Human Reproduction, Vol.36, No.3, pp. 756–770, 2021
Advance Access Publication on December 13, 2020 doi:10.1093/humrep/deaa335

human reproduction

ORIGINAL ARTICLE *Reproductive biology*

Proteome of fluid from human ovarian small antral follicles reveals insights in folliculogenesis and oocyte maturation

Indira Pla[1,2,†], Aniel Sanchez[1,2,*,†], Susanne Elisabeth Pors[3,*,†], Krzysztof Pawlowski[2,4], Roger Appelqvist[2], K. Barbara Sahlin[1,2], Liv La Cour Poulsen[5], György Marko-Varga[2,6], Claus Yding Andersen[3], and Johan Malm[1,2]

[1]Section for Clinical Chemistry, Department of Translational Medicine, Lund University, Skåne University Hospital Malmö, 205 02 Malmö, Sweden [2]Clinical Protein Science & Imaging, Biomedical Centre, Department of Biomedical Engineering, Lund University, BMC D13, 221 84 Lund, Sweden [3]Laboratory of Reproductive Biology, The Juliane Marie Centre for Women, Children and Reproduction, University Hospital of Copenhagen, 2100 Copenhagen, Denmark [4]Department of Experimental Design and Bioinformatics, Faculty of Agriculture and Biology, Warsaw University of Life Sciences SGGW, Warszawa 02-787, Poland [5]Fertility Clinic, Department of Gynaecology and Obstetrics, Zealand University Hospital, Lykkebækvej 14, 4600 Køge, Denmark [6]First Department of Surgery, Tokyo Medical University, Shinjiku-ku, Japan

*Correspondence address: Clinical Protein Science & Imaging, Biomedical Centre, Department of Biomedical Engineering, Lund University Klinikgatan 32, D13., BMC D13, 221 84 Lund, Sweden. E-mail: aniel.sanchez@med.lu.se (A.S.); Laboratory of Reproductive Biology, University Hospital of Copenhagen, Rigshospitalet, Section 5701, Henrik Harpestrengsvej 6A, DK-2100 Copenhagen, Denmark. E-mail: susanne.elisabeth.pors@regionh.dk (S.E.P.)

Submitted on March 17, 2020; resubmitted on November 3, 2020; editorial decision on November 10, 2020

STUDY QUESTION: Is it possible to identify by mass spectrometry a wider range of proteins and key proteins involved in folliculogenesis and oocyte growth and development by studying follicular fluid (FF) from human small antral follicles (hSAF)?

SUMMARY ANSWER: The largest number of proteins currently reported in human FF was identified in this study analysing hSAF where several proteins showed a strong relationship with follicular developmental processes.

WHAT IS KNOWN ALREADY: Protein composition of human ovarian FF constitutes the microenvironment for oocyte development. Previous proteomics studies have analysed fluids from pre-ovulatory follicles, where large numbers of plasma constituents are transferred through the follicular basal membrane. This attenuates the detection of low abundant proteins, however, the basal membrane of small antral follicles is less permeable, making it possible to detect a large number of proteins, and thereby offering further insights in folliculogenesis.

STUDY DESIGN, SIZE, DURATION: Proteins in FF from unstimulated hSAF (size 6.1 ± 0.4 mm) were characterised by mass spectrometry, supported by high-throughput and targeted proteomics and bioinformatics. The FF protein profiles from hSAF containing oocytes, capable or not of maturing to metaphase II of the second meiotic division during an IVM (n = 13, from 6 women), were also analysed.

PARTICIPANTS/MATERIALS, SETTING, METHODS: We collected FF from hSAF of ovaries that had been surgically removed from 31 women (~28.5 years old) undergoing unilateral ovariectomy for fertility preservation.

MAIN RESULTS AND THE ROLE OF CHANCE: In total, 2461 proteins were identified, of which 1108 identified for the first time in FF. Of the identified proteins, 24 were related to follicular regulatory processes. A total of 35 and 65 proteins were down- and up-regulated, respectively, in fluid from hSAF surrounding oocytes capable of maturing (to MII). We found that changes at the protein level occur already in FF from small antral follicles related to subsequent oocyte maturation.

LIMITATIONS, REASONS FOR CAUTION: A possible limitation of our study is the uncertainty of the proportion of the sampled follicles that are undergoing atresia. Although the FF samples were carefully aspirated and processed to remove possible contaminants, we

[†]The authors consider that the first three authors should be regarded as joint first authors.

cannot ensure the absence of some proteins derived from cellular lysis provoked by technical reasons.

WIDER IMPLICATIONS OF THE FINDINGS: This study is, to our knowledge, the first proteomics characterisation of FF from hSAF obtained from women in their natural menstrual cycle. We demonstrated that the analysis by mass spectrometry of FF from hSAF allows the identification of a greater number of proteins compared to the results obtained from previous analyses of larger follicles. Significant differences found at the protein level in hSAF fluid could predict the ability of the enclosed oocyte to sustain meiotic resumption. If this can be confirmed in further studies, it demonstrates that the viability of the oocyte is determined early on in follicular development and this may open up new pathways for augmenting or attenuating subsequent oocyte viability in the pre-ovulatory follicle ready to undergo ovulation.

STUDY FUNDING/COMPETING INTEREST(S): The authors thank the financial support from ReproUnion, which is funded by the Interreg V EU programme. No conflict of interest was reported by the authors.

TRIAL REGISTRATION NUMBER: N/A

Key words: follicular fluid / small antral follicles / proteomics / oocyte maturation / mass spectrometry

Introduction

A woman's fertility is based on the pool of resting follicles in the ovaries, i.e. the ovarian reserve. This reserve ensures the generation of menstrual cycles and release of fertilisable oocytes that may result in new offspring. Follicular development is initiated by activation of resting primordial follicles and is completed with ovulation of fully mature oocytes approximately half a year later (Gougeon, 2010). During this lengthy period of follicular development, the diameter of the follicle increases from 45 μm to ~20 mm. This process involves several developmental stages, including the formation of the follicular fluid (FF)-filled antrum that begins to form when human ovarian follicles reach a diameter around 250 μm (Rodgers and Irving-Rodgers, 2010). The FF is comprised of secretions from the oocyte and granulosa cells (GC) within the follicles and from theca cells (TC) surrounding the follicle as well as transudates from circulation filtered through the basal membrane. The basal membrane acts as a molecular filter, which means that proteins with a relatively high molecular weight can only enter the FF to a limited extent (Rodgers and Irving-Rodgers, 2010; Siu and Cheng, 2012). This is the reason why FF contains low concentrations of the high molecular weight protein fibrinogen and, therefore, does not coagulate. Additionally, the basal membrane has been described to be both charge and size selective in mouse ovaries (Hess *et al.*, 1998; Siu and Cheng, 2012). The composition of FF is highly variable and is associated with the developmental stage of follicles. In particular, FF reflects GC activity, which is regulated by gonadotropins, steroids, peptide hormones and growth factors. Anti-Müllerian hormone (AMH), part of the transforming growth factor β (TGF-β) superfamily, is present at very high concentrations in small antral follicles with a peak in follicular content around a diameter of 8 mm (Jeppesen *et al.*, 2013). Conversely, sex-steroids, such as estradiol and progesterone, accumulate at very high concentrations in the pre-ovulatory follicles, in orders of magnitude higher than in small antral follicles (Jeppesen *et al.*, 2013).

The FF constitutes the microenvironment in which oocytes develop and as a consequence, the protein composition of FF has attracted considerable interest and many proteomics studies have been conducted (Spitzer *et al.*, 1996; Anahory *et al.*, 2002; Lee *et al.*, 2005; Schweigert *et al.*, 2006; Angelucci *et al.*, 2006; Hanrieder *et al.*, 2008; Estes *et al.*, 2009; Gougeon, 2010; Jarkovska *et al.*, 2010; Twigt *et al.*, 2012, 2015; Ambekar *et al.*, 2013, 2015; Bianchi *et al.*, 2013; Jeppesen *et al.*, 2013; Zamah *et al.*, 2015; Oh *et al.*, 2017). All of these studies have focused on FF collected just prior to ovulation, in connection with assisted reproduction techniques. The fluid from pre-ovulatory follicles has a high dynamic range in terms of protein concentrations, which reduces the possibility of detecting low abundant proteins. At this follicular stage, the FF contains a high number of plasma constituents transferred through the follicular basal membrane during follicular expansion. For this reason, the number of proteins identified in previous studies has been in the range of several hundred. By using FF from human small antral follicles (hSAF), this limitation would be avoided. Furthermore, the study of the FF proteome by mass spectrometry (MS) allows the simultaneous analysis of hundreds of proteins and the current bioinformatics advances can provide, even from early follicular stages, the functional network of the proteins and their role in the follicular development. Hence, the larger the number of identified proteins, the greater the knowledge acquired. In addition, an increased focus on the earlier stages of follicles (i.e. small antral follicles), has shown that these could be a potential source of immature oocytes for women with a low ovarian reserve (Kristensen *et al.*, 2017). Details on the composition of FF from small antral follicles will provide essential knowledge about factors impacting the developmental capacity of the immature oocyte. By contributing to a better understanding of basic follicular processes, central proteins present in small antral follicles could be valuable in creating new physiological and effective methods for maturing oocytes *in vitro* and advancing culture conditions for human follicles, ultimately advancing fertility treatment by augmenting the number and quality of oocytes available for treatment.

The present study aimed to create a detailed fingerprint of proteins present in FF from hSAF to identify from the early follicular stage, candidate proteins that support follicular growth and development.

Materials and methods

An overview of the study workflow is presented in Supplementary Fig. S1.

Reagents and solutions

Unless otherwise specified, all chemical reagents were purchased from Sigma Aldrich (St. Louis, MO, USA). Modified porcine trypsin was

obtained from Promega (Madison, WI, USA) and the water was from a Milli-Q Ultrapure Water System (Millipore, Billerica, MA, USA). Liquid chromatography-mass spectrometry (LC-MS) grade water and organic solvents were supplied by Merck (Darmstadt, Germany). Prior to use, all LC-MS solutions were degassed by sonication. Synthetic peptides were provided by the University of Victoria-Genome BC Proteomics Centre, British Columbia, Canada.

Strong cation exchange (SCX) chromatography microspin columns (MA SEM HIL-SCX, 10–100 mg capacity) and silica C18 ultra-microspin columns (SUM SS18V, 3–30 mg capacity) were purchased from the Nest Group Inc. (South Borough, MA, USA). MARS 7 immunoaffinity spin columns were purchased from Agilent (Agilent Technologies, Inc., CA, USA) and were used to deplete the top seven most abundant proteins, i.e. albumin, IgG, antitrypsin, IgA, transferrin, haptoglobin and fibrinogen.

FF samples acquisition from small antral follicles

Samples of FF from hSAF were collected from ovaries that had been surgically removed from 25 women (25.6 ± 6.1 years old (mean ± SD)) undergoing unilateral ovariectomy for fertility preservation at the Laboratory of Reproductive Biology, Rigshospitalet, Denmark. Cryopreservation of the ovarian cortex was offered to women facing a potentially gonadotoxic treatment and thereby at risk of becoming sterile (Andersen *et al.*, 2008; Schmidt *et al.*, 2011). Cryopreservation of the ovarian cortical tissue provides these women with the option of later transplantation of the frozen-thawed tissue if they become menopausal due to the gonadotoxic treatment. Only patients with diseases unrelated to the ovary were included, and in all cases, the ovary had a macroscopically normal appearance. The diagnosis of the women included breast cancer (n = 11), Hodgkin's lymphoma (n = 3), sarcoma (n = 4), lymphoma (n = 4), brain cancer (n = 1), cervical cancer (n = 1) and systemic lupus erythematous (n = 1). Even though the women had a concurrent cancer diagnosis, they were eligible for fertility cryopreservation after an evaluation of reproductive parameters and ovarian reserve (AMH levels and antral follicle count); therefore, we assumed that the women and their ovaries were reproductively normal. Small antral follicles exposed on the surface of the ovary or visible during the isolation of ovarian cortex were aspirated with a 1-ml syringe fitted with a 26-gauge needle (Becton Dickinson, Brøndby, Denmark). From each ovary, FF from one small antral follicle was collected and the mean size of the follicles was 6.1 ± 0.4 mm (mean ± SD). The diameter was calculated based on the total volume of fluid drawn from the follicle using the calculation of spherical shape ($V = 4/3 * \pi * r^3$). Aspiration of FF had no effect on the fertility preservation procedure. The FF samples had no visible blood contamination but were immediately centrifuged to remove debris and cells. The FF were collected at random times during the menstrual cycle as the dynamics of the hSAF appears to be similar throughout the menstrual cycle (Mcnatty *et al.*, 1983; Westergaard, 1985; Kuang *et al.*, 2014). From 15 of the follicles, 50-µl FF was pooled whilst the remaining 10 follicles were analysed individually.

Furthermore, six women were selected (women with breast cancer (n = 4, sarcoma (n = 1) and multiple sclerosis (n = 1) and an average age of 31.4 ± 1.8 (mean ± SD)) and from each one, fluid from two or three hSAF (size 6 ± 1.5 mm (mean ± SD)) was obtained. Each follicle contained an oocyte that underwent IVM. From each woman, at least one follicle contained an oocyte that matured to metaphase II of the second meiotic division after IVM and another that contained an oocyte unable to mature. Oocytes were matured during a 48 h culture period using the MediCult IVM medium (Origio A/S, Denmark) supplemented with 75 mIU/ml rFSH (Puregon, MSD, the Netherlands), 100 mIU/ml rLH (Luveris, Serono, Germany) and 10 mg/ml human serum albumin. Oocytes were cultured individually in 25-µl drops and separate data for each oocyte was obtained. At the end of the culture period, all oocytes were denudated, and the developmental stage was classified as either germinal vesicle, metaphase I (MI) or metaphase II (MII).

The study was approved by the ethics committee of the municipalities of Copenhagen and Frederiksberg (H-2-2011-044). Informed consent was obtained from all participants.

Sample preparation to MS

A schematic representation of the general strategy followed for the analysis of the samples is shown in Supplementary Fig. S2. For all experiments, the quantitation of total proteins was performed using the bicinchoninic acid (BCA) assay.

Preparation of samples without depletion

Samples were dissolved in 1.6% sodium deoxycholate (SDC) in 50 mM NH_4HCO_3. The disulphide bonds were reduced by adding dithiothreitol (DTT) to a final concentration of 10 mM and incubated at 37°C for 1 h. The free thiol groups were alkylated by adding iodoacetamide (IAA) to a final concentration of 25 mM and incubated for 30 min at room temperature in the dark. The SDC was diluted to 0.5% before digestion with trypsin at an enzyme-to-substrate ratio of 1:100 (w/w) for 16 h at 37°C. The SDC was precipitated by adding 20% formic acid to a polypropylene filter plate with a hydrophilic polyvinylidene difluoride (PVDF) membrane (mean pore size 0.45 µm, Porvair Filtration Group, Fareham, UK).

Sample depletion

To deplete the top 7 or 14 (Supplementary Fig. S2) most abundant proteins, 10 or 25 µl of FF were used per sample, respectively. The depletions, with MARS 7 (spin columns) or MARS 14 column (Human-14 (4.6 × 100 mm) coupled to an 1260 Infinity LC System) both from Agilent technologies, were carried out according to instructions supplied by the manufacturer. After depletion, the buffer was exchanged to SDC 1.6%, 50 mM of Ambic using Amicon Ultra Centrifugal filter (0.5 ml – 10 kDa, Millipore, Tullagreen, Ireland).

Samples were reduced and alkylated as done with non-depleted samples and it took place in the Amicon filter. The buffer was exchanged to Ambic 50 mM and the samples were re-suspended in 100 µl of 50 mM AmBic (30 µg of proteins after BCA quantification). Then samples were digested with trypsin at an enzyme-to-substrate mass ratio of 1/30 for 16 h at 37 °C. The remaining SDC precipitated by adding 20% of formic acid prior to filtering the samples through a polypropylene filter plate with hydrophilic PVDF membrane (mean pore size 0.45 µm, Porvair Filtration Group, Fareham, United Kingdom).

Fractionation of peptides by strong cation exchange chromatography

For SCX chromatography, three aliquots (10 µl each) from the pool were depleted and separately digested. The digests were combined before fractionation. Mainly, a step-wise gradient of potassium chloride: 20, 40, 60, 100, 500 mM and 1 M KCl in 10 mM potassium phosphate, pH 2.8 (or 0.01% phosphoric acid) containing 20% acetonitrile was used to fractionate the peptides.

According to the instructions from the manufacturer, salt was removed from the samples by silica C18 ultra-microspin columns (SUM SS18V, 3–30 mg capacity, The Nest Group Inc., South Borough, MA, USA). After elution with 50% acetonitrile/0.1% TFA, the fractions were dried in a centrifugal evaporator and re-suspended in 0.1% formic acid prior to analysis by LC-MS/MS.

Fractionation of peptides by basic reversed-phase chromatography

Three aliquots (10 µl each) from the pooled FF were depleted and digested separately. We performed a Filter Aided Sample Preparation for protein digestion (Wiśniewski *et al.*, 2009). The digestions were combined prior to drying the sample. The sample was reconstituted in 100 µl ammonium formate/20% acetonitrile, pH 10, loaded on the column and fractionated using an Agilent 1100 series HPLC instrument by reversed-phase chromatography (RP-HPLC) at a flow rate of 50 µl/min. The mobile phase consisted of 20 mM ammonium formate, pH 10 (buffer A) and 20 mM ammonium formate, 80% acetonitrile, pH 10 (buffer B). The peptides were separated using the following gradient: 15 min isocratic hold at 1% solvent B, 1–20% solvent B in 1 min; 20–60% solvent B in 44 min; 60–80% solvent B in 5 min; finally, 15 min isocratic flow at 80% solvent B. Using 96 × 2 ml well plates, fractions were collected every 2 min for a total of 20 fractions. Finally, the fractions were recombined by pooling, and 12 were analysed by LC-MS.

MS acquisition methods

The samples were analysed by data-dependent acquisition (DDA), parallel reaction monitoring (PRM) and multiple reaction monitoring (MRM). Three different mass spectrometers from Thermo Fisher Scientific (San José, CA, USA) were used for the analyses: The Q-Exactive Plus (DDA and PRM, see Supplementary Fig. S2a and b), the Q-Exactive HF-X (DDA) and TSQ Quantiva (MRM) equipped with an easy-spray NG ion source. The spectrometers were connected to an easy-nLC 1000 pump (Thermo Scientific, San José, CA, USA).

In the Q-Exactive Plus mass spectrometer, a top 10 method was used for DDA. Full MS1 spectra were acquired in the Orbitrap mass analyser from m/z 400–1600 at a resolution of 70 000 (at m/z 200), a target automatic gain control (AGC) value of 1e6 and a maximum injection time (IT) of 100 ms over a 60-min HPLC gradient. The 10 most intense peaks with charge state ≥2 were fragmented in the HCD collision cell with a normalised collision energy of 26%. Tandem MS2 spectra were acquired in the Orbitrap mass analyser at a resolution of 35 000 (at m/z 200), a target AGC value of 5e4 and a maximum IT of 100 ms. The underfill ratio was set to 10% and dynamic exclusion was 45 s.

For the Q-Exactive HF-X mass spectrometer, the full MS scans were set with an acquisition range of m/z 375-1500, resolution of 120 000 (at m/z 200), target AGC value as 3 × 106, maximum injection time of 100 ms and normalised collision energy of 28. The top 20 precursors were selected for fragmentation. For the MS2 acquisition, we used a resolution of 15 000 (at m/z 200), target AGC value of 1 × 106, maximum injection time of 50 ms, isolation window of 1.2 m/z and fixed first mass at 110 m/z. Peptide elution was performed with a gradient of 2% of 80% acetonitrile/0.1% formic acid (Solvent B) and 98% of 0.1% formic acid (Solvent A) with a flow of 0.300 µl/min during the first 3 min. This step was followed by an increase in the percentage of solvent B to 25% in 112 min, to 32% in 10 min, and to 45% in 7 min. After 132 min of the gradient, the solvent B percentage was increased to 90% in 8 min and kept constant for 5 min. Finally, the solvent B content was reduced to 2% in 1 min and kept constant for 14 min.

For targeted proteomics (PRM and MRM), we built a spectral library using the MS/MS spectra from synthetic peptides. In addition, we measured a 'reference sample' by adding the synthetic peptides to a pool of all samples.

For PRM, 1 µg in 2 µl was loaded per sample onto the column. The MS2 resolution was 70 000 with an AGC value of 5e5 and a maximum IT of 200 ms. The normalised collision energy was 26%. Peptides were separated on an easy-spray column (25 cm × 75 µm ID, PepMap C18 2 µm, 100 Å) with the flow rate of 300 nl/min and the column temperature at 35°C. Solvent A (0.1% formic acid) and solvent B (0.1% formic acid in acetonitrile) were used to create a non-linear gradient to elute the peptides (60 min).

In the MRM method, a scheduled mode was used with 5-min detection windows. Peptides (1 µg) were loaded onto an Acclaim PepMap 100 pre-column (100 µm × 2 cm, Thermo Scientific, San José, CA, USA). Peptides were separated on an easy-spray column (15 cm × 75 µm ID, PepMap C18 3 µm, 100 Å) with the flow rate set to 300 nl/min and the column temperature at 35°C. Solvent A (0.1% formic acid) and solvent B (0.1% formic acid in acetonitrile) were used to create a non-linear gradient to elute the peptides (60 min). Selected reaction monitoring (SRM) transitions were acquired in Q1 and Q3 operated at unit resolution (0.7 fwhm); the collision gas pressure in Q2 was set to 1.5 mTorr. The cycle time was 2 s, and calibrated RF and S-lens values were used. At least three transitions per precursor were monitored.

MS data analysis

Raw files were analysed with Proteome Discoverer v2.2 or 2.4 (Thermo Scientific, San José, CA, USA). To identify the peptides, the MS data were searched against the UniProtKB human database (Released 20180207, 42213 sequences including isoforms). For the analysis of the data generated from FF that contained oocytes capable of maturing or not, we combined an FF spectral database built from the LC-MS/MS analysis of a top 14 depleted pool (hSAF MS1 spectral library) and the MSPepSearch node plus SEQUEST HT. We used the human spectral library 'ProteomeTools_HCD28_PD' and UniprotKB human database (Date: 28 January 2020), respectively. The search was performed with the following parameters: carbamidomethylation of cysteine residues and oxidation of methionine residues as static and dynamic modifications, respectively. Precursor and fragment ion tolerances were 10 ppm and 0.02 Da, respectively. Up to one missed cleavage site for tryptic peptides was allowed. The filters applied were

high (false discovery rate (FDR) < 1%) and medium (FDR < 5%) confidence at peptide and protein level, respectively. The peptide/protein quantification was based on the MS peptide signals (label-free quantification). For label-free quantification, the 'Minora Feature Detector' node was included in the processing workflow, and the nodes 'Precursor Ions Quantifier' and 'Feature Mapper' were included in the consensus workflow.

Comparison of the identified proteins with previous studies

The total number of proteins identified in the current study was compared to 30 previous proteomics studies performed in human FF. Out of 30, 20 studies were found after performing a PubMed and Embase searching using the term 'human follicular fluid proteome'. Considering that one of the previous papers (Bianchi *et al.*, 2016) summarised proteins identified up to 2014 (including 10 proteomics studies), we looked for proteomic studies published between 2013 and 2020. Only proteins reported as 'reviewed' in the UniProtKB/Swiss-Prot protein sequence database were selected. The studies were: Severino *et al.* (2013), Bayasula *et al.* (2013), Regiani *et al.* (2015), Twigt *et al.* (2015), Zamah *et al.* (2015), Ambekar *et al.* (2015), Chen *et al.* (2016), Bianchi *et al.* (2016), Shen *et al.* (2017), Oh *et al.* (2017), Lim *et al.* (2017), Lewandowska *et al.* (2017), Lewandowska *et al.* (2019), Poulsen *et al.* (2019), Li *et al.* (2019) , Zhang *et al.* (2019), Domingues *et al.* (2019), Liu *et al.* (2020a,b) and Zakerkish *et al.* (2020).

Furthermore, the list of identified proteins was compared with a list of proteins from oocytes reported by Virant-Klun *et al.* (2016) and with transcripts from GC published by Kõks *et al.* (2010).

Bioinformatics and statistical analyses

Gene Ontology (GO) and functional enrichment analyses were performed utilising the bioinformatics tool FunRich (Pathan *et al.*, 2015). The bioinformatics web tool DAVID (https://david.ncifcrf.gov/) (Huang *et al.*, 2009) was employed to perform a functional annotation clustering using as input 226 proteins from the high abundant proteome (163 only identified in hSAF, 28 only identified in pre-ovulatory follicles and 35 non-plasma proteins identified in both hSAF and large follicles). This tool was also employed to perform functional annotation clustering of the top 100 dysregulated proteins in FF that contained oocyte capable of reaching MII. Protein relationship networks were generated and analysed using the Ingenuity Pathway Analysis (IPA) software (QIAGEN, Germany).

To perform proteomics quantitative analyses, the intensities of the protein were normalised by $\log_2$ transformation and then standardised by subtracting the median of the sample. Statistical analyses were performed in RStudio software (R Core Team, 2016; RStudio Team, 2016). Coefficient of correlation between MDK, vimentin (VIM) and other proteins was determined by Pearson correlation test using the $\log_2$ intensities values of the protein quantified across 10 individual samples. The *P*-values were adjusted to control the FDR produced by multiple testing. Correlations with $-0.7 \geq r \geq 0.7$ and adjusted *P*-value <0.10 were considered significant. Proteins that significantly correlated with MDK/VIM were subjected to a biological pathway enrichment analysis in FunRich (background: FunRich database) and pathways with a significant enrichment score (BH method: adj. *P*-value <0.05) were selected. A comparative enrichment analysis based on 'cellular components' annotations was performed between positively and negatively correlated protein and a Q-value (Storey-Tibshirani method) <0.05 was considered significant.

To assess whether there were differences at protein level between FF that surrounded oocytes that matured or remained immature, a sparse partial least squares discriminant analysis (sPLS-DA) (Chung and Keles, 2010) was performed using 'mixOmics' R package. This is a multivariate analysis that classifies the samples by performing a multivariate regression using the protein expression matrix (749 proteins quantified in all samples) as predictors and the sampling origin (FF surrounding oocytes mature or immature) as the response. To select the top 100 most informative predictors (e.g. proteins) for discriminating samples, an LASSO penalisation was applied. With the top 100 proteins, a hierarchical clustering plus heatmap was performed using 'ComplexHeatmap' R library. In addition to the multivariable analysis, a Student *t*-test (two-tails) followed by FDR correction was performed to determine differentially expressed proteins. Proteins with an adjusted *P*-value <0.05 were considered significant. Since FF samples surrounding the immature oocyte and FF samples surrounding the MII oocyte originated in the same woman, the analyses (the multivariable and univariate) were performed considering the paired nature of the samples. This was achieved by subtracting from the protein intensity of a given sample the mean of the two or three samples belonging to the woman from which they were extracted. In this analysis, the expression values of VIM were determined by MRM. Differences between groups in term of follicular size were assessed by performing a paired Student *t*-test (two-tails). Secreted proteins dysregulated in FF that contained oocyte capable of reaching MII were correlated with the remaining proteins using a Pearson correlation test and adjusted *P*-values <0.05 were considered significant.

Selection of proteins potentially more concentrated and accessible by MS in FF from hSAF

The selection of proteins potentially more concentrated and accessible by MS in FF from hSAF is presented in the workflow in Supplementary Fig. S3. Proteins identified in the analysis of pool were filtered by proteins identified in the non-depleted samples and those also identified by another method, e.g. SCX or basic RP. Then 368 resulting proteins were compared against the proteins identified in the two largest previous data sets that used similar MS strategies (Zamah *et al.*, 2015; Oh *et al.*, 2017). From the resulting 35 proteins unique to our study, those known as 'classical plasma proteins' (Anderson and Anderson, 2002) (mainly immunoglobulins and complement components) were excluded. Then the resulting proteins were compared with the list of proteins identified in FF by MS up to 2020.

Results

General protein characterisation

In this study, a total of 2461 proteins were identified in FF from hSAF (Supplementary Table SI). A pool of 15 FF samples was first evaluated

to deepen the protein identification and subsequently, 23 separated samples (10 from different women and 13 from six other women) were analysed. Supplementary Tables SII, SIII and SIV contain the list of proteins identified in each case. Detailed information on the protein identification is described in Supplementary File S1 (Supplementary Figs S4, S5 and S6). The list of identified proteins was compared with previous human FF proteomics studies conducted up to 2020 (Supplementary Table SV). A total of 1108 proteins were detected for the first time in our study (Venn diagram: Fig. 1).

The entire set of proteins identified in this study was analysed to determine GO annotations. Figure 2 shows the protein distribution according to biological process, molecular function, cellular compartment and protein class of the FF proteome in hSAF. The generic terms: metabolic process and biological regulation were the most represented biological presses with 26% and 16%, respectively, whilst almost half of the identified proteins were related to catalytic (36%) and binding (43%) molecular functions. The classification based on cellular compartment indicated that 38% of the proteins were mainly nucleus and cytosolic (cell part), whilst 19% were extracellular.

To detect proteins possibly secreted from GC or oocyte, we compared our results with data obtained by transcriptomics in human GC and by proteomics in human oocytes (Kõks *et al.*, 2010; Virant-Klun *et al.*, 2016). In those studies, samples were obtained from pre-ovulatory follicles from women undergoing IVF. A total of 1940 proteins (94%) were found at the transcript level in GC whilst 793 proteins (39%) were identified in oocytes (Supplementary Table SI).

Comparing FF proteome from hSAF and large follicles: high abundance proteome

To our knowledge, this is the first proteomic study of FF from hSAF (<8 mm). To detect proteins possibly more accessible by MS in hSAF, a comparison in terms of the most abundant proteins identified in FF from hSAF and large follicles was carried out. The 'high abundance proteome' (non-depleted samples) identified in our study (413 proteins, Supplementary Table SVI) was compared to the 'high abundance proteome' (400 proteins) of a recent proteomics study performed by our group in large, pre-ovulatory follicles (>14 mm) (Poulsen *et al.*, 2019). Out of the 413 proteins from hSAF, 231 (56%) were also identified in FF from large pre-ovulatory follicles. From these, 196 are commonly detected in plasma (as compared with the human plasma proteome database) and may, therefore, represent plasma-filtrated proteins. According to the GO analysis, the remaining 35 proteins are

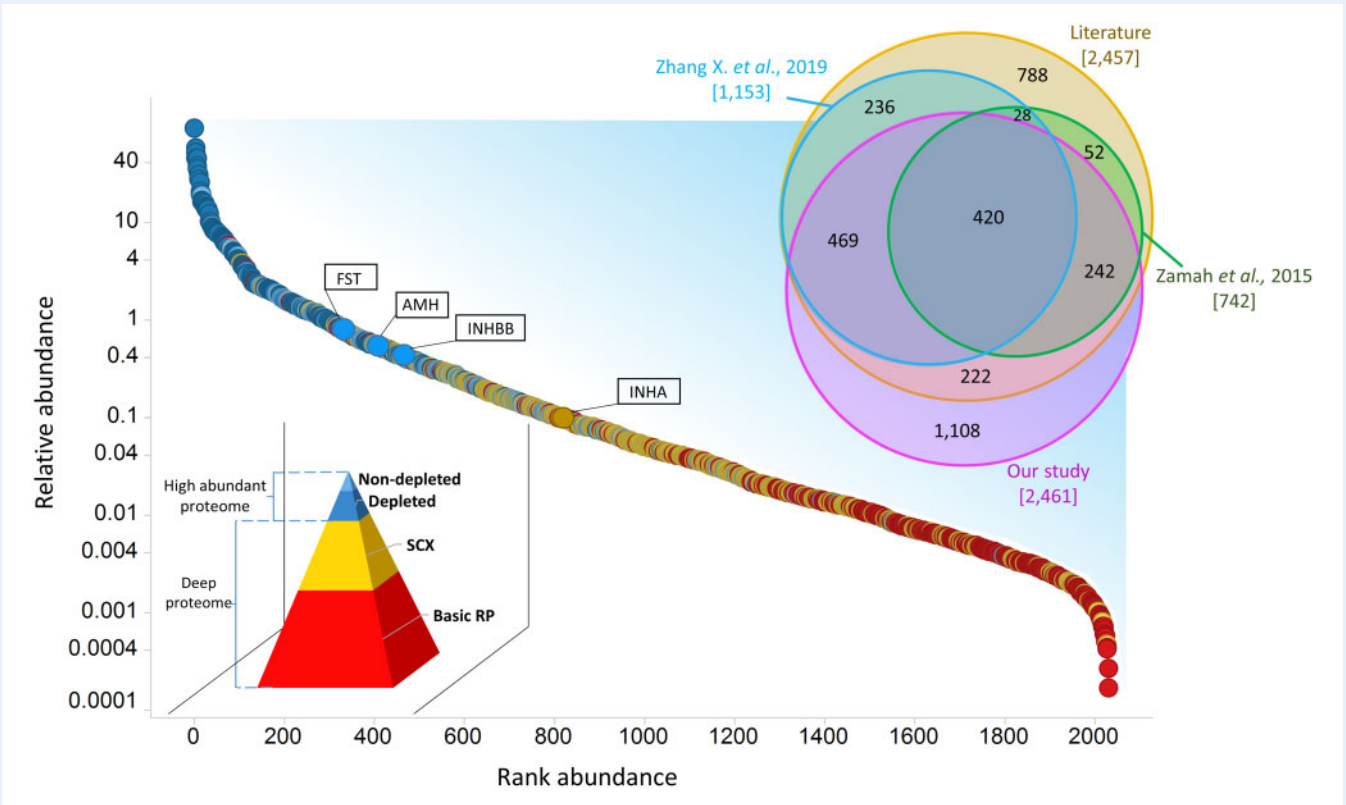

Figure 1. Dynamic rank of proteins identified in our study and Venn diagram of total proteins identified in human follicular fluid (FF) studies. In the Venn diagram, 'literature' (yellow) denotes proteins identified in previous FF studies (up to 2020). The circles in light blue and green colour, represent the two FF studies that currently have identified the highest number of proteins (Zamah *et al.*, 2015; Zhang et al., 2019). In total, 3565 proteins have been identified in FF samples by proteomics. The colours in the pyramid correspond to the colours in the dynamic rank plot. When a more complex method is applied (SCX and basic RP), the results yield a greater number of identifications in total and, in particular, increased identifications of low abundance proteins. Proteins highlighted are known to be relevant in the reproductive system.

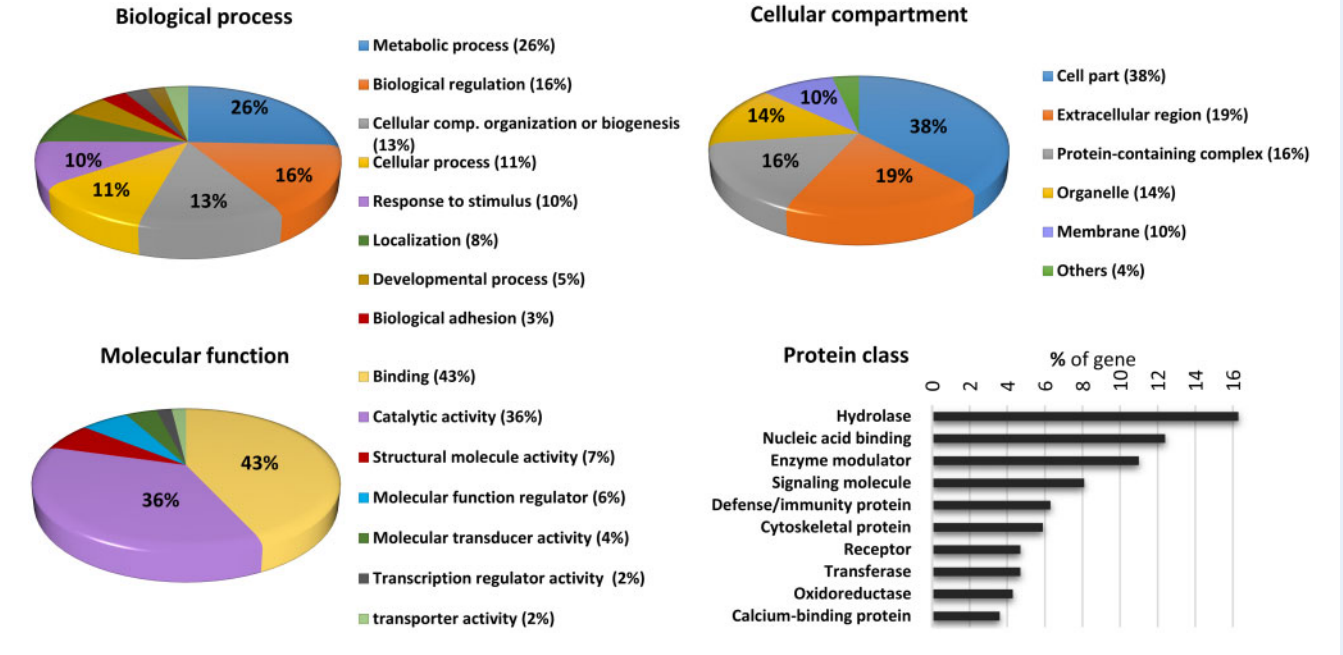

Figure 2. Gene ontology analysis for all follicular fluid (FF) proteins. Distribution of all proteins identified in this study according to three different GO categories and protein class.

distributed across the cell and almost half were secreted proteins (Fig. 3a).

Proteins only identified in hSAF and completely missing in FF from large follicles (163 proteins referred as 'on in hSAF' in Fig. 3a (red colour)) were mostly intracellular proteins (e.g. cytoplasmic and/or nuclear) whilst the extracellular proteins were distributed indistinctly in FF from small follicles and large follicles. However, proteins only identified in large follicles (28 proteins referred as 'off in hSAF' in Fig. 3b) are mostly allocated in the intracellular part. Interestingly, the proportion of secreted proteins was higher in proteins not identified in small follicles, i.e. only identified in large follicles (35/163 versus 13/28).

Furthermore, from the list of proteins identified in hSAF, a group of 24 proteins were highlighted (Supplementary Table SVII). This was carried out following a designed work-flow (see Materials and methods section, Supplementary Fig. S3), which allowed us to access proteins possibly more concentrated or accessible in FF from hSAF. Figure 3c shows the cellular compartments where these proteins are allocated according to the GO analysis. More than 90% of these proteins are intracellular (nucleus and cytosol).

The presence of the 24 proteins in FF from hSAF was verified in 23 samples analysed in this study (10 from different women and 13 from six other women). All of the 24 proteins were identified across the samples (Supplementary Fig. S7, Supplementary Table SVII), with the majority present in more than 70% of the samples.

Interestingly, four of the proteins (AMH, HTRA1, LOXL2 and MDK) are secreted by cells, which could indicate that they play an important role in hSAF, as it is well-known in the case of AMH (Jeppesen *et al.*, 2013). In order to inquire about the function of these proteins, a functional annotation clustering was performed on DAVID (Huang *et al.*, 2009) using an enriched list of proteins from the high abundant proteome that includes the 24 above selected proteins (see Materials and methods section). A cluster of 18 proteins involved in the ovarian follicle development came out (cluster 19 Supplementary Table SVIII, Fig. 3d). AMH and inhibin-B (associated to growth factor activity) are well-known to be present in small follicles at a high concentration (Andersen and Byskov, 2006), which was corroborated in this study by MS. Conversely, among others, the proteins inhibin-A and amphiregulin (AREG) were not identified in the 'high abundance proteome' of hSAF. Two of the four previously mentioned secreted proteins were involved in this cluster: the first one is the well-known AMH and the second one is midkine (MDK).

According to the bioinformatics analyses, specifically, MDK showed relevant evidence that indicated that this protein plays an important role in the follicle. MDK appears to be, out of 24 proteins, the only cell-secreted protein that functionally clusters together with well-known proteins involved in the ovarian follicle development (e.g. inhibin proteins (A and B), AREG, AMH) (Fig. 3d). On the other hand, MDK is involved in biological processes such as growth factor activity and cell differentiation. In a previous study, this protein was suggested to play a role in oocyte maturation in pre-ovulatory follicles (Poulsen *et al.*, 2019). Also, the addition of MDK in the culture medium during IVM seems to improve the maturation rate of oocytes (Nikiforov *et al.*, 2020). Apart from MDK, VIM was another of the 24 proteins more accessible in hSAF that came out in the cluster of ovarian follicle development. This protein is involved in the nuclear reprogramming (Kong *et al.*, 2014; Zhao *et al.*, 2015) and is well-known its function in the cytoskeletal organisation.

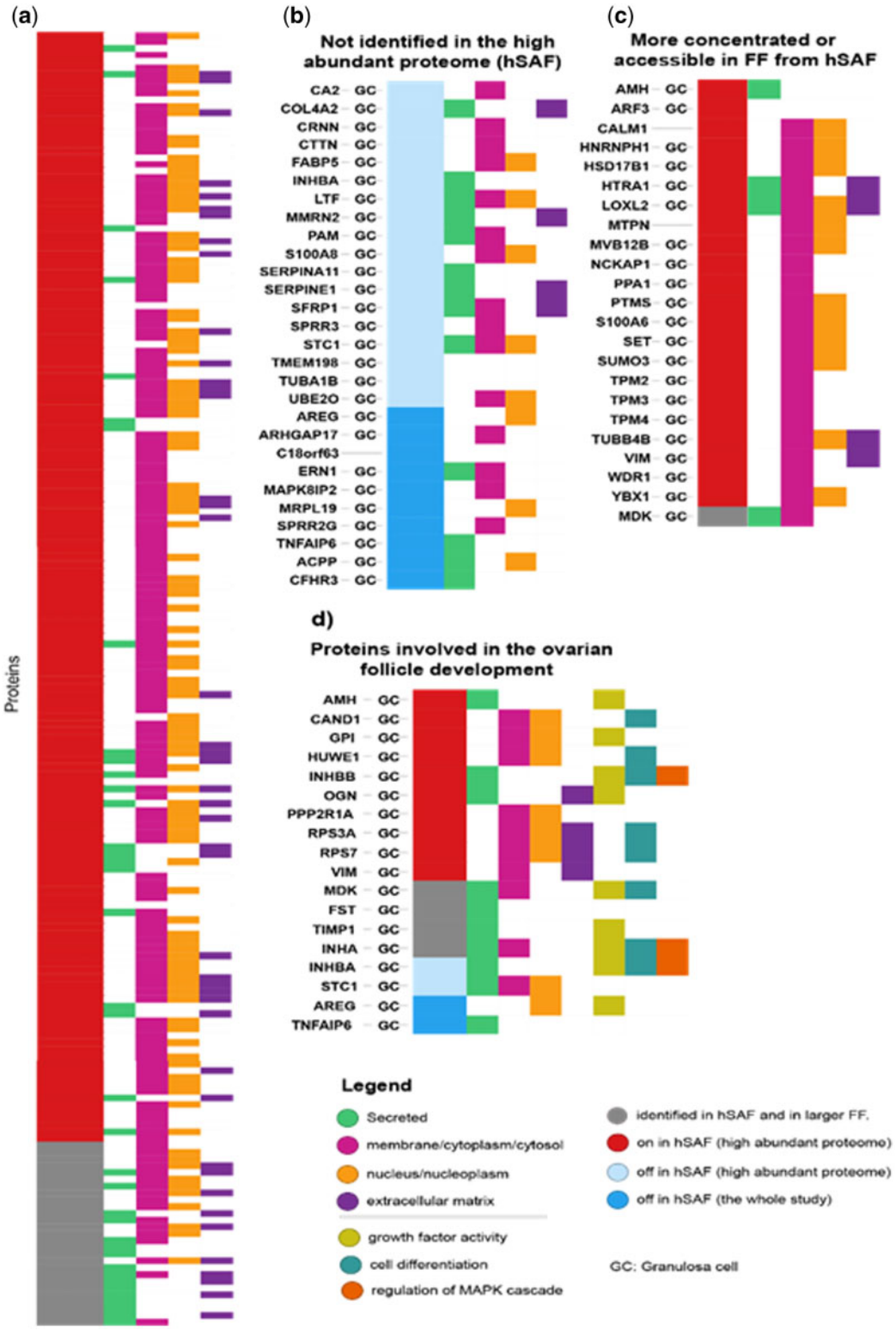

Figure 3. High abundance proteome identified in follicular fluid (FF) from human small antral follicles (hSAF) compared with proteins identified by Poulsen *et al.* (2019) **in large follicles**. In the two studies, protein identification was carried out following the same methodology. (**a**) Red: 163 proteins identified in FF from hSAF and not identified in large follicles. Grey: 35 proteins identified in both studies. (**b**) Proteins identified in FF from large follicles and not identified in hSAF. (**c**) Proteins more concentrated or accessible in hSAF. **(d)** Cluster of 18 proteins grouped by DAVID according to their functional role in the ovarian follicle development.

Functional relationship of MDK and selected proteins

To deepen our understanding of the role that MDK plays in hSAF for follicular development, we focus on proteins strongly correlated with this protein. Expression correlation is known as an indication of a functional association between genes or proteins (Pita-Juárez *et al.*, 2018). Accordingly, we set out to identify proteins and pathways that would likely act together with MDK during follicle development and oocyte maturation using the protein expression quantified in the individual samples. Interestingly, a significant positive strong correlation was found between MDK and VIM ($r = 0.727$, $P = 0.017$). These two proteins, apparently more accessible in hSAF, were also functionally clustered in the previous analysis. Considering the functionality that VIM has in the nuclear reprogramming and the cytoskeletal organisation, we also looked for proteins significantly correlated with VIM. A set of proteins significantly correlated (Pearson correlation: $-0.7 \geq r \geq 0.7$, $P < 0.02$) with MDK and VIM was generated to explore the functional network related to these proteins. The set included 72 and 25 positively and negatively correlated proteins, respectively (Supplementary Table SIX). The proteins that correlated negatively with MDK and VIM were primarily extracellular and cytosolic (Fig. 4b). The proteins that showed a positive correlation were predominantly nuclear proteins.

Proteins that significantly correlated with MDK/VIM (Supplementary Table SIX) were subjected to a biological pathway enrichment analysis. The analysis showed a significant enrichment of proteins positively correlated with MDK and VIM which are associated with gene regulation processes, such as transcription, chromosome maintenance and meiosis (Fig. 4a). These proteins are also associated with processes related to the development and normal function of female reproductive organs (ovaries and uterus), e.g. the epithelial-to-mesenchymal transition (EMT) process (Bilyk *et al.*, 2017). Taken together, these results suggest that MDK and the MDK/VIM protein pair could play a fundamental role in follicle development from early antral follicles and in oocyte maturation at later stages.

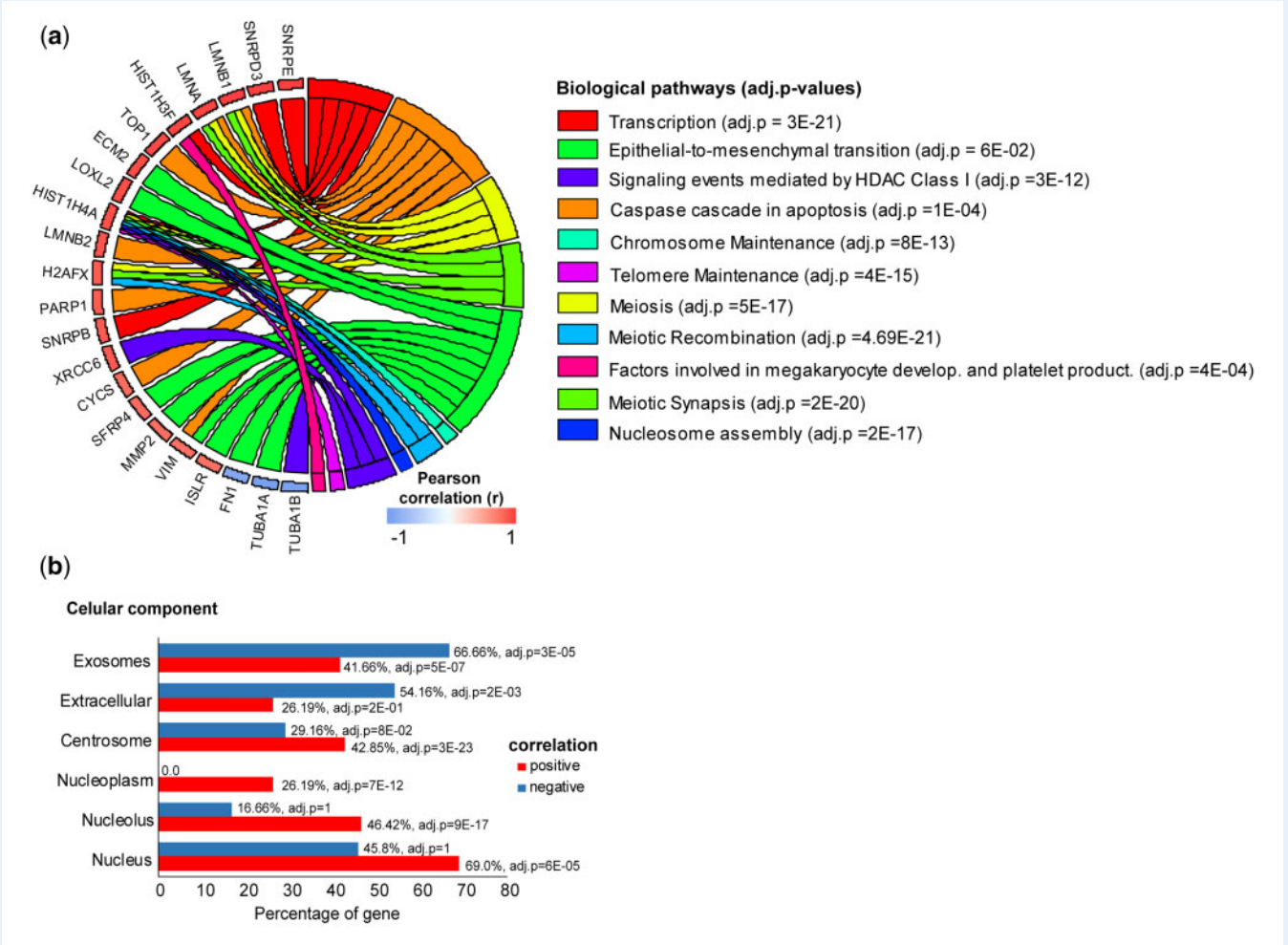

Figure 4. Functional enrichment analysis among proteins correlated positively or negatively with MDK and VIM. (**a**) Biological pathways significantly enriched (BH method: adjusted *P*-value <0.05). (**b**) Comparative enrichment analysis between positively and negatively correlated proteins based on 'cellular components' annotations. Bars represent the percentage of genes in each cellular component. VIM, vimentin.

Proteomic changes in FF from small antral follicles related to subsequent oocyte maturation

We set out to evaluate possible proteomics changes in FF from hSAF related to oocyte maturation. The protein profile of FF that surrounded oocytes capable of reaching MII (n = 7) with the protein profile of FF that surrounded oocytes that remained immature after IVM (n = 6) were compared. The samples were collected from hSAF extracted from six women, from which two or three samples were extracted (see Materials and methods section). The comparison was based on both, a multivariate (sPLS-DA) and univariate analyses (*t*-test). The sPLS-DA method perfectly classified the samples according to their

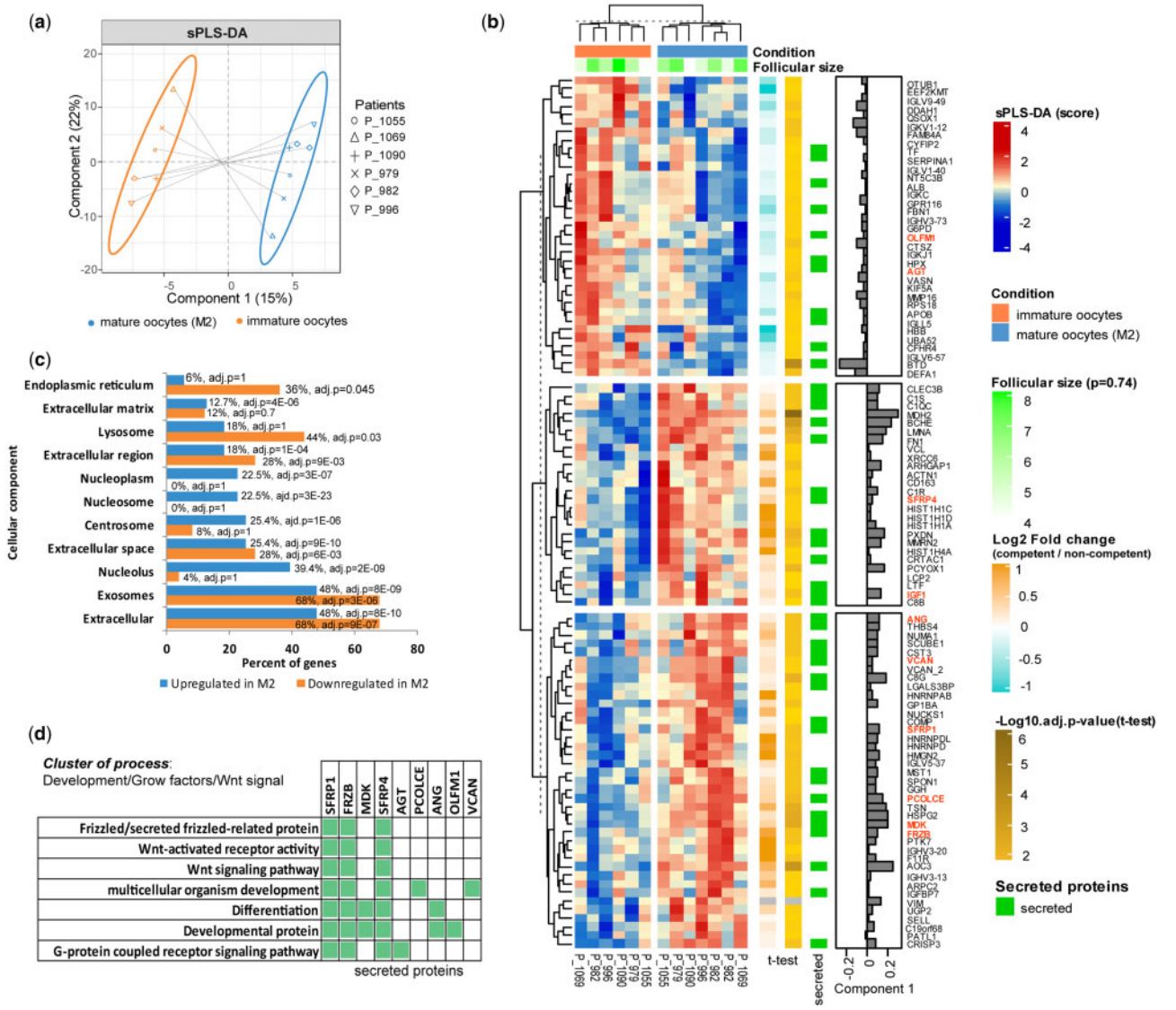

Figure 5. Proteins from human small antral follicles (hSAF) associated with upcoming oocyte maturation. (**a**) Sparse partial squares discriminant analysis performed with 750 proteins quantified in 13 paired FF samples extracted from small antral follicles coming from six women. The analysis discriminated between FF surrounding oocytes capable of achieving metaphase II (MII) after IVM (n = 7, blue) and FF surrounding oocytes unable to mature (n = 6, orange) after IVM. (**b**) Top 100 proteins that contributed to Component 1 of sPLS-DA to discriminate between FF samples. Positive and negative sPLS-DA scores values mean that the protein is up- and down-regulated in FF surrounding oocytes capable to reach M2, respectively. The bar chart indicates the contribution that each protein had in Component 1 (sPLS-DA) to discriminate between groups. (**c**) Comparative enrichment analysis between down- and up-regulated proteins in MII based on 'cellular components' annotations. Bars represent the percentage of genes in each cellular component. (**d**) Functional cluster of secreted proteins involved in development, growth factors and Wnt signal functions (clustered by DAVID software). These proteins were highlighted in red colour in the heat map shown in (b). sPLS-DA, sparse partial least squares discriminant analysis.

sample origin (i.e. FF surrounding oocytes mature or FF surrounding immature) (see Fig. 5a). The top 100 most informative proteins involved in the sample discrimination (shown by Component 1) are displayed in Fig. 5b. Since FF is composed of secretions from granulosa-theca cells and the oocyte, we looked for secreted proteins according to 'The human protein atlas' (https://www.proteinatlas.org/). Out of 100, 42 proteins were secreted (Fig. 5b). The *t*-test also revealed these proteins (Supplementary Table SIV) as significantly dysregulated (adjusted *P*-value <0.05) and the fold changes of the intensities matched with the sPLS-DA scores changes (heat map). Both, down- (35) and up-regulated (65) proteins in FF from small follicles containing oocytes capable of reaching MII were primarily exosome and extracellular proteins (Fig. 5c), however, the per cent of down-regulated proteins was higher. On the other hand, unlike down-regulated proteins, the up-regulated were significantly enriched in several nuclear compartments (Fig. 5c).

As we expected, MDK (adj. $P=0.003$) and VIM (adj. $P=0.016$) were included in the top 100 most dysregulated proteins. Within the groups of up-regulated proteins, there were 11 proteins significantly correlated to MDK and VIM as evaluated above (MDH2, LMNA, FN1, SFRP4, XRCC6, HIST1H4A, ANG, LGALS3BP, HNRNPD, SPON1 and PTK7).

To inquire into the functionality of the top 100 proteins, we performed a functional annotation clustering on the DAVID bioinformatics tool (https://david.ncifcrf.gov/) (Supplementary Table SX). Figure 5d shows one of the functional clusters integrated of secreted proteins (SFRP1, SFRP4, FRZB, MDK, AGT, PCOLCE, ANG, OLFM1, VCAN) involved in development, growth factors and Wnt signal processes. Two and seven of these secreted proteins were down- and up-regulated, respectively.

To further understand the role that these secreted proteins play in the follicle, we performed a biological pathways analysis not only with these proteins, but also with proteins that correlate to them (Supplementary Table SXI). In this case, we wanted to include IGF family proteins since they are stimulators of ovarian follicular development (Mazerbourg and Monget, 2018). IGF1 was included for the correlation since it was significantly up-regulated (adj. $P=0.018$, FC = 0.28) in FF containing oocytes capable of maturation. Interestingly, IGF2 (more commonly related to human folliculogenesis) did not change significantly. The pathways enrichment analysis revealed biological pathways such as transcription, signalling by NOTCH andepitelial-to-mesenchymal transition, among others (see Fig. 6). Proteins involved in the analysis were mostly enriched in the metabolism of protein pathway.

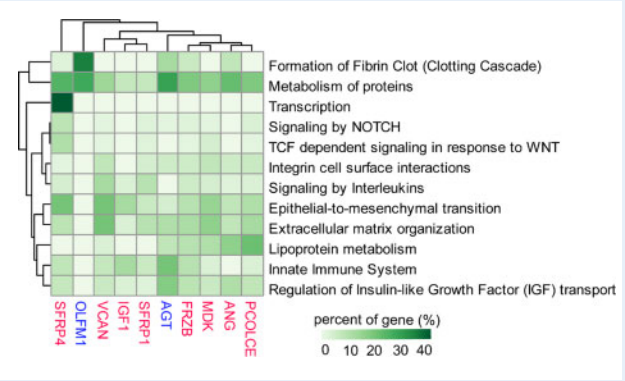

Figure 6. Pathways analysis of proteins that correlate (adjusted $P<0.05$) to secreted proteins involved in development, growth factors and Wnt signal processes. Rows represent the pathways in which the proteins were enriched. The heat map represents the per cent of enriched proteins that correlated to each secreted protein (columns). Proteins that significantly correlate with the secreted proteins are mostly enriched in the metabolism of proteins pathways. Most of the proteins correlated with SFRP4 were enriched in transcription pathways. To see if the correlation is positive or negative, see Supplementary Table SXI. Secreted proteins coloured in red and blue were up- and down-regulated in metaphase II, respectively.

Discussion

This study is to our knowledge, the first proteomics analysis of FF from hSAF obtained from women in their natural menstrual cycle. We were able to identify 2461 proteins of which more than 1108 were new to FF. The data generated constitute the largest number of proteins reported to date in human FF. We present an up-to-date compilation of proteins found in FF (Supplementary Table SV). Furthermore, we identified a signature of FF proteins significantly associated with the ability of the enclosed oocytes to sustain meiotic resumption. This suggests that oocyte viability is affected by the FF already at the early antral stage of follicular development. These results open up for a better understanding of the regulation of human folliculogenesis at earlier stages.

We demonstrated that the analysis by MS of FF from hSAF allows the identification of a greater number of proteins compared to the results obtained from previous analyses of larger follicles. It is well-known that in large follicles, there are more abundant proteins, due to the transfer of plasma constituents through the follicular basal membrane as the follicle expands during the late part of folliculogenesis (Anderson and Anderson, 2002; Zamah *et al.*, 2015). This fact makes the access to low abundance proteins more difficult in large follicles. Most of the new proteins identified in this study come from the cytosol and the nucleus of the cells that surround the antrum.

Follicles used in this study were obtained from ovaries removed surgically from women undergoing fertility preservation. These women did not have diseases related to the ovary and, overall, the ovaries appeared normal at surgery. Thus, we anticipate that the identified proteins could provide relevant physiological information related to folliculogenesis. A possible limitation of our study is the uncertainty of the proportion of the sampled follicles that are undergoing atresia.

On the other hand, we confirmed by MS that AMH and inhibin-B are more concentrated in FF from hSAF compared to large follicles as previously described in an ELISA experiment (Andersen and Byskov, 2006). Conversely, some proteins which were over-represented in the large pre-ovulatory follicles compared to hSAF, may play physiological roles for ovulation rather than for early follicular development. These proteins included inhibin-A and AREG, which are well-known for their determinant functional role in pre-ovulatory follicles (Zamah *et al.*, 2010; Poulsen *et al.*, 2019). Several of these proteins were even up-

regulated during ovulation, including TNFAIP6, AREG, SERPINE1 and ACPP, further highlighting their role in ovulation (Poulsen *et al.*, 2019).

Furthermore, it was possible to identify several signalling pathways and critical components affecting follicular growth and development. Similar biological pathways and processes were identified when we analysed proteins well-correlated to MDK/VIM and proteins well-correlated to those secreted and dysregulated in FF that contain oocytes capable of reaching MII. The biological pathways involved processes linked to follicular development. The chromosome maintenance and nucleosome assembly pathways appeared as part of the chromosome organisation of the GC that supports the progression of follicular growth and maturation. In addition, the EMT process is important for folliculogenesis (Kim *et al.*, 2014) and VIM is a well-known marker of the EMT process (Mirantes *et al.*, 2013). Numerous signalling pathways participate together in EMT up-regulation including PI3K-Akt, EGF, TGF-β, hepatocyte growth factor, MAPK-ERK, NF-kβ, Wnt, Notch, estrogen-receptor-α (ER-α) and HIF-1α. A similar analysis can be performed for the family of frizzled-related proteins (SFRP1, SFRP3, SFRP4) which in addition were up-regulated in FF surrounding oocytes capable of maturing to MII. SFRPs family function as modulators of Wnt signalling through direct interaction with Wnts. This is a targeted pathway that involves secreted glycoproteins that control development in organisms (Mikels and Nusse, 2006) and is also up-regulated in EMT. Thus, the observation of the extracellular matrix protein 2 (ECM2) in the enriched EMT pathway is not surprising. These ECM proteins regulate EMT via interactions with specific integrin receptors (Chen *et al.*, 2013).

Specifically, MDK has been reported to have a pro-survival effect on cumulus-GC (Ikeda and Yamada, 2014). This could explain why proteins significantly correlated (positively) with MDK were enriched in the caspase cascade in the apoptosis pathway. Generally, proteins involved in this pathway are commonly degraded and become markers for apoptosis (VIM, PARP1, LAMNB2 and TOP1). For example, the caspase proteolysis of VIM promotes apoptosis by dismantling intermediate filaments (Byun *et al.*, 2001).

For its part, VIM is a protein responsible for maintaining cell shape and cytoplasm integrity, as well as stabilising cytoskeletal interactions (Eriksson *et al.*, 2009). In addition, VIM was found to be expressed in mouse GC during folliculogenesis and the strongest expression was at earliest stages of follicle growth (Mora *et al.*, 2012). Other studies demonstrated that VIM is required for successful nuclear reprogramming in porcine cloned embryos (Kong *et al.*, 2014).

With the significant dysregulation of 100 proteins, it was demonstrated that changes occur at early follicular stages that could affect the subsequent oocyte maturation process. Interestingly, within the IGF family proteins, IGF1, well-known to be mainly active in non-human species, was up-regulated in this study in FF containing oocytes capable of reaching MII. This finding could be linked with the effect that IGF1 has at early follicular stages. Stubbs *et al.* (2013) demonstrated that IGF-1 stimulates the initiation of follicle growth in cultured ovarian tissue from the normal human ovary (Stubbs *et al.*, 2013). On the other hand, IGF1 has been described to play an important role in mouse oocyte competence and IVM (Toori *et al.*, 2014).

The dysregulated secreted proteins and their correlated network of proteins were involved in pathways such as EMT, extracellular matrix organisation, transcription and metabolism of proteins, among others. In this case, MDK appeared to be up-regulated in FF surrounding oocytes capable of maturing (to MII), which is a piece of valuable information to add to the wide range of studies conducted on this protein in the field of reproductive medicine.

MDK is a member of a family of neurotrophic factors that functions in the central nervous system, and which has been discovered to be expressed in ovarian follicles (Muramatsu *et al.*, 1993; Hirota *et al.*, 2005, 2007; Rauvala, 1989). It has been suggested that MDK is involved in cytoplasmatic maturation of bovine ovarian oocytes and it has been related to the promotion of oocyte developmental competence (Ikeda and Yamada, 2014). This protein acts indirectly via GCs, as no effect is observed in naked oocytes (Ikeda *et al.*, 2006). This is the first time that MDK has been identified in hSAF. Previously, it has been identified in FF from large follicles by western blotting (Hirota *et al.*, 2007) and recently it was identified in large follicles by MS (Poulsen *et al.*, 2019; Zhang *et al.*, 2019; Li *et al.*, 2020). In the study performed by Poulsen *et al.* (2019), MDK was suggested to play a role in oocyte maturation and, recently, the direct impact of MDK in oocyte maturation was assessed by Nikiforov *et al.* (2020). In their study, it was shown that the addition of MDK in the culture medium during IVM significantly improved the maturation rate of oocytes collected from surplus ovarian tissue after fertility preservation (Nikiforov *et al.*, 2020). On the other hand, Özdemir et al. found that the MDK levels in FF and serum may lead to an increase in blastocyst development. They also found that the level of MDK is higher in pregnant than in non-pregnant IVF-ICSI patients (Özdemir *et al.*, 2020).

Collectively, we present in this study, the first characterisation of proteins from hSAF in their natural state by MS. Several proteins observed in this study may have a strong relationship with the follicular developmental process. Furthermore, these data enforce that MDK and VIM are intimately involved in the regulation of oocyte performance in connection with meiotic resumption as observed close to ovulation. Most importantly, however, is that these results demonstrate that profound and significant differences exist in FF from follicles already at a non-selected stage with a diameter of below 9 mm, predicting the ability of the enclosed oocyte to sustain meiotic resumption. If this can be confirmed in further studies, it demonstrates that the viability of the oocyte is determined early on in follicular development and may open up new pathways for augmenting or attenuating subsequent oocyte viability in the pre-ovulatory follicle when ready to undergo ovulation.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Data availability

The mass spectrometry proteomics data have been deposited at the ProteomeXchange Consortium via the PRIDE(Perez-Riverol *et al.*, 2019) partner repository with the dataset identifiers PXD013344 and PXD021338.

Acknowledgements

The authors thank the patients who participated in this research project, without whom this research would not be possible. All

personnel, including the clinical activities in fertility preservation, are also thanked for their devoted work.

Authors' roles

C.Y.A. conceived the idea and designed the study together with J.M., I.P., A.S., S.E.P. and G.M.. S.E.P. collected FF and oocytes samples together with L.L.C.P. A.S. designed and executed the mass spectrometry (MS) experiments and acquired the data. I.P. performed bioinformatics and statistical analyses of the proteomics data together with K.P. who performed IPA functional analysis and supervised the bioinformatics analyses. I.P. also interpreted the proteomics results and drafted the manuscript together with A.S. and S.E.P. G.M. and J.M. supervised, contributed and approved the MS experimental design and statistical analyses of proteomics data. K.B.S. participated in the design of MS experiments and did samples management and preparation. C.Y.A. supervised the sample collection and interpreted the results in a biological context together with S.E.P. and L.L.C.P. R.A. managed the collaborative study together with J.M. All the authors revised the manuscript and approved the final version.

Funding

This study is part of the ReproUnion collaborative study, co-financed by Interreg V (European Union programme).

Conflict of interest

The authors have no conflicts of interest to disclose.

References

Ambekar AS, Kelkar DS, Pinto SM, Sharma R, Hinduja I, Zaveri K, Pandey A, Prasad TSK, Gowda H, Mukherjee S. Proteomics of follicular fluid from women with polycystic ovary syndrome suggests molecular defects in follicular development. *J Clin Endocrinol Metab* 2015;**100**:744–753.

Ambekar AS, Nirujogi RS, Srikanth SM, Chavan S, Kelkar DS, Hinduja I, Zaveri K, Prasad TSK, Harsha HC, Pandey A *et al.* Proteomic analysis of human follicular fluid: a new perspective towards understanding folliculogenesis. *J Proteomics* 2013;**87**:68–77.

Anahory T, Dechaud H, Bennes R, Marin P, Lamb NJ, Laoudj D, Pppelmann M, Becker WM, Petersen A. Identification of new proteins in follicular fluid of mature human follicles. *Electrophoresis* 2002;**23**:1197–1202.

Andersen CY, Byskov AG. Estradiol and regulation of Anti-Müllerian hormone, inhibin-A, and inhibin-B secretion: analysis of small antral and preovulatory human follicles' fluid. *J Clin Endocrinol Metab* 2006;**91**:4064–4069.

Andersen CY, Rosendahl M, Byskov AG, Loft A, Ottosen C, Dueholm M, Schmidt KLTT, Andersen AN, Ernst E. Two successful pregnancies following autotransplantation of frozen/thawed ovarian tissue. *Hum Reprod* 2008;**23**:2266–2272.

Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics* 2002;**1**: 845–867.

Angelucci S, Ciavardelli D, Di GF, Eleuterio E, Sulpizio M, Tiboni GM, Giampietro F, Palumbo P, Di IC. Proteome analysis of human follicular fluid. *Biochim Biophys Acta* 2006;**1764**:1775–1785.

Bayasula IA, Kobayashi H, Goto M, Nakahara T, Nakamura T, Kondo M, Nagatomo Y, Kotani T, Kikkawa F. A proteomic analysis of human follicular fluid: comparison between fertilized oocytes and non-fertilized oocytes in the same patient. *J Assist Reprod Genet* 2013;**30**:1231–1238.

Bianchi L, Gagliardi A, Campanella G, Landi C, Capaldo A, Carleo A, Armini A, De LV, Piomboni P, Focarelli R *et al.* A methodological and functional proteomic approach of human follicular fluid en route for oocyte quality evaluation. *J Proteomics* 2013;**90**:61–76.

Bianchi L, Gagliardi A, Landi C, Focarelli R, Leo V, De Luddi A, Bini L, Piomboni P. Protein pathways working in human follicular fluid: the future for tailored IVF? *Expert Rev Mol Med* 2016;**18**:e9.

Bilyk O, Coatham M, Jewer M, Postovit L-M. Epithelial-to-mesenchymal transition in the female reproductive tract: from normal functioning to disease pathology. *Front Oncol* 2017;**7**:145.

Byun Y, Chen F, Chang R, Trivedi M, Green KJ, Cryns VL. Caspase cleavage of vimentin disrupts intermediate filaments and promotes apoptosis. *Cell Death Differ* 2001;**8**:443–450.

Chen F, Spiessens C, D'Hooghe T, Peeraer K, Carpentier S. Follicular fluid biomarkers for human in vitro fertilization outcome: proof of principle. *Proteome Sci* 2016;**14**:17.

Chen QK, Lee K, Radisky DC, Nelson CM. Extracellular matrix proteins regulate epithelial-mesenchymal transition in mammary epithelial cells. *Differentiation* 2013;**86**:126–132.

Chung D, Keles S. Sparse partial least squares classification for high dimensional data. *Stat Appl Genet Mol Biol* 2010;**9**:17.

Domingues TS, Bonetti TCS, Pimenta DC, Mariano DOC, Barros B, Aquino AP, Motta ELA. Proteomic profile of follicular fluid from patients with polycystic ovary syndrome (PCOS) submitted to in vitro fertilization (IVF) compared to oocyte donors. *J Bras Reprod Assist* 2019;**23**:367–391.

Eriksson JE, Dechat T, Grin B, Helfand B, Mendez M, Pallari HM, Goldman RD. Introducing intermediate filaments: from discovery to disease. *J Clin Invest* 2009;**119**:1763–1771.

Estes SJ, Ye B, Qiu W, Cramer D, Hornstein MD, Missmer SA. A proteomic analysis of IVF follicular fluid in women ≤32 years old. *Fertil Steril* 2009;**92**:1569–1578.

Gougeon A. Human ovarian follicular development: from activation of resting follicles to preovulatory maturation. *Ann Endocrinol (Paris)* 2010;**71**:132–143.

Hanrieder J, Nyakas A, Naessén T, Bergquist J. Proteomic analysis of human follicular fluid using an alternative bottom-up approach. *J Proteome Res* 2008;**7**:443–449.

Hess KA, Chen L, Larsen WJ. The ovarian blood follicle barrier is both charge- and size-selective in mice1. *Biol Reprod* 1998;**58**: 705–711.

Hirota Y, Osuga Y, Koga K, Yoshino O, Hirata T, Harada M, Morimoto C, Yano T, Tsutsumi O, Sakuma S *et al.* Possible implication of midkine in the development of endometriosis. *Hum Reprod* 2005;**20**:1084–1089.

Hirota Y, Osuga Y, Nose E, Koga K, Yoshino O, Hirata T, Yano T, Tsutsumi O, Sakuma S, Muramatsu T *et al.* The presence of midkine and its possible implication in human ovarian follicles. *Am J Reprod Immunol* 2007;**58**:367–373.

Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;**4**:44–57.

Ikeda S, Saeki K, Imai H, Yamada M. Abilities of cumulus and granulosa cells to enhance the developmental competence of bovine oocytes during in vitro maturation period are promoted by midkine; a possible implication of its apoptosis suppressing effects. *Reproduction* 2006;**132**:549–557.

Ikeda S, Yamada M. Midkine and cytoplasmic maturation of mammalian oocytes in the context of ovarian follicle physiology. *Br J Pharmacol* 2014;**171**:827–836.

Jarkovska K, Martinkova J, Liskova L, Haiada P, Moos J, Rezabek K, Gadher SJ, Kovarova H, Halada P, Moos J *et al*. Proteome mining of human follicular fluid reveals a crucial role of complement cascade and key biological pathways in women undergoing in vitro fertilization. *J Proteome Res* 2010;**9**:1289–1301.

Jeppesen JV, Anderson RA, Kelsey TW, Christiansen SL, Kristensen SG, Jayaprakasan K, Raine-Fenning N, Campbell BK, Yding Andersen C. Which follicles make the most anti-Müllerian hormone in humans? Evidence for an abrupt decline in AMH production at the time of follicle selection. *Mol Hum Reprod* 2013;**19**:519–527.

Kim YS, Yi BR, Kim NH, Choi KC. Role of the epithelial-mesenchymal transition and its effects on embryonic stem cells. *Exp Mol Med* 2014;**46**:e108.

Kõks S, Velthut A, Sarapik A, Altmäe S, Reinmaa E, Schalkwyk LC, Fernandes C, Lad HV, Soomets U, Jaakma Ü *et al*. The differential transcriptome and ontology profiles of floating and cumulus granulosa cells in stimulated human antral follicles. *MHR Basic Sci Reprod Med* 2010;**16**:229–240.

Kong Q, Xie B, Li J, Huan Y, Huang T, Wei R, Lv J, Liu S, Liu Z. Identification and characterization of an oocyte factor required for porcine nuclear reprogramming. *J Biol Chem* 2014;**289**:6960–6968.

Kristensen SG, Pors SE, Andersen CY. Diving into the oocyte pool. *Curr Opin Obstet Gynecol* 2017;**29**:112–118.

Kuang T-M, Liu CJ-L, Ko Y-C, Lee S-M, Cheng C-Y, Chou P. Distribution and associated factors of optic disc diameter and cup-to-disc ratio in an elderly Chinese population. *J Chinese Med Assoc* 2014;**77**:203–208.

Lee HC, Lee SW, Kyo WL, Lee SW, Cha KY, Kye HK, Lee S. Identification of new proteins in follicular fluid from mature human follicles by direct sample rehydration method of two-dimensional polyacrylamide gel electrophoresis. *J Korean Med Sci* 2005;**20**:456–460.

Lewandowska AE, Macur K, Czaplewska P, Liss J, Łukaszuk K, Ołdziej S. Qualitative and quantitative analysis of proteome and peptidome of human follicular fluid using multiple samples from single donor with LC−MS and SWATH methodology. *J Proteome Res* 2017;**16**:3053–3067.

Lewandowska AE, Macur K, Czaplewska P, Liss J, Łukaszuk K, Ołdziej S. Human follicular fluid proteomic and peptidomic composition quantitative studies by SWATH-MS methodology. Applicability of high pH RP-HPLC fractionation. *J Proteomics* 2019;**191**:131–142.

Li H, Huang X, Chang X, Yao J, He Q, Shen Z, Ji Y, Wang K. S100-A9 protein in exosomes derived from follicular fluid promotes inflammation via activation of NF-κB pathway in polycystic ovary syndrome. *J Cell Mol Med* 2019;**24**:114–125.

Lim H-JJ, Seok AE, Han J, Lee J, Lee S, Kang H-GG, Cha BH, Yang Y. N-glycoproteomic analysis of human follicular fluid during natural and stimulated cycles in patients undergoing in vitro fertilization. *Clin Exp Reprod Med* 2017;**44**:63–72.

Liu X, Wang Y, Zhu P, Wang J, Liu J, Li N, Wang W, Zhang W, Zhang C, Wang Y *et al*. Human follicular fluid proteome reveals association between overweight status and oocyte maturation abnormality. *Clin Proteom* 2020a;**17**:22.

Liu Y, Wu Y, Tian M, Luo W, Zhang C, Liu Y, Li K, Cheng W, Liu D. Protein expression profile in IVF follicular fluid and pregnancy outcome analysis in euthyroid women with thyroid autoimmunity. *ACS Omega* 2020b;**5**:11439–11447.

Mazerbourg S, Monget P. Insulin-like growth factor binding proteins and IGFBP proteases: A dynamic system regulating the ovarian folliculogenesis. *Front Endocrinol (Lausanne)* 2018;**9**:134. doi: 10.3389/fendo.2018.0013

Mcnatty KP, Hillier SG, Boogaard AMJ, van den Trimbos-Kemper TCM, Reichert LE, Hall EV. van Follicular development during the luteal phase of the human menstrual cycle. *J Clin Endocrinol Metab* 1983;**56**:1022–1031.

Mikels AJ, Nusse R. Wnts as ligands: processing, secretion and reception. *Oncogene* 2006;**25**:7461–7468.

Mirantes C, Espinosa I, Ferrer I, Dolcet X, Prat J, Matias-Guiu X. Epithelial-to-mesenchymal transition and stem cells in endometrial cancer. *Hum Pathol* 2013;**44**:1973–1981.

Mora JM, Fenwick MA, Castle L, Baithun M, Ryder TA, Mobberley M, Carzaniga R, Franks S, Hardy K. Characterization and significance of adhesion and junction-related proteins in mouse ovarian follicles. *Biol Reprod* 2012;**86**:1–14.

Muramatsu H, Shirahama H, Yonezawa S, Maruta H, Muramatsu T, Midkine A. Retinoic acid-inducible growth/differentiation factor: immunochemical evidence for the function and distribution. *Dev Biol* 1993;**159**:392–402.

Nikiforov D, Junping C, Cadenas J, Shukla V, Blanshard R, Pors SE, Kristensen SG, Macklon KT, Colmorn L, Ernst E *et al*. Improving the maturation rate of human oocytes collected ex vivo during the cryopreservation of ovarian tissue. *J Assist Reprod Genet* 2020;**37**:891–904.

Oh JW, Kim SK, Cho K-C, Kim M-S, Suh CS, Lee JR, Kim KP. Proteomic analysis of human follicular fluid in poor ovarian responders during in vitro fertilization. *Proteomics* 2017;**17**:1600333.

Özdemir A, Karli P, Avci B. Do midkine levels in serum and follicular fluid affect IVF-ICSI outcome? *J Heal Sci Med* 2020;**3**:301–306.

Pathan M, Keerthikumar S, Ang C-S, Gangoda L, Quek CYJ, Williamson NA, Mouradov D, Sieber OM, Simpson RJ, Salim A *et al*. FunRich: an open access standalone functional enrichment and interaction network analysis tool. *Proteomics* 2015;**15**:2597–2601.

Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M *et al*. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* 2019;**47**:D442–D450.

Pita-Juárez Y, Altschuler G, Kariotis S, Wei W, Koler K, Green C, Tanzi RE, Hide W. The pathway coexpression network: revealing pathway relationships. *PLoS Comput Biol* 2018;**14**:e1006042.

Poulsen L, la C, Pla I, Sanchez A, Grøndahl ML, Marko-Varga G, Yding Andersen C, Englund ALM, Malm J. Progressive changes in human follicular fluid composition over the course of ovulation: quantitative proteomic analyses. *Mol Cell Endocrinol* 2019;**495**:110522.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna. https://www.r-project.org/. version 3.6.3 (22 February 2020, date accessed) 2016.

Rauvala H. An 18-kd heparin-binding protein of developing brain that is distinct from fibroblast growth factors. *EMBO J* 1989;**8**:2933–2941.

Regiani T, Cordeiro FB, Costa L, do V. D, Salgueiro J, Cardozo K, Carvalho VM, Perkel KJ, Zylbersztejn DS, Cedenho AP, Turco EG. Lo Follicular fluid alterations in endometriosis: label-free proteomics by MS ^{E} as a functional tool for endometriosis. *Syst Biol Reprod Med* 2015;**61**:263–276.

Rodgers RJ, Irving-Rodgers HF. Formation of the ovarian follicular antrum and follicular fluid. *Biol Reprod* 2010;**82**:1021–1029.

RStudio Team. *RStudio: Integrated Development for R*. RStudio.Boston. http://www.rstudio.com/. version 1.1.463. (13 September 2019, date accessed) 2016.

Schmidt KT, Rosendahl M, Ernst E, Loft A, Andersen AN, Dueholm M, Ottosen C, Andersen CY. Autotransplantation of cryopreserved ovarian tissue in 12 women with chemotherapy-induced premature ovarian failure: the Danish experience. *Fertil Steril* 2011; **95**:695–701.

Schweigert FJ, Gericke B, Wolfram W, Kaisers U, Dudenhausen JW. Peptide and protein profiles in serum and follicular fluid of women undergoing IVF. *Hum Reprod* 2006;**21**:2960–2968.

Severino V, Malorni L, Cicatiello AE, D'Esposito V, Longobardi S, Colacurci N, Miraglia N, Sannolo N, Farina A, Chambery A. Array and iTRAQ labeling for in-depth identification of pathways associated to IVF outcome. *PLoS One* 2013;**8**:e77303.

Shen X, Liu X, Zhu P, Zhang Y, Wang J, Wang Y, Wang W, Liu J, Li N, Liu F. Proteomic analysis of human follicular fluid associated with successful in vitro fertilization. *Reprod Biol Endocrinol* 2017;**15**:58.

Siu MKY, Cheng CY. The Blood-Follicle Barrier (BFB) in disease and in ovarian function. *Adv Exp Med Biol* 2012;**763**:186–192.

Spitzer D, Murach KF, Lottspeich F, Staudach A, Illmensee K. Different protein patterns derived from follicular fluid of mature and immature human follicles. *Hum Reprod* 1996;**11**: 798–807.

Stubbs SA, Webber LJ, Stark J, Rice S, Margara R, Lavery S, Trew GH, Hardy K, Franks S. Role of Insulin-like growth factors in initiation of follicle growth in normal and polycystic human ovaries. *J Clin Endocrinol Metab* 2013;**98**:3298–3305.

Toori MA, Mosavi E, Nikseresht M, Barmak MJ, Mahmoudi R. Influence of insulin-like growth factor-i on maturation and fertilization rate of immature oocyte and embryo development in NMRI mouse with TCM199 and α-MEM medium. *J Clin Diagnostic Res* 2014;**8**:AC05–AC08.

Twigt JM, Bezstarosti K, Demmers J, Lindemans J, Laven JSE, Steegers-Theunissen RP. Preconception folic acid use influences the follicle fluid proteome. *Eur J Clin Invest* 2015;**45**:833–841.

Twigt JM, Bolhuis MEC, Steegers EAP, Hammiche F, van Inzen WG, Laven JSE, Steegers-Theunissen RPM. The preconception diet is associated with the chance of ongoing pregnancy in women undergoing IVF/ICSI treatment. *Hum Reprod* 2012;**27**:2526–2531.

Virant-Klun I, Leicht S, Hughes C, Krijgsveld J. Identification of maturation-specific proteins by single-cell proteomics of human oocytes. *Mol Cell Proteomics* 2016;**15**:2616–2627.

Westergaard H. The sprue syndromes. *Am J Med Sci* 1985;**290**: 249–262.

Wiśniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. *Nat Methods* 2009;**6**:359–362.

Zakerkish F, Brännström M, Carlsohn E, Sihlbom C, Post S, Thoroddsen A. Proteomic analysis of follicular fluid during human ovulation. *Acta Obstet Gynecol Scand* 2020;**99**:917–924.

Zamah AM, Hassis ME, Albertolle ME, Williams KE. Proteomic analysis of human follicular fluid from fertile women. *Clin Proteom* 2015; **12**:1–12.

Zamah AM, Hsieh M, Chen J, Vigne JL, Rosen MP, Cedars MI, Conti M. Human oocyte maturation is dependent on LH-stimulated accumulation of the epidermal growth factor-like growth factor, amphiregulin. *Hum Reprod* 2010;**25**:2569–2578.

Zhang X, Xu X, Li P, Zhou F, Kong L, Qiu J, Yuan Z, Tan J. TMT based proteomic analysis of human follicular fluid from overweight/obese and normal-weight patients with polycystic ovary syndrome. *Front Endocrinol (Lausanne)* 2019;**10**:821.

Zhao Q, Guo Z, Piao S, Wang C, An T. Discovery of porcine maternal factors related to nuclear reprogramming and early embryo development by proteomic analysis. *Proteome Sci* 2015;**13**:18.