptide N-acetylgalactosaminyltransferase 9                             |           | MYO1E          | Unconventional myosin-Ie               |

|                      |   |     |                       |
|----------------------|---|-----|-----------------------|
| SLC2A5               | Solute carrier family 2, facilitated glucose transporter member 5 | EVL | Ena/VASP-like protein |
| CDSN                 | Corneodesmosin  | JUP | Junction plakoglobin  |
| <b>INS-IGF2;INS</b>  | Insulin;Insulin B chain;Insulin A chain                           |     |                       |
| <b>S100A8</b>        | Protein S100-A8   |     |                       |
| <b>BAIAP2</b>        | Brain-specific angiogenesis inhibitor 1-associated protein 2      |     |                       |
| <b>TNRC6C</b>        | Trinucleotide repeat-containing gene 6C protein                   |     |                       |
| MEIS1                | Homeobox protein Meis1  |     |                       |
| <b>BORA;FLJ22624</b> | Protein aurora borealis   |     |                       |
| CEP250               | Centrosome-associated protein CEP250                              |     |                       |
| ALDH1A2              | Isoform 3 of Retinal dehydrogenase 2                              |     |                       |
| PAM                  | Peptidyl-glycine alpha-amidating monooxygenase...                 |     |                       |
| <b>PRSS1</b>         | Trypsin-1;Alpha-trypsin chain 1;Alpha-trypsin chain 2             |     |                       |
| <b>HMCN1</b>         | Hemicentin-1  |     |                       |
| <b>SBSN</b>          | Suprabasin  |     |                       |
| <b>LCN1;LCN1P1</b>   | Lipocalin-1;Putative lipocalin 1-like protein 1                   |     |                       |
| <b>DMBT1</b>         | Deleted in malignant brain tumors 1 protein                       |     |                       |
| <b>PMCH</b>          | Pro-MCH;Neuropeptide-glycine-glutamic acid...                     |     |                       |

|                                 |        |              |  |        |  |  |
|---------------------------------|--------|--------------|--|--------|--|--|
| <b>Nucleic Acid interaction</b> | only A | <b>PCGF1</b> | Polycomb group RING finger protein 1     | only B | <b>H2AFX;HIST1H2AA;HIST1H2AB...</b>    | Histone H2A.x;Histone H2A type 1-A;Histone H2A type 1-B/E;Histone H2A type 1.....                                  |
|                                 |        | <b>NRBF2</b> | Nuclear receptor-binding factor 2        | only B | <b>HIST1H4A</b>                        | Histone H4   |
|                                 |        | NFIB         | Nuclear factor 1;Nuclear factor 1 B-type | only B | <b>EEF1A2;EEF1A1;EEF1A1P5</b>          | Elongation factor 1-alpha 2;Elongation factor 1-alpha 1;Putative elongation factor 1-alpha-like 3                  |
|                                 |        |              |  | only B | <b>HIST2H2BF;HIST1H2BK...</b><br>BEND4 | Histone H2B;Histone H2B type 1-K;Histone H2B type 1-J;Histone H2B type 1-O.....<br>BEN domain-containing protein 4 |

|                   |        |                 |   |        |               |  |
|-------------------|--------|-----------------|---|--------|---------------|--|
| <b>Signalling</b> | only A | <b>DOCK4</b>    | Dedicator of cytokinesis protein 4                              | only B | <b>PLCB4</b>  | 1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta-4 |
|                   | only A | <b>EIF2AK1</b>  | Eukaryotic translation initiation factor 2-alpha kinase 1       | only B | <b>GPR160</b> | Probable G-protein coupled receptor 160                          |
|                   | only A | <b>EIF4EBP2</b> | Eukaryotic translation initiation factor 4E-binding protein 2   |        | <b>SH2B1</b>  | SH2B adapter protein 1   |
|                   | only A | <b>TREML2</b>   | Trem-like transcript 2 protein                                  |        | <b>PPP1R8</b> | Isoform Gamma of Nuclear inhibitor of protein phosphatase 1      |
|                   | only A | <b>ZAK</b>      | Isoform 2 of Mitogen-activated protein kinase kinase kinase MLT |        |               |  |
|                   | only A | <b>ZFAND6</b>   | AN1-type zinc finger protein 6                                  |        |               |  |

|                        |        |                                 |  |        |              |  |
|------------------------|--------|---------------------------------|--|--------|--------------|--|
|                        | only A | NGFRAP1<br><b>PTCHI</b>         | Isoform 2 of Protein BEX3<br>Isoform S of Protein patched homolog 1  |        |              |  |
| <b>Uncharacterized</b> | only A | <b>C3orf51</b>                  | Putative uncharacterized protein C3orf51                             | only B | NCRNA00204   | Putative uncharacterized protein CXorf52B                  |
|                        | only A | <b>XKR4</b>                     | XK-related protein 4   | only B | PCNXL3       | Pecanex-like protein 3                                     |
|                        | only A | FAM189A1                        | Isoform 2 of Protein FAM189A1  |        |              |  |
|                        |        | C9orf43                         | Uncharacterized protein C9orf43                                      |        |              |  |
|                        |        | KIAA1671                        | Uncharacterized protein KIAA1671                                     |        |              |  |
|                        |        | <b>HEATR8</b>                   | Maestro heat-like repeat-containing protein family member 7          |        |              |  |
|                        |        | <b>CCDC77</b>                   | Coiled-coil domain-containing protein 77                             |        |              |  |
|                        |        | <b>MORC4</b>                    | MORC family CW-type zinc finger protein 4                            |        |              |  |
| <b>Others</b>          | only A | ISCU                            | Isoform 2 of Iron-sulfur cluster assembly enzyme ISCU, mitochondrial | only B | C1orf61      | Protein CROC-4   |
|                        | only A | <b>KCTD2</b>                    | BTB/POZ domain-containing protein KCTD2                              |        | MAT2B        | Isoform 5 of Methionine adenosyltransferase 2 subunit beta |
|                        | only A | KLHL17                          | Kelch-like protein 17  |        | <b>CHGB</b>  | Secretogranin-1;GAWK peptide;CCB peptide                   |
|                        | only A | MAGEC3                          | MAGEC3 protein   |        | SLC35B3      | Adenosine 3-phospho 5-phosphosulfate transporter 2         |
|                        | only A | <b>MPDZ</b>                     | Multiple PDZ domain protein  |        | <b>SSFA2</b> | Sperm-specific antigen 2                                   |
|                        | only A | SCGB2A1                         | Mammaglobin-B  |        | AANAT        | Serotonin N-acetyltransferase                              |
|                        | only A | <b>SPINK5</b>                   | Isoform long of Serine protease inhibitor Kazal-type 5               |        | <b>KPRP</b>  | Keratinocyte proline-rich protein                          |
|                        | only A | AKT1S1                          | Isoform 3 of Proline-rich AKT1 substrate 1                           |        | <b>PDZD4</b> | PDZ domain-containing protein 4                            |
|                        |        | <b>LCN10</b>                    | Epididymal-specific lipocalin-10                                     |        | EPG5         | Ectopic P granules protein 5 homolog                       |
|                        |        | <b>UBC;UBB;RPS27A;UBA52</b>     | Polyubiquitin-C;Ubiquitin;Polyubiquitin-B;Ubiquitin.....             |        | INADL        | Isoform 5 of InaD-like protein                             |
|                        |        | <b>MFSD9</b>                    | Major facilitator superfamily domain-containing protein 9            |        |              |  |
|                        |        | <b>IFNA10;IFNA21;IFNA17....</b> | Interferon alpha-10;Interferon alpha-21....                          |        |              |  |
|                        |        | <b>RNF223</b>                   | RING finger protein 223  |        |              |  |
|                        |        | CEP112                          | Centrosomal protein of 112 kDa                                       |        |              |  |
|                        |        | <b>LYZ</b>                      | Lysozyme C   |        |              |  |
|                        |        | <b>DCD</b>                      | Dermeidin;Survival-promoting peptide;DCD-1                           |        |              |  |
|                        |        | CALML5                          | Calmodulin-like protein 5  |        |              |  |
|                        |        | <b>S100A7;S100A7A</b>           | Protein S100-A7;Protein S100-A7A                                     |        |              |  |
|                        |        | LETM1                           | LETM1 and EF-hand domain-containing protein 1, mitochondrial         |        |              |  |



**ADAMTS17**

A disintegrin and metalloproteinase with  
thrombospondin motifs 17

**GOLM1**

Golgi membrane protein 1

**PIP**

Prolactin-inducible protein

---

**Table 1: List of the proteins increased in group B.** Proteins were considered increased or decreased if they showed significant t-test difference (cut-off at 5% permutation-based False Discovery Rate) between group A and B. Previous identification: in the Table the proteins previously identified in human blastocoel fluid (16, 17) are indicated with a capital X, while x indicates gene products previously experimentally detected at the protein level as secreted and/or present in extracellular region and/or in exosome and/or in microvesicles according to public available databases and dataset ([www.neXtProt.org](http://www.neXtProt.org), [www.exocarta.org](http://www.exocarta.org), [microvesicle.org](http://microvesicle.org), [geneontology.org](http://geneontology.org), [www.proteinatlas.org](http://www.proteinatlas.org), 29) or at the gene expression level from microarray data (16, 92) or at the transcription level in blastocystis or embryonic tissue according to <http://genomewidepdb.proteomix.org> or <https://www.ncbi.nlm.nih.gov/geoprofiles> databases. The proteins ascribed to fertilization and embryo implantation or nucleic acid interaction categories of Fig. 4 are highlighted in light and dark grey, respectively. The table lists only the proteins present and quantified in at least 3 out of 5 technical repeats using a false positive rate of 0.05 and minimum peptide length 9 amino acids, reporting in a column the proteins identified with FDR 0.01.

| <b>Protein IDs</b> | <b>Protein names</b>  | <b>Gene names</b> | <b>t-test Difference</b> | <b>FDR 1%</b> | <b>Previous identification</b> |
|--------------------|---|-------------------|--------------------------|---------------|--------------------------------|
| Q96NH3             | Protein broad-minded  | BROMI;C6orf170    | 1.612                    |               | x                              |
| Q9UHD9             | Ubiquilin-2   | UBQLN2            | 1.451                    | X             | x                              |
| Q9NZL9-5           | Isoform 5 of Methionine adenosyltransferase 2 subunit beta  | MAT2B             | 1.217                    |               | x                              |
| P05060             | Secretogranin-1;GAWK peptide;CCB peptide                    | CHGB              | 1.159                    | X             | x                              |
| Q9H1N7             | Adenosine 3-phospho 5-phosphosulfate transporter 2          | SLC35B3           | 0.932                    |               | x                              |
| Q6ZU67             | BEN domain-containing protein 4                             | BEND4             | 0.837                    |               | x                              |
| P28290             | Sperm-specific antigen 2                                    | SSFA2             | 0.792                    | X             | x                              |
| J3KMZ5             | Serotonin N-acetyltransferase                               | AANAT             | 0.717                    |               | x                              |
| H0YNB0             | Unconventional myosin-Ie                                    | MYO1E             | 0.712                    |               | x                              |
| H3BUN7             | SH2B adapter protein 1                                      | SH2B1             | 0.617                    | X             | x                              |
| Q5T749             | Keratinocyte proline-rich protein                           | KPRP              | 0.600                    | X             | X                              |
| Q76G19             | PDZ domain-containing protein 4                             | PDZD4             | 0.582                    | X             | x                              |
| Q9HCE0             | Ectopic P granules protein 5 homolog                        | EPG5              | 0.561                    |               | x                              |
| Q12972-3           | Isoform Gamma of Nuclear inhibitor of protein phosphatase 1 | PPP1R8            | 0.553                    |               | x                              |
| G3V314             | Ena/VASP-like protein                                       | EVL               | 0.491                    |               | x                              |
| Q8NI35-5           | Isoform 5 of InaD-like protein                              | INADL             | 0.450                    |               | x                              |
| P14923             | Junction plakoglobin  | JUP               | 0.445                    | X             | x                              |

**Table 2: List of the proteins decreased in group B.** Proteins were considered increased or decreased if they showed significant t-test difference (cut-off at 5% permutation-based False Discovery Rate) between group A and B. Previous identification: in the Table the proteins previously identified in human blastocoel fluid (16, 17) are indicated with a capital X, while x indicates gene products previously experimentally detected at the protein level as secreted and/or present in extracellular region and/or in exosome and/or in microvesicles according to public available databases and dataset (www.neXtProt.org, www.exocarta.org, microvesicle.org, geneontology.org, www.proteinatlas.org, 29) or at the gene expression level from microarray data (16, 92) or at the transcription level in blastocystis or embryonic tissue according to <http://genomewidepdb.proteomix.org> or <https://www.ncbi.nlm.nih.gov/geoprofiles> databases. The proteins ascribed to fertilization and embryo implantation or nucleic acid interaction categories of Fig. 4 are highlighted in light and dark grey, respectively. The table lists only the proteins present and quantified in at least 3 out of 5 technical repeats using a false positive rate of 0.05 and minimum peptide length 9 amino acids, reporting in a column the proteins identified with FDR 0.01.

| Protein IDs | Protein names   | Gene names    | t-test Difference | FDR 1% | Previous identification |
|-------------|---|---------------|-------------------|--------|-------------------------|
| Q02413      | Desmoglein-1  | DSG1          | -0.158            | X      | X                       |
| J3KNN1      | Polypeptide N-acetylgalactosaminyltransferase 9                   | GALNT9        | -0.194            | X      |                         |
| Q8TAL5      | Uncharacterized protein C9orf43                                   | C9orf43       | -0.341            |        | x                       |
| P22732      | Solute carrier family 2, facilitated glucose transporter member 5 | SLC2A5        | -0.349            |        | x                       |
| P01566      | Interferon alpha-10;Interferon alpha-21                           | IFNA10;IFNA21 | -0.368            | X      | x                       |
| G8JLG2      | Corneodesmosin  | CDSN          | -0.379            |        | X                       |
| F8WCM5      | Insulin;Insulin B chain;Insulin A chain                           | INS-IGF2;INS  | -0.386            | X      | X                       |
| Q8NBP5      | Major facilitator superfamily domain-containing protein 9         | MFSD9         | -0.420            | X      | x                       |
| Q13635-4    | Isoform S of Protein patched homolog 1                            | PTCH1         | -0.421            | X      | x                       |
| P05109      | Protein S100-A8   | S100A8        | -0.438            | X      | X                       |
| F8W878      | Brain-specific angiogenesis inhibitor 1-associated protein 2      | BAIAP2        | -0.464            | X      | x                       |
| Q9HCJ0      | Trinucleotide repeat-containing gene 6C protein                   | TNRC6C        | -0.476            | X      | x                       |
| Q6JVE6      | Epididymal-specific lipocalin-10                                  | LCN10         | -0.504            |        | x                       |
| P0CG48      | Polyubiquitin-C;Ubiquitin;Polyubiquitin-B                         | UBC;UBB       | -0.539            | X      | X                       |
| F8W8U3      | Homeobox protein Meis1  | MEIS1         | -0.558            |        | x                       |
| E7ERA6      | RING finger protein 223   | RNF223        | -0.558            | X      |                         |
| Q96F24-2    | Nuclear receptor-binding factor 2                                 | NRBF2         | -0.634            | X      | x                       |
| Q9BY89      | Uncharacterized protein KIAA1671                                  | KIAA1671      | -0.645            |        | x                       |
| Q8N8E3      | Centrosomal protein of 112 kDa                                    | CEP112        | -0.713            |        | x                       |
| P61626      | Lysozyme C  | LYZ           | -0.760            | X      | X                       |
| C9K0L2      | Maestro heat-like repeat-containing protein family member 7       | HEATR8        | -0.760            | X      |                         |
| Q6PGQ7      | Protein aurora borealis   | BORA;FLJ2624  | -0.770            | X      | x                       |
| P81605      | Dermcidin;Survival-promoting peptide;DCD-1                        | DCD           | -0.793            | X      | X                       |
| Q9NZN1      | Calmodulin-like protein 5   | CALML5        | -0.829            |        | X                       |
| Q9BV73      | Centrosome-associated protein CEP250                              | CEP250        | -0.991            |        | X                       |

|          |   |                |        |   |   |
|----------|---|----------------|--------|---|---|
| O00712   | Nuclear factor 1;Nuclear factor 1 B-type                          | NFIB           | -1.041 |   | x |
| O94788-3 | Isoform 3 of Retinal dehydrogenase 2                              | ALDH1A2        | -1.065 |   | x |
| P19021   | Peptidyl-glycine alpha-amidating monooxygenase                    | PAM            | -1.140 |   | x |
| Q6GMR7   | Fatty-acid amide hydrolase 2                                      | FAAH2          | -1.178 | X | x |
| H0Y7A7   | Calmodulin  | CALM2;CALM1    | -1.209 | X | x |
| P31151   | Protein S100-A7;Protein S100-A7A                                  | S100A7;S100A7A | -1.240 | X | X |
| Q96RW7   | Hemicentin-1  | HMCN1          | -1.250 | X | x |
| Q9BR77   | Coiled-coil domain-containing protein 77                          | CCDC77         | -1.315 | X | x |
| Q6UWP8   | Suprabasin  | SBSN           | -1.407 | X | X |
| P31025   | Lipocalin-1;Putative lipocalin 1-like protein 1                   | LCN1;LCN1P1    | -1.419 | X | X |
| Q9UGM3   | Deleted in malignant brain tumors 1 protein                       | DMBT1          | -1.508 | X | X |
| O95202   | LETM1 and EF-hand domain-containing protein 1, mitochondrial      | LETM1          | -1.622 |   | x |
| D6R956   | Ubiquitin carboxyl-terminal hydrolase isozyme L1                  | UCHL1          | -1.796 |   | X |
| Q8TE56   | A disintegrin and metalloproteinase with thrombospondin motifs 17 | ADAMTS17       | -1.825 | X | x |
| Q8TE76   | MORC family CW-type zinc finger protein 4                         | MORC4          | -1.857 | X | x |
| P20382   | Pro-MCH;Neuropeptide-glycine-glutamic acid                        | PMCH           | -2.149 | X | x |
| P01040   | Cystatin-A  | CSTA           | -2.310 | X | X |
| Q8NBJ4   | Golgi membrane protein 1  | GOLM1          | -2.427 | X | x |
| P12273   | Prolactin-inducible protein                                       | PIP            | -2.704 | X | X |

**Table 3: List of the proteins present only in group A.** Previous identification: in the Table the proteins previously identified in human blastocoel fluid (16, 17) are indicated with a capital X, while x indicates gene products previously experimentally detected at the protein level as secreted and/or present in extracellular region and/or in exosome and/or in microvesicles according to public available databases and dataset ([www.neXtProt.org](http://www.neXtProt.org), [www.exocarta.org](http://www.exocarta.org), [microvesicle.org](http://microvesicle.org), [geneontology.org](http://geneontology.org), [www.proteinatlas.org](http://www.proteinatlas.org), 29) or at the gene expression level from microarray data (16, 92) or at the transcription level in blastocystis or embryonic tissue according to <http://genomewidepdb.proteomix.org> or <https://www.ncbi.nlm.nih.gov/geoprofiles> databases. The proteins ascribed to fertilization and embryo implantation or nucleic acid interaction categories of Fig. 4 are highlighted in light and dark grey, respectively. The table lists only the proteins present and quantified in at least 3 out of 5 technical repeats using a false positive rate of 0.05 and minimum peptide length 9 amino acids, reporting in a column the proteins identified with FDR 0.01.

| Protein IDs | Protein names   | Gene names  | FDR 1% | Previous identification |
|-------------|---|-------------|--------|-------------------------|
| P04745      | Alpha-amylase 1;Pancreatic alpha-amylase                                    | AMY1A;AMY2A |        | X                       |
| Q13535      | Serine/threonine-protein kinase ATR   | ATR         |        | x                       |
| J9JIC5      | Protein Njmu-R1   | C17orf75    |        | x                       |
| O76042      | Putative uncharacterized protein C3orf51                                    | C3orf51     | X      |                         |
| H0Y7H7      | Dedicator of cytokinesis protein 4  | DOCK4       |        | x                       |
| B4DIP4      | Eukaryotic translation initiation factor 2-alpha kinase 1                   | EIF2AK1     |        | x                       |
| Q13542      | Eukaryotic translation initiation factor 4E-binding protein 2               | EIF4EBP2    | X      | x                       |
| Q9H1K1-2    | Isoform 2 of Iron-sulfur cluster assembly enzyme ISCU, mitochondrial        | ISCU        |        | x                       |
| H0Y3B9      | BTB/POZ domain-containing protein KCTD2                                     | KCTD2       | X      | x                       |
| Q7Z4S6      | Kinesin-like protein KIF21A   | KIF21A      |        | x                       |
| Q6TDP4      | Kelch-like protein 17   | KLHL17      |        | x                       |
| Q9UKR0      | Kallikrein-12   | KLK12       |        | x                       |
| P11047      | Laminin subunit gamma-1   | LAMC1       |        | x                       |
| P02788      | Lactotransferrin;Kaliocin-1;Lactoferroxin-A;Lactoferroxin-B;Lactoferroxin-C | LTF         | X      | X                       |
| Q3SYA6      | MAGEC3 protein  | MAGEC3      |        |                         |
| F5H1U9      | Multiple PDZ domain protein   | MPDZ        | X      | x                       |
| Q96DH6      | RNA-binding protein Musashi homolog 2                                       | MSI2        |        | x                       |
| P51843      | Nuclear receptor subfamily 0 group B member 1                               | NR0B1       |        | x                       |
| Q9BSM1      | Polycomb group RING finger protein 1  | PCGF1       | X      | x                       |
| Q5VWC4      | 26S proteasome non-ATPase regulatory subunit 4                              | PSMD4       |        | x                       |
| P46060      | Ran GTPase-activating protein 1   | RANGAP1     |        | x                       |
| O75556      | Mammaglobin-B   | SCGB2A1     |        | x                       |
| C9JNL6      | Electrogenic sodium bicarbonate cotransporter 4                             | SLC4A5      |        | x                       |
| Q9NQ38      | Serine protease inhibitor Kazal-type 5                                      | SPINK5      | X      | x                       |
| Q5T2D2      | Trem-like transcript 2 protein  | TREML2      |        | x                       |
| Q9H3U1      | Protein unc-45 homolog A  | UNC45A      |        | x                       |
| Q5GH76      | XK-related protein 4  | XKR4        | X      |                         |
| Q9NYL2-2    | Isoform 2 of Mitogen-activated protein kinase kinase kinase MLT             | ZAK         |        | x                       |
| Q6FIF0      | AN1-type zinc finger protein 6  | ZFAND6      |        | x                       |

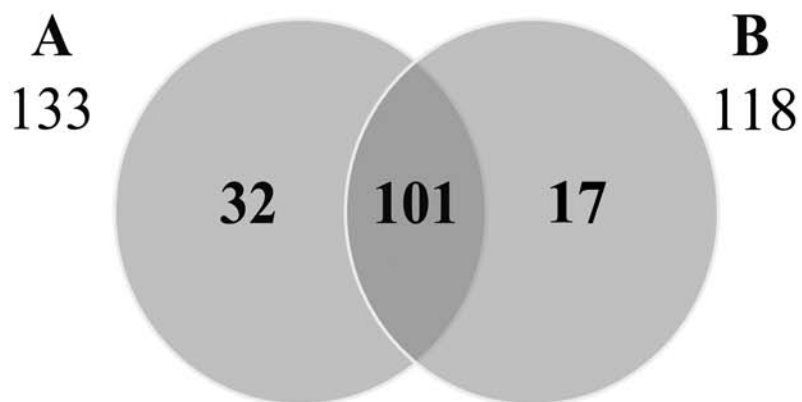
|          |  |          |   |
|----------|--|----------|---|
| O60320-2 | Isoform 2 of Protein FAM189A1              | FAM189A1 | x |
| Q00994-2 | Isoform 2 of Protein BEX3                  | NGFRAP1  | x |
| Q96B36-3 | Isoform 3 of Proline-rich AKT1 substrate 1 | AKT1S1   | x |

---

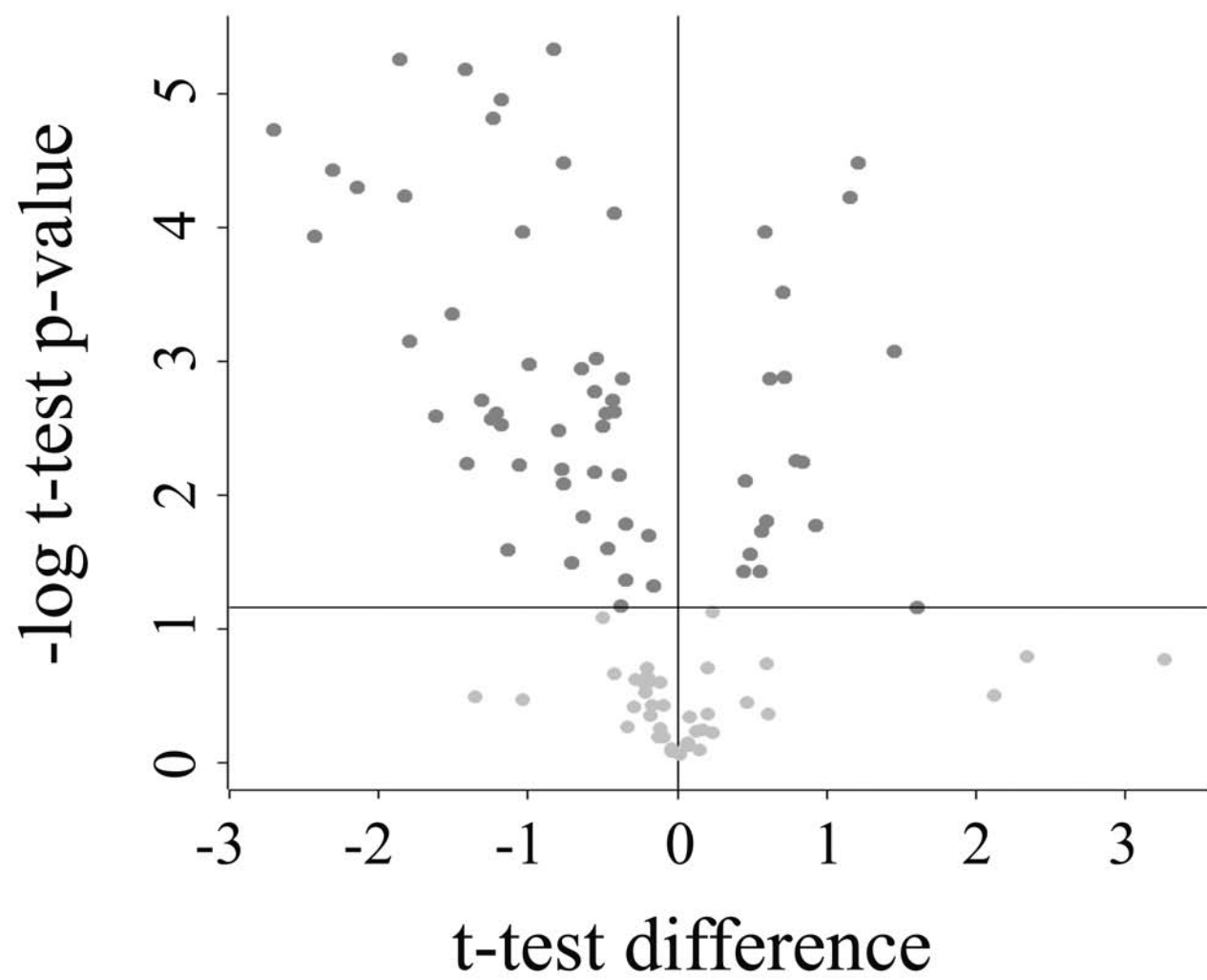
**Table 4: List of the proteins present only in group B.** Previous identification: in the Table the proteins previously identified in human blastocoel fluid (16, 17) are indicated with a capital X, while x indicates gene products previously experimentally detected at the protein level as secreted and/or present in extracellular region and/or in exosome and/or in microvesicles according to public available databases and dataset ([www.neXtProt.org](http://www.neXtProt.org), [www.exocarta.org](http://www.exocarta.org), [microvesicle.org](http://microvesicle.org), [geneontology.org](http://geneontology.org), [www.proteinatlas.org](http://www.proteinatlas.org), 29) or at the gene expression level from microarray data (16, 92) or at the transcription level in blastocystis or embryonic tissue according to <http://genomewidepdb.proteomix.org> or <https://www.ncbi.nlm.nih.gov/geoprofiles> databases. The proteins ascribed to fertilization and embryo implantation or nucleic acid interaction categories of Fig. 4 are highlighted in light and dark grey, respectively. The table lists only the proteins present and quantified in at least 3 out of 5 technical repeats using a false positive rate of 0.05 and minimum peptide length 9 amino acids, reporting in a column the proteins identified with FDR 0.01.

| <b>Protein IDs</b> | <b>Protein names</b>  | <b>Gene names</b>          | <b>FDR 1%</b> | <b>Previous identification</b> |
|--------------------|---|----------------------------|---------------|--------------------------------|
| A6NKX1             | Putative uncharacterized protein CXorf52B   | NCRNA00204                 |               |                                |
| O75553-5           | Isoform DAB553 of Disabled homolog 1  | DAB1                       |               | x                              |
| O75665             | Oral-facial-digital syndrome 1 protein  | OFD1                       |               | x                              |
| P16104             | Histone H2A.x;Histone H2A type 1-A  | H2AFX;HIST1H2A<br>A        | X             | X                              |
| P48431             | Transcription factor SOX-2  | SOX2                       | X             | x                              |
| P51654             | Glypican-3;Secreted glypican-3  | GPC3                       |               | x                              |
| P62805             | Histone H4  | HIST1H4A                   | X             | X                              |
| Q05639             | Elongation factor 1-alpha 2;Elongation factor 1-alpha 1;Putative elongation factor 1-alpha-like 3 | EEF1A2;EEF1A1;E<br>EF1A1P5 | X             | X                              |
| Q15147             | 1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta-4                                  | PLCB4                      |               | x                              |
| Q5J5C9             | Beta-defensin 121   | DEFB121                    |               | x                              |
| Q96DR8             | Mucin-like protein 1  | MUCL1                      | X             | X                              |
| Q96ST8             | Centrosomal protein of 89 kDa   | CEP89                      |               | x                              |
| Q9H6A9             | Pecanex-like protein 3  | PCNXL3                     |               | x                              |
| Q9UJ42             | Probable G-protein coupled receptor 160   | GPR160                     |               | x                              |
| B4DR52             | Histone H2B;Histone H2B type 1-K  | HIST2H2BF;HIST1<br>H2BK    | X             | X                              |
| H0Y3T1             | Protein CROC-4  | C1orf61                    |               | x                              |

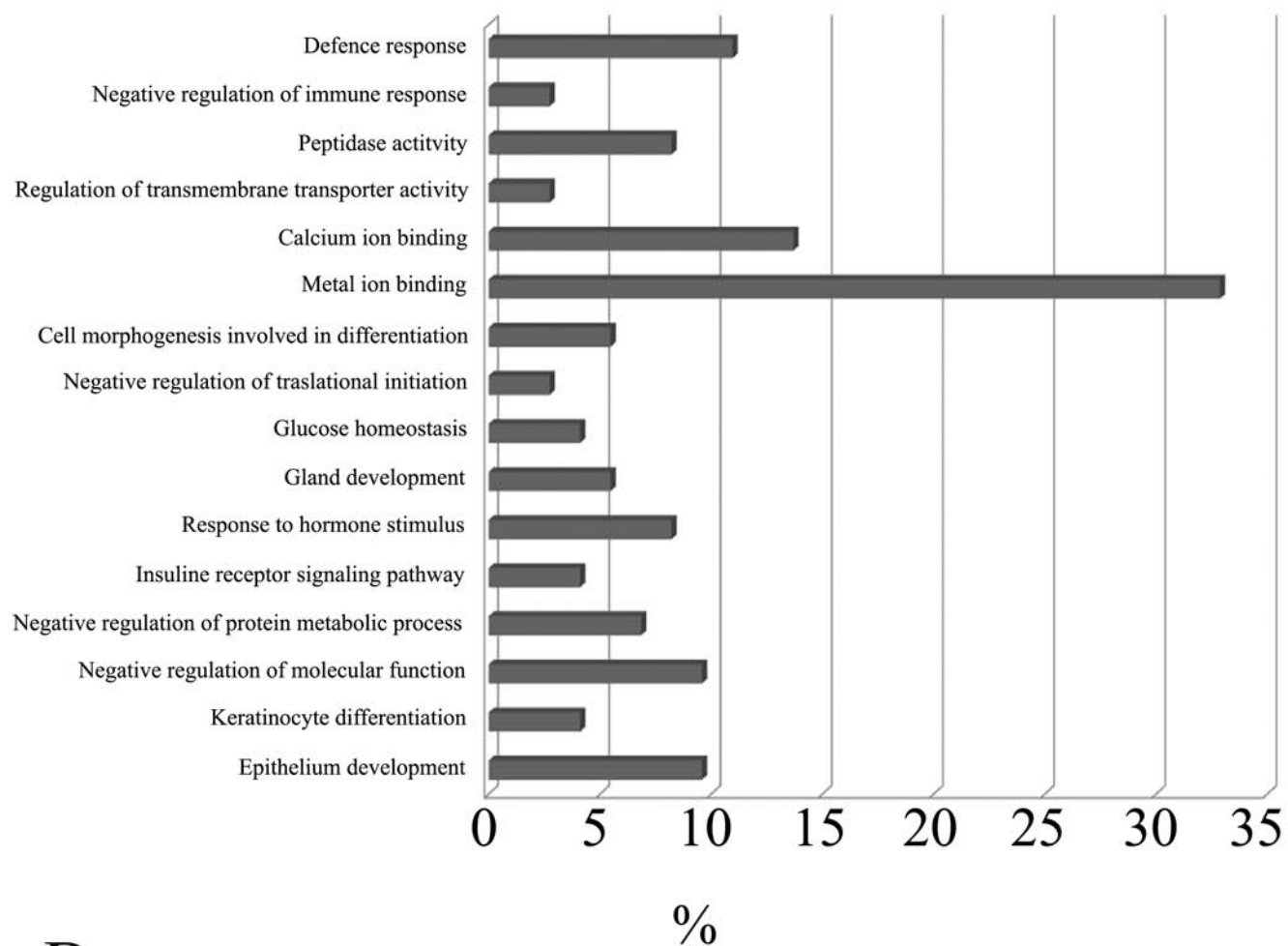
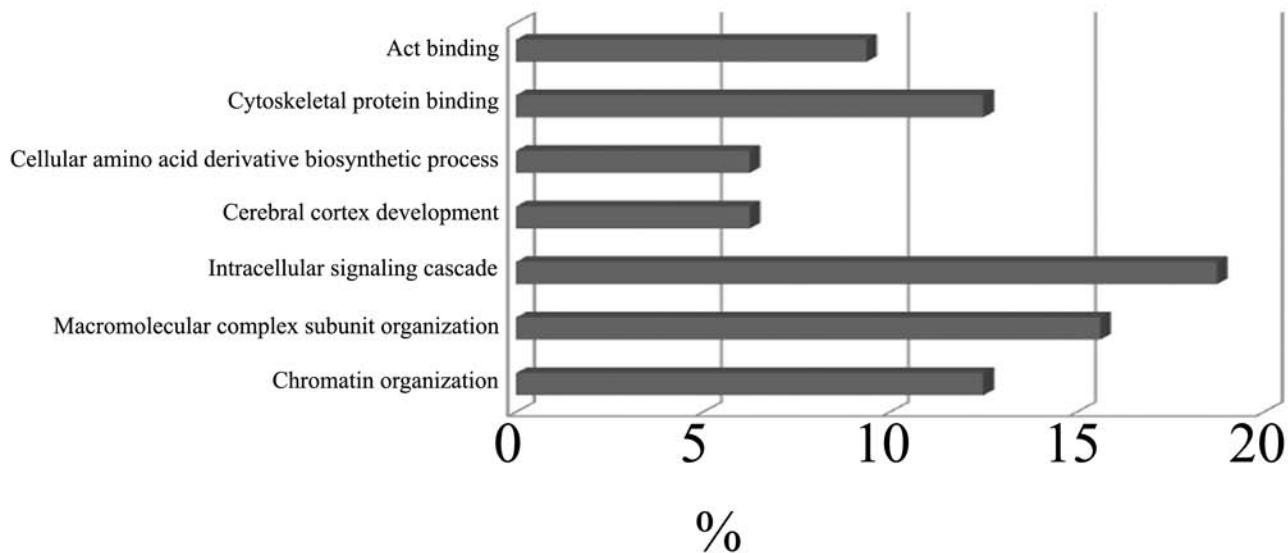
A



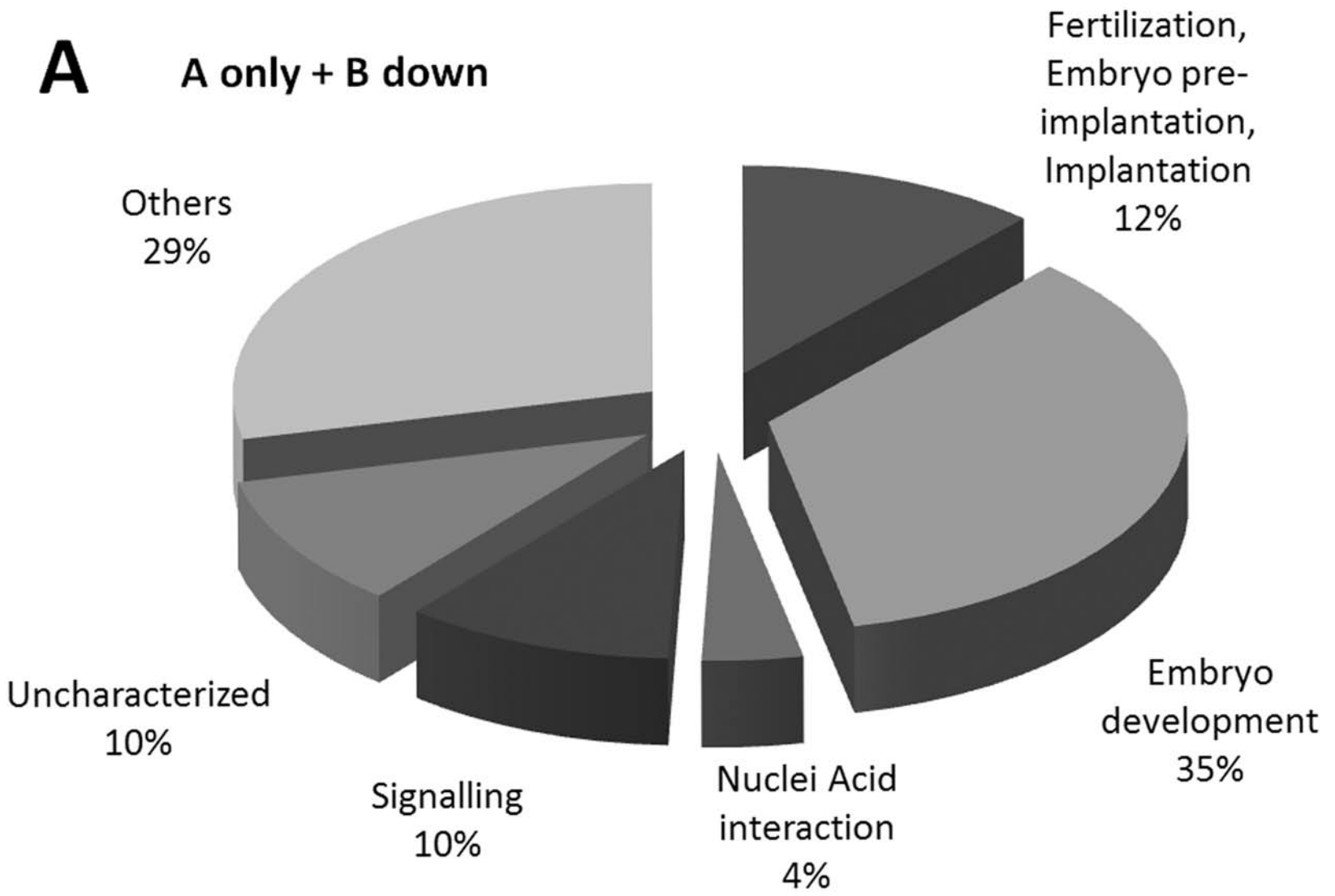
B





**A****B down + A only****B****B up + B only**

**A** A only + B down



**B** B only+B up

