Reproductive Sciences (2023) 30:642–655
https://doi.org/10.1007/s43032-022-01046-2

EMBRYOLOGY: ORIGINAL ARTICLE

Bioinformatic Analysis of Human Cumulus Cells to Unravel Cellular's Processes that Could Be Used to Establish Oocyte Quality Biomarkers with Clinical Application

Lucia von Mengden[1,2,3] · Marco Antônio De Bastiani[1,2,3] · Lucas Kich Grun[2,3,4,5] · Florencia Barbé-Tuana[3,4,5] · Tom Adriaenssens[6] · Johan Smitz[6] · Leticia Schmidt Arruda[7] · Carlos Alberto Link[7] · Fábio Klamt[1,2,3]

Received: 3 February 2022 / Accepted: 18 July 2022 / Published online: 26 July 2022

Abstract

Metadata analysis of public microarray datasets using bioinformatics tools has been successfully used in several biomedical fields in the search for biomarkers. In reproductive science, there is an urgent need for the establishment of oocyte quality biomarkers that could be used in the clinical environment to increase the chances of successful outcomes in treatment cycles. Adaptive cellular processes observed in *cumulus oophorus* cells reflect the conditions of the follicular microenvironment and may thus bring relevant information of oocyte's conditions. Here we analyzed human *cumulus* cells gene expression datasets in search of predictors of oocyte quality, a strategy which uncovered several cellular processes positively and negatively associated with embryo development and pregnancy potential. Secondly, the expression levels of genes that were present in the majority of processes observed were validated *in house* with clinical samples. Our data confirmed the association of the selected biomarkers with blastocyst formation and pregnancy potential rates, independently of patients' clinical characteristics such as diagnosis, age, BMI, and stimulation protocol applied. This study shows that bioinformatic analysis of cellular processes can be successfully used to elucidate possible oocyte quality biomarkers. Our data reinforces the need to consider clinical characteristics of patients when selecting relevant biomarkers to be used in the clinical environment and suggests a combination of positive (*PTGS2*) and negative (*CYPB1*) quality biomarkers as a robust strategy for a complementary oocyte selection tool, potentially increasing assisted reproduction success rates. Also, *GPX4* expression as pregnancy potential biomarker is indicated here as a possibility for further investigations.

Keywords *Cumulus oophorus* cells · Bioinformatic · Oocyte quality · Blastocyst formation · IVF · Functional enrichment analysis

Introduction

Despite its exponentially increasing popularity, the rate of assisted reproduction technique (ART) success is still relatively low (around 23%) [43] and the outcome is hardly predictable. In this scenario, appropriate oocyte selection would improve in vitro fertilization outcomes, limit embryo overproduction, and abbreviate time for *take-home* babies. Nowadays, oocyte selection relies mainly in morphological analyses [60], which is not an unbiased method and may fail to reveal the gamete real competence status [10, 11].

Cumulus oophorus cells (CCs) are somatic cells that surround the oocyte in the antral follicle. These cells represent the interface of the gamete with the ovarian environment [56]. Connected to the oocyte through several specialized junctions, the CCs deliver essential compounds to the oocyte, receive metabolic products from the gamete, and protect [4, 67] and participate in the oocyte maturation process [5, 44]. Therefore, analysis of CCs may provide valuable information on the quality and genetics [32, 33] of the oocyte and its environment. Since CCs are discarded before fertilization by intracytoplasmic sperm injection (ICSI) procedure, the use of CC offers no ethical barriers, being a non-invasive, easy-to-access surrogate tissue for oocyte evaluation.

Lucia von Mengden and Marco Antônio De Bastiani should be regarded as joint first authors.

✉ Lucia von Mengden
luciavonmengden@ufrgs.br

Extended author information available on the last page of the article

Gene expression patterns in CCs have previously been considered a predictive tool for oocyte quality in several studies [1, 9, 23, 27, 28, 31, 38, 45, 51, 62, 66, 70, 73]. However, there is little consensus about which biomarkers would actually be clinically relevant [16, 37, 48, 50, 57, 65]. This might be a consequence of the strong influence that the patient characteristics (such as diagnosis, age, body mass index (BMI), and stimulation protocol applied) have over oocyte's and CC's biology, as observed when the same biomarkers were analyzed in different clinical profiles [7, 8, 15, 17, 19, 22, 30, 39, 49, 66, 76]. Thus, biomarkers for oocyte quality that considers individual patient's characteristics are highly needed and still not available.

In this context, high-throughput genomic scanning technologies, such as microarray gene expression analysis, allow the study of a large variety of gene expression patterns, obtaining a systemic understanding of several biological phenomena and conditions [14, 20, 71]. Besides facilitating the identification of differentially expressed genes (DEGs) in various conditions, it is possible to extract biological features/meaning associated with groups of DEG (such as molecular and cellular processes that analyzes the differences through functional enrichment analyses). These approaches increase the likelihood of identifying biological processes and markers most relevant to the event of interest, saving time and resources by enabling data-driven hypothesis generation and research [40], and have been applied successfully in distinct fields, such as neuroscience [21, 58, 71] and oncology [13]. Furthermore, this kind of analysis can drastically abbreviate research costs and time, since it can be done with public available datasets, and provide a very robust approach to search for possible targets to be validated in the laboratory, identifying not only the most expressed genes but also the genes involved in the majority of biological processes responsible for the success of the outcome of interest [12, 24, 47, 75].

Here, we performed functional enrichment analyses in human CC datasets in order to identify a novel and robust technique for enlightment of *cumulus*-oocyte-embryo dinamic and to determine if this approach could be applied for biomarker selection, highlighting biologically relevant genes, representatives of complex processes, that could further be easily targeted in the clinical environment for several patient profiles. Afterwards, selected genes were experimentally validated *in house* with patient samples as possible oocyte quality biomarkers. In this sense, we also took into consideration patients'clinical caracteristics. This approach uncovered relevant biomarkers that can be used as a support tool for oocyte selection in clinical scenario.

Materials and Methods

Microarray Data Acquisition

The entire bioinformatics pipeline was conducted in *R* statistical environment [59] with Bioconductor packages [42], and is summarized in Figure S1.

Microarray datasets were obtained from Gene Expression Omnibus (GEO) public platform [12], under the accession numbers GSE37277 [30] and GSE55654 [16]. Data acquisition was done using the GEOquery package [63]. Inclusion criterion was human CC microarray data related to oocyte quality. Sample classification into good quality group (GQ) or in poor quality group (PQ) in each dataset is summarized in Fig. S1 and Fig. 1A.

Differentially Expressed Genes and Functional Enrichment Analysis

For each dataset, uninformative probes were removed and duplicated probes were filtered according to their variance using genefilter package [36]. Next, LIMMA package [61] was used to assess the DEGs for good *versus* poor oocyte quality comparisons, according to the group definitions in each study. Thus, positive log fold change (logFC) values represent genes upregulated in good-quality oocytes and negative logFC values genes upregulated in poor-quality oocytes. For subsequent analyses, we limited our study to DEG showing an absolute expression fold change of 1.5 and above (Supplementary Tables S1, S2, S3, and S4).

Finally, for the functional enrichment analysis (FEA), we divided the DEG list into positive and negative and employed the FGNet package [3, 6] and topGO feature for biological processes, with a significance cutoff threshold of 0.001 and node size 50. The network graphical representation of FEA results were built using RedeR package [18].

Ethical Considerations

This study was approved by the Research Ethics Committee (#68081017.2.0000.5347). CCs were obtained as waste products of a local fertility clinic's ICSI procedures and had no other destination beyond the experiments described here. They were supplied anonymously to the laboratory after the patient signed an informed consent.

Patients and Samples

All patient samples were retrieved from the local fertility clinic. Sample collection is summarized in Fig. 1.

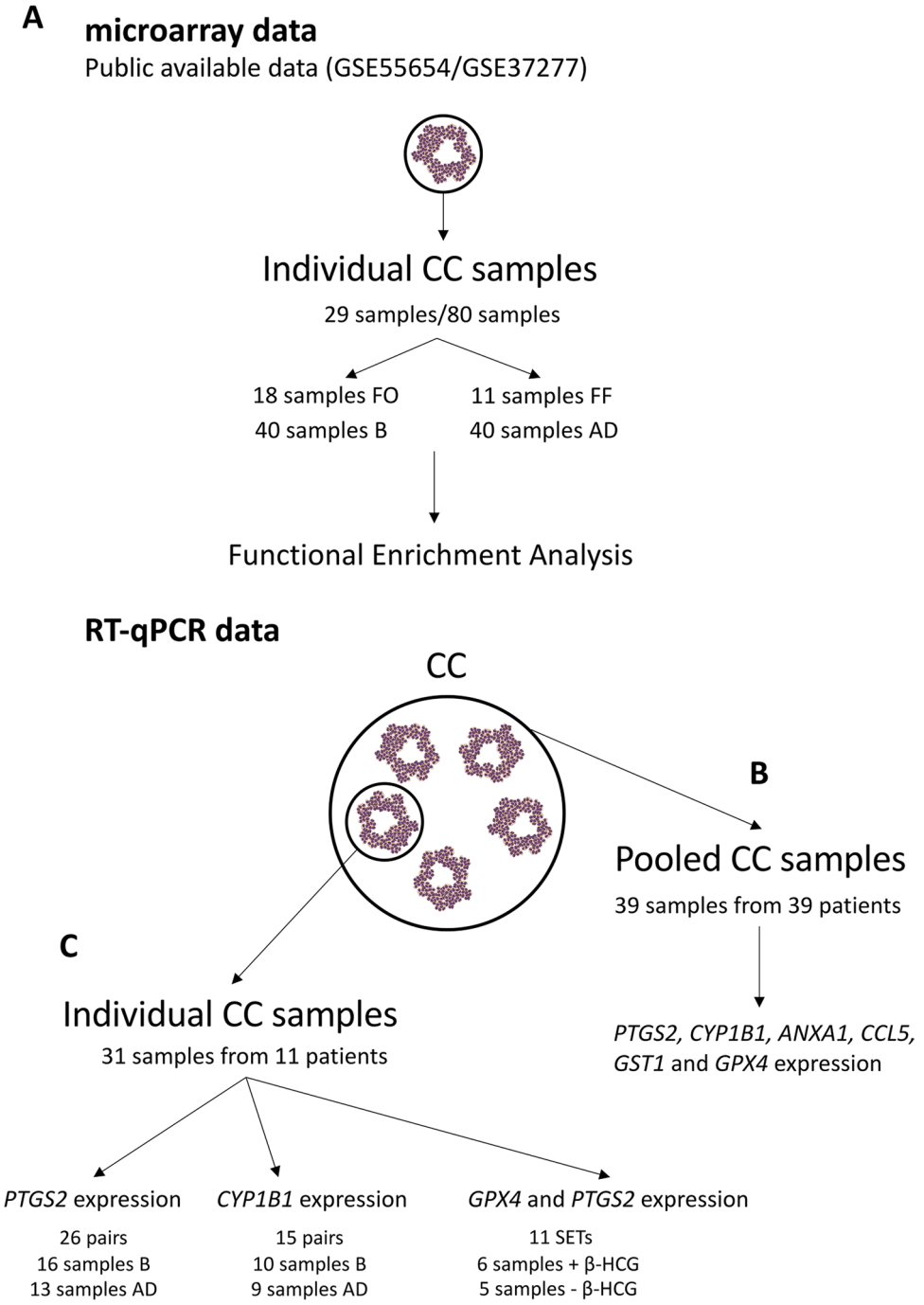

Fig. 1 Samples flowchart. (**A**) 29 and 80 individually collected samples were included in microarray data obtained from GSE55654 and GSE37277. Probes were filtered and differentially expressed genes (DEGs) were assessed through log fold change (logFC) values. Samples were categorized according to successful fertilization (FO, $n=18$) and blastocyst formation (B, $n=40$) or failed to fertilize (FF, $n=11$) and arrested development (AD, $n=40$) from the corresponding oocytes after fertilization. (**B**) 39 pooled CC samples from all follicles of 39 patients were collected for analysis of *PTGS2*, *CYP1B1*, *ANXA1*, *CCL5*, *GST1*, and *GPX4* expression levels. (**C**) Individual samples included at q-rtPCR experimental phase. 31 individually collected CC samples from 31 follicles of 11 patients were included in this study. *PTGS2* expression levels were determined in 16 samples derived from follicles whose oocytes developed into blastocysts (**B**) 5 days after fertilization and 13 samples whose oocytes presented an arrested development (AD) after fertilization. For *CYP1B1* expression levels, 10 samples were included in the B group and 9 in the AD group. *GPX4* expression levels could be analyzed in 11 samples from 11 patients submitted to single-embryo transfer (SET): 6 samples related to successful implantations (+βHCG) and 5 samples that failed to implant (−βHCG). The analysis was made through paired tests combining samples from the same patient. *PTGS2* expression levels could also be analyzed in 6 samples from 6 patients submitted to single-embryo transfer (SET): 3 samples related to successful implantations (+βHCG) and 3 samples that failed to implant (−βHCG). The analysis was made through paired tests combining samples from the same patient

Seventy CC samples from 50 patients were included in the validation phase of this study. Since individual collection of CCs requires a longer manipulation time of oocytes, two types of sample collection were used to minimize the influence of the study: pooled CC samples from all follicles of the same patient and individually collected *cumulus* complexes from each follicle. Therefore, pooled samples were used for analysis of patient-related characteristics, since those are not affected by the oocytes, and individually collected samples were used for analysis of oocyte-related and intra-patient characteristics. Thirty-nine patients provided pooled samples, with all CC being collected from all *cumulus*-oocyte complex (COC) of the patient during the same stimulation cycle. Eleven patients provided 31 individualized CC samples, composed of CC from each COC retrieved during a single stimulation cycle. Clinical information about age, infertility diagnosis, BMI, and ovarian stimulation protocol of samples were retrieved from patient records, and are listed in Table 1.

The corresponding oocytes were tracked and the number of retrieved, injected, and fertilized oocytes were computed. Embryos were tracked individually until day 5 of culture and analyzed for developmental capacity. For pooled samples, samples containing ≥ 50% CCs corresponding to oocytes that generated embryos presenting a well-defined blastocoel at day 5 and considered top quality blastocysts (using Gardner criteria) [34] were classified as blastocysts (B). Samples containing < 50% CCs corresponding to oocytes that did not generate embryos that reach blastocyst stage and did not present a blastocoel were classified as arrested development (AD). Of the 39 pooled samples, 23 corresponded to women who were submitted to embryo transfer and were analyzed for β-HCG detection at day 14 after transfer and were divided between positive and negative groups. All the corresponding embryos from the 31 individualized CC samples were accompanied until day 5 of culture after ICSI, and samples were characterized according to the embryo development potential (blastocyst formation or arrested development) and β-HCG results when single embryo transfer was performed.

Table 1 Clinocopatological information of *cumulus oophorus* samples

Pooled samples ($n = 39$)	
Age (years)	
Mean (min.–max.)	35.30 (25–42)
Median	42
Body mass index	
Mean (min.–max.)	24.62 (20.28–33.05)
Median	24.61
Diagnosis	
Male factor	14
Endometriosis	3
Polycystic ovary	4
Tubarian factor	14
Other	4
Stimulation protocol	
Short (GnRH antagonist + FSH + hMG)	26
Short (GnRH antagonist + FSH + hMG + CC)	11
Long (GnRH agonist + FSH)	2
Individual CC samples (from 10 patients) ($n = 29$)	
Age (years)	
Mean (min.–max.)	33.9 (29–39)
Median	33.5
Body mass index	
Mean (min.–max.)	24.10 (15.98–32.25)
Median	24.24
Diagnosis	
Male factor	4
Endometriosis	1
Polycystic ovary	1
Tubarian factor	4
Stimulation protocol	
Short (GnRH antagonist + FSH + hMG)	2
Short (GnRH antagonist + FSH + hMG + CC)	5
Short (GnRH agonist + FSH + CC)	1
Short (GhRH antagonist + LH)	1
Short (GnRN antagonist + hMG)	1

Thirty-nine pooled samples from 39 patients and 29 individual samples from 10 patients were considered in the experimental phase of this study. *BMI* body mass index, *GnRH* gonadotropin-releasing hormone, *FSH* follicle-stimulating hormone, *hMG* human menopausal gonadotropin, *CC* clomiphene citrate, *LH* luteinizing hormone

Ovarian Stimulation and Luteal Phase Support

Controlled ovarian stimulation followed the methods adopted by the clinic (Table 1), which consisted of short protocols, with administration of gonadotropin-releasing hormone (GnRH) antagonist (Orgalutran®, Schering-Plough, Brazil), or the GnRH agonist, with or without recombinant (Puregon®, Organon, Holland) or urinary follicle-stimulating hormone (FSH) (Fostimon®, IBSA Institut Biochimique S.A., Switzerland), and human highly purified menopausal gonadotropin (HP-hMG) (Menopur®, Ferring Pharmaceuticals, Copenhagen, Denmark), with or without clomiphene citrate (Clomid®, Medley, Brazil). Long protocol (approximately 30 days) was also used with some patients, with GnRH agonist and FSH.

Ultrasonography follow-up of the cycle initiated on the seventh day of stimulation was performed daily or at every 2 days, and the gonadotropin dose was adjusted according to the follicular growth observed (between 225 and 300 IU). About 34 to 36 h after administration of recombinant human chorionic gonadotropin (HCG) (Ovidrel®, Serono, Brazil), each patient underwent oocyte retrieval under intravenous sedation with propofol (Diprivan®,

AstraZeneca, Brazil) and fentanyl citrate (Fentanyl, Janssen-Cilag, Brazil).

Oocyte Retrieval

Follicle aspiration was performed with ultrasonography with a 5-MHz transvaginal transducer coupled to a puncture guide. Retrieved COCs were placed on cell culture plates (2004 FIV; Ingamed, Brazil) filled with human tubal fluid-HEPES culture medium (HTF) (Irvine Scientific, USA) supplemented with 10% synthetic serum substitute (SSS; Irvine Scientific), covered with mineral oil (Sigma-Aldrich, Brazil), and incubated at 37 °C in 5.8% CO_2 and 95% humidity for 2 h. After this period, the oocytes were denuded by exposure of COC to hyaluronidase (H4272 type IV-S, Sigma; 40 IU/mL) for 30 s, and CCs were mechanically removed in HTF-SSS with the aid of a stripper pipette (130 mm; Denuding Pipette, Cook). Pooled samples were provided from COCs denuded altogether in the same media drops, while individualized samples were provided from COCs denuded individually on its own media drop and using each its own disposable denuding pipette. Each drop of media containing the CCs were centrifuged (2000 *g*/10 min). After centrifugation, the supernatants were discarded and CC samples were conditioned in 500 µL TRIzol® Reagent and stored at − 80 °C until experimentation.

Intracytoplasmic Sperm Injection and Fertilization, Cleavage, Implantation, and Pregnancy Assessment

Mature oocytes characterized by the extrusion of the first polar body were submitted to ICSI 2 to 4 h after oocyte retrieval. About 16 to 18 h after ICSI, fertilization was assessed on the basis of the presence of two pronuclei and two polar bodies. Embryos were cultivated in Global® Total® culture media (LifeGlobal®, Brazil). At day 5 after ICSI, the presence of blastocoel was determined. The total percentage of retrieved mature, injected, and fertilized oocytes and cleaved and produced embryos was determined for each oocyte.

Reagents and Equipment

All reagents were obtained from Sigma-Aldrich (São Paulo, Brazil), except when indicated. All quantitative real-time polymerase chain reaction (rt-qPCR) experiments were run in 96-well plates on a StepOnePlus™ (Applied Biosystems, USA).

Total RNA Extraction and cDNA Synthesis

Total RNA was extracted from 39 pooled CC samples and 31 individualized CC samples with TRIzol® Reagent (Invitrogen, Paisley, United Kingdom), according to the manufacturer's instructions. Total RNA quality was determined spectrophotometrically using a BioPhotometer Plus (Eppendorf, Germany) and analyzing the 260/280 nm absorbance ratio.

Complementary DNA (cDNA) was synthetized from 2 µg of total RNA from each sample using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, USA) and stored at − 20 °C until use.

Quantitative Real-Time Polymerase Chain Reaction

After single-stranded cDNA synthesis, 3 µL (1:20 dilution) of the cDNA from each sample was used as a template for quantitative polymerase chain reaction (qPCR) using 1 U of enzyme Platinum® Taq DNA polymerase (5 U/µL, Invitrogen, USA), 100 nM of each specific primer and 2 mM of $MgCl_2$, and 2 µL of SYBR Green (1000X, Molecular Probes), in a final volume of 20 µL.

Oligonucleotides were selected to be RNA specific and complementary to the human sequence of annexin 1 (*ANXA1*), prostaglandin-endoperoxide synthase 2 (*PTGS2*), glutathione peroxidase 4 (*GPX4*), glutathione-*S*-transferase 1 (*GST1*), C–C motif chemokine ligand 5 (*CCL5*), cytochrome P450 family 1 subfamily B member 1 (*CYP1B1*), B-cell CLL/lymphoma 5 (*BCL5*), guanine nucleotide binding protein beta polypeptide 2-like 1 (*GNB2L1*), and hypoxanthine phosphoribosyltransferase 1 (*HPRT1*). *GNB2L1* and *HPRT1* were used as endogenous controls. Gene sequence information was collected from databases (www.ncbi.nlm.nih.gov/, www.ensembl.org/) and used for primer design with software from Integrated DNA Technologies (www.idtdna.com/).

Samples were run in triplicate in 96-well plates on a StepOnePlus™ (Applied Biosystems, USA). The thermal cycling profile for all genes was an initial denaturation step at 94 °C for 10 min followed by 40 cycles of 15 s at 94 °C, 15 s at 60 °C, and 15 s at 72 °C for data acquisition. The specificity of the amplified products was confirmed by analyzing the dissociation curves at the end of each reaction. The relative expression (real quantitative [RQ]) of the genes analyzed was calculated for each sample by the relative comparative (ΔΔCT) method [52]. Triplicates with standard deviation ≥ 0.3 threshold cycles (Ct) were excluded and rerun. A random sample pool was used as positive control in order to monitor interplacement variation. All negative controls performed as expected and all results mentioned later are normalized values.

Multiple Regression Analysis

The thirty-nine CC samples collected as a pool from all the COCs of the same patient were submitted for multiple

Table 2 Statistical models generated using a multiple stepwise regression

Model	Predictors	PR(> *F*)
Baseline	%FO ~ Age + Diag + Stim_Prot + BMI	–
ANXA1	%FO ~ Age + Diag + Stim_Prot + BMI + *ANXA1*	0.4548503672
PTGS2	%FO ~ Age + Diag + Stim_Prot + BMI + *PTGS2*	0.487972144
CYP1B1	%FO ~ Age + Diag + Stim_Prot + BMI + *CYP1B1*	–
GPX4	%FO ~ Age + Diag + Stim_Prot + BMI + *GPX4*	0.029765708
GST1	%FO ~ Age + Diag + Stim_Prot + BMI + *GST1*	0.668770931

Statistical models were generated using a multiple stepwise regression. A significant model ($P > 0.05$), containing four variables, indicated that *GPX4* expression levels are significantly different between pooled cumulus cells related to oocytes that fertilized successfully and cells corresponding to oocytes that failed to fertilize, and this significance is independent of the clinical variables of each patient. *%FO* percentage of fertilized oocytes relative to cumulus sample, *Diag* infertility diagnosis, *Stim_Prot* stimulation protocol, *BMI* body mass index, *ANXA1* annexin 1, *PTGS2* prostaglandin-endoperoxide synthase 2, *CYP1B1* cytochrome P 450 1B1, *GPX4* glutathione peroxidase 4, *GST1* glutathione-S-transferase 1

Table 3 Statistical models generated using a multiple stepwise regression

Model	Predictors	PR(> *F*)
Baseline	%Blast ~ Age + Diag + Stim_Prot + BMI	–
ANXA1	%Blast ~ Age + Diag + Stim_Prot + BMI + *ANXA1*	0.699613807
PTGS2	%Blast ~ Age + Diag + Stim_Prot + BMI + *PTGS2*	0.031927186
CYP1B1	%Blast ~ Age + Diag + Stim_Prot + BMI + *CYP1B1*	–
GPX4	%Blast ~ Age + Diag + Stim_Prot + BMI + *GPX4*	0.185043755
GST1	%Blast ~ Age + Diag + Stim_Prot + BMI + *GST1*	0.151647716

Statistical models were generated using a multiple stepwise regression. A significant model ($P > 0.05$), containing four variables, indicated that PTGS2 expression levels are significantly different in pooled cumulus cells related to embryos that reached the blastocyst stage at day 5, and this significance is independent of the clinical variables of each patient. *% Blast* percentage of blastocysts relative to cumulus sample, *Diag* infertility diagnosis, *Stim_Prot* stimulation protocol, *BMI* body mass index, *ANXA1* annexin 1, *PTGS2* prostaglandin-endoperoxide synthase 2, *CYP1B1* cytochrome P 450 1B1, *GPX4* glutathione peroxidase 4, *GST1* glutathione-S-transferase 1

regression analysis to analyze whether patients' clinical characteristics or samples' heterogeneity would affect biomarkers' performance. Clinical and experimental data were combined, and missing values were imputed applying predictive mean matching algorithm using mice package [69]. Afterwards, we built four baseline models using the *percent of fertilized oocytes*, *percent of good quality embryos*, *percent of blastocysts*, or *embryo transfer result (pregnancy)* as dependent variables against *age*, *diagnosis*, *stimulation protocol*, and *BMI* as independent variables. After we obtained the experimental data, we made four test models using the *percent of fertilized oocytes*, *percent of good quality embryos*, *percent of blastocysts*, or *embryo transfer result (pregnancy)* as dependent variables against *age*, *diagnosis*, *stimulation protocol*, *BMI*, and *rt-qPCR results* as independent variables. Finally, we compared test models composed of the four clinical variables and each rt-qPCR assay data for each gene against the baseline model (Tables 2, 3, and 4). All procedures and computations were performed in *R* statistical environment [59].

Table 4 ANOVA model comparison using molecular predictors

Model	Predictors	PR(> *F*)
Baseline	Transf ~ Age + Diag + Stim Prot + BMI	–
ANXA1	Transf ~ Age + Diag + Stim Prot + BMI + *ANXA1*	0. 730378509
PTGS2	Transf ~ Age + Diag + Stim Prot + BMI + *PTGS2*	0. 196586106
CYP1B1	Transf ~ Age + Diag + Stim Prot + BMI + *CYP1B1*	–
GPX4	Transf ~ Age + Diag + Stim Prot + BMI + *GPX4*	0. 010541996
GST1	Transf ~ Age + Diag + Stim Prot + BMI + *GST1*	0. 739441369

Statistical models were generated using a multiple stepwise regression. A significant model ($P > 0.05$), containing four variables, indicated that *GPX4* expression levels are significantly different in pooled cumulus cells from patients with successful embryo transfer, and this significance is independent of the clinical variables of each patient. *Transf* embryo transfer result, *Diag* infertility diagnosis, *Stim Prot* stimulation protocol, *BMI* body mass index, *ANXA1* annexin 1, *PTGS2* prostaglandin-endoperoxide synthase 2, *CYP1B1* cytochrome P 450 1B1, *GPX4* glutathione peroxidase 4, *GST1* glutathione-S-transferase 1

Statistical Analysis

Experimental data were expressed as means ± SD and *P* values were considered significant for $P < 0.05$.

For pooled samples, the influence of patients and oocyte characteristics in gene expression was determined by one-way analysis of variance (ANOVA). For individual samples, all the 11 patients included in this phase of the study presented CC samples corresponding to COCs yielding an oocyte of good quality (B) and samples from follicles from poor-quality oocytes (AD), each patient ranging from 1 to 4 B samples and 1 to 3 AD samples. Therefore, for each patient, it was possible to combine between several different pairs (B *vs*. AD). A paired Wilcoxon test was used, based on group normality, analyzed through D'Agostino-Pearson and Shapiro–Wilk normality tests (GraphPad® Software 5.0).

Results

Differentially Expressed Genes and Functional Enrichment Analysis

Our metadata analysis of microarray datasets identified differentially expressed genes associated with oocyte competence that were further investigated for enriched biological processes in each outcome (Fig. 2). In samples from oocytes that successfully fertilized (good quality) or failed to fertilize (poor quality), obtained from GSE55654 GEO dataset, a plethora of cellular processes were observed (Tables S1 and S2 and Fig. 2). Notably, CCs associated with oocytes that failed to fertilize showed enrichment of biological process related to oxidative stress (GO0006979) (Fig. 2B).

In samples from oocytes that further generated blastocysts (B) *vs*. embryos that presented AD (raw data extracted from GSE37277 GEO dataset), genes associated with tissue development (GO0009888) and cell differentiation (GO0045597) were highlighted (Table S3 and Fig. 3A). On the other hand, in the AD group, we identified differentially expressed genes related to steroid metabolism (GO0008202), response to external stimulus (GO0009605), response to organic cyclic compounds (GO0014070), response to oxygen-containing compounds (GO1901701), response to nutrients (GO0007584), response to chemical stimulus (GO0070887), and response to stress (GO0006950) (Table S4 and Fig. 3B). Thus, in a general manner, CCs related to good-quality oocytes presented augumented development processes, while the ones related to bad-quality oocytes seems to be responding to an inhospitable, stressful environment.

Validation of Selected Targets by Real-Time Quantitative PCR

Among the differentially expressed genes obtained from metadata analysis of microarray datasets, six were selected for *in house* validation in clinical samples by rt-qPCR: genes involved in the majority of each GQ group, being *CCL5* for GSE55465 dataset, involved in 43 of the 45 processes observed (Table S1), and *PTGS2*, for GSE37277 dataset, involved in all 6 processes observed (Table S3), and genes involved in the majority of processes in each PQ group, being *ANXA1* for GSE55465 dataset, involved in 25 of the 40 processes observed (Table S2), and *CYP1B1* for GSE37277 dataset, involved in 17 of the 24 processes observed (Table S4). Two other genes, glutathione-*S*-transferase (*GST1*) and glutathione peroxidase (*GPX4*), were chosen for analysis based on the processes highlighted in both PQ groups, as response to oxidative stress (GO0006979), for GSE55465 dataset (Table S2), and RESPONSE to oxygen-containing compound (GO1901700), for GSE37277 dataset (Table S4), and based on literature [25, 46, 72].

Pooled CC Samples

Thirty-nine pooled CC samples were used for rt-qPCR validation. Of the 6 genes analyzed, two (*CCL5* and *CYP1B1*) were not properly detected in pooled samples (possibly due to primer design) and were not further analyzed.

Patients' clinical data are listed in Table 1. Pooled samples were classified according to (1) fertilization rates: samples were included in the fertilized oocytes (FO) group, if more than 75% of the corresponding oocytes fertilized successfully ($n = 24$), or in the failed to fertilize (FF) group, if less than 25% of the corresponding oocytes fertilized successfully ($n = 15$); (2) blastocyst formation rates: samples were included in the blastocyst group (B) ($n = 13$), if the pooled CC sample corresponded to oocytes that generated 50% or more of blastocysts, or in the AD group ($n = 26$), if less than 50% of the corresponding oocytes generated blastocysts after ICSI; and (3) implantation potential: samples were included in positive (P) ($n = 8$) or negative (N) ($n = 15$) groups based on β-HCG results after embryo transfer.

Mann–Whitney tests did not show statistical differences between groups for any gene. However, gene expression results were submitted to multiple regression models considering the patient's clinical data, to determine if the observed results were confused by patient's clinical profiles. This analysis revealed a series of differences in gene expression levels between groups.

A significant model containing four variables indicated that *GPX4* expression levels are altered depending on oocyte quality, being significantly lower in CCs from good-quality oocytes, with better fertilizing rates ($P = 0.0297$) (Table 2)

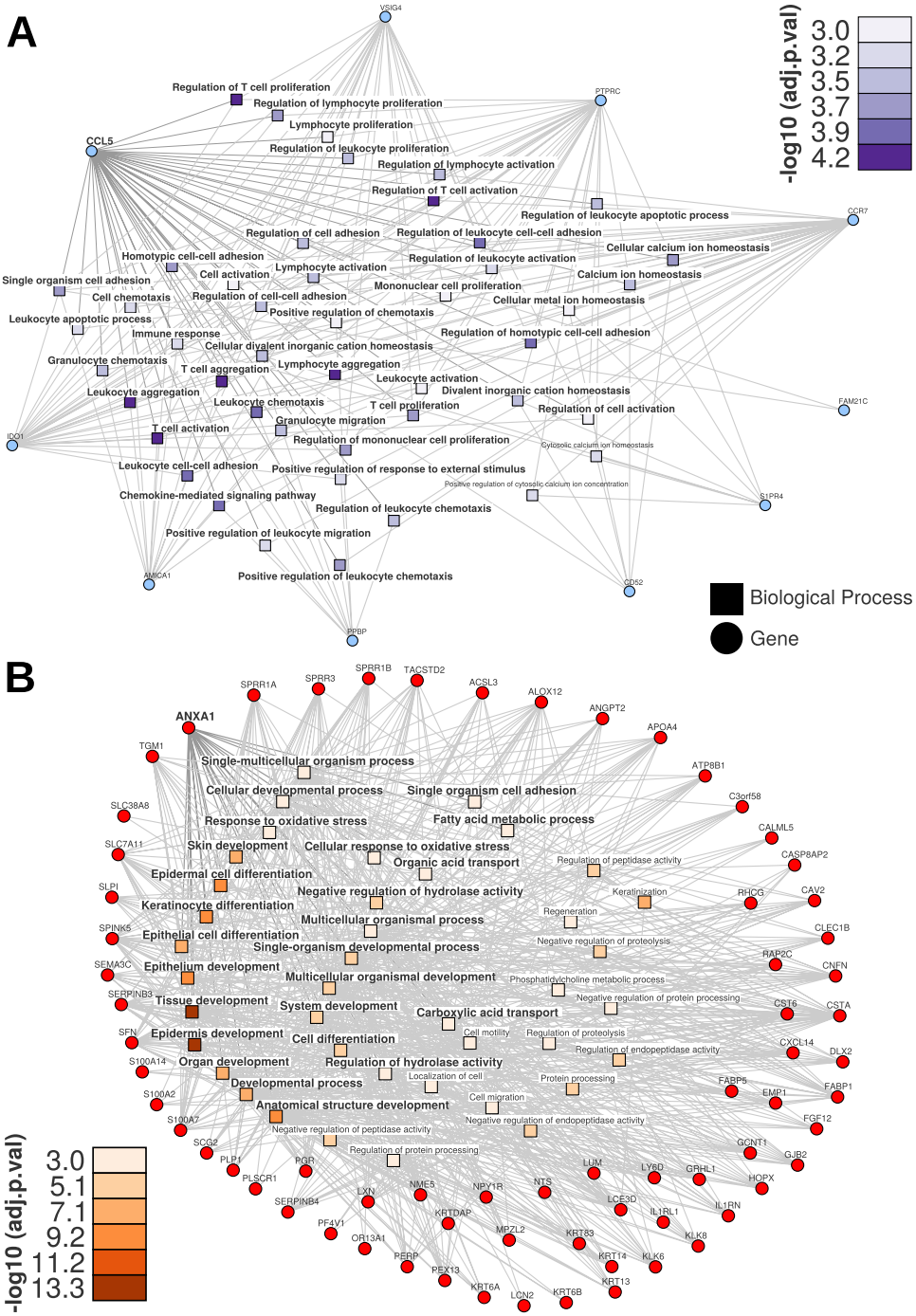

Fig. 2 Biological profile of *cumulus* cells associated with oocyte fertilization potential (GSE55654). Microarray data were obtained from publicly available databanks. Probes were filtered and differentially expressed genes (DEGs) were assessed through log fold change (logFC) values. FGNet package was employed for biological processes. (**A**) Biological processes overexpressed in *cumulus* cells (CCs) of oocytes in the good-quality group that fertilized successfully ($n=18$). (**B**) Biological processes overexpressed in CC of oocytes in the poor-quality group that failed to fertilize ($n=11$). In squares are represented biological processes whereas in circles are the genes

and in positive β-HCG group, with better pregnancy rates ($P=0.0105$) (Table 4). *PTGS2* was shown to be a potential blastocyst development biomarker ($P=0.0319$) (Table 3).

Individualized CC Samples

Based on rt-qPCR results of pooled samples, 31 individually collected CC samples from 11 patients were retrieved for validation of the potential biomarkers. The paired analysis revealed that CC associated with blastocyst-stage embryos have higher levels of *PTGS2* gene expression than CC samples associated with arrested development (poor-quality group) ($P=0.0002$) from the same patient (Fig. 3A). On the other hand, CC associated with blastocyst-stage embryos had lower levels of *CYP1B1* expression ($P=0.0084$) than poor-quality CC samples (Fig. 3B), indicating *CYP1B1* as a possible biomarker of development failure. These results are in accordance with the bioinformatics analysis of microarray data (Fig. 3, Tables S3 and S4). Paired analysis comparing samples from the same

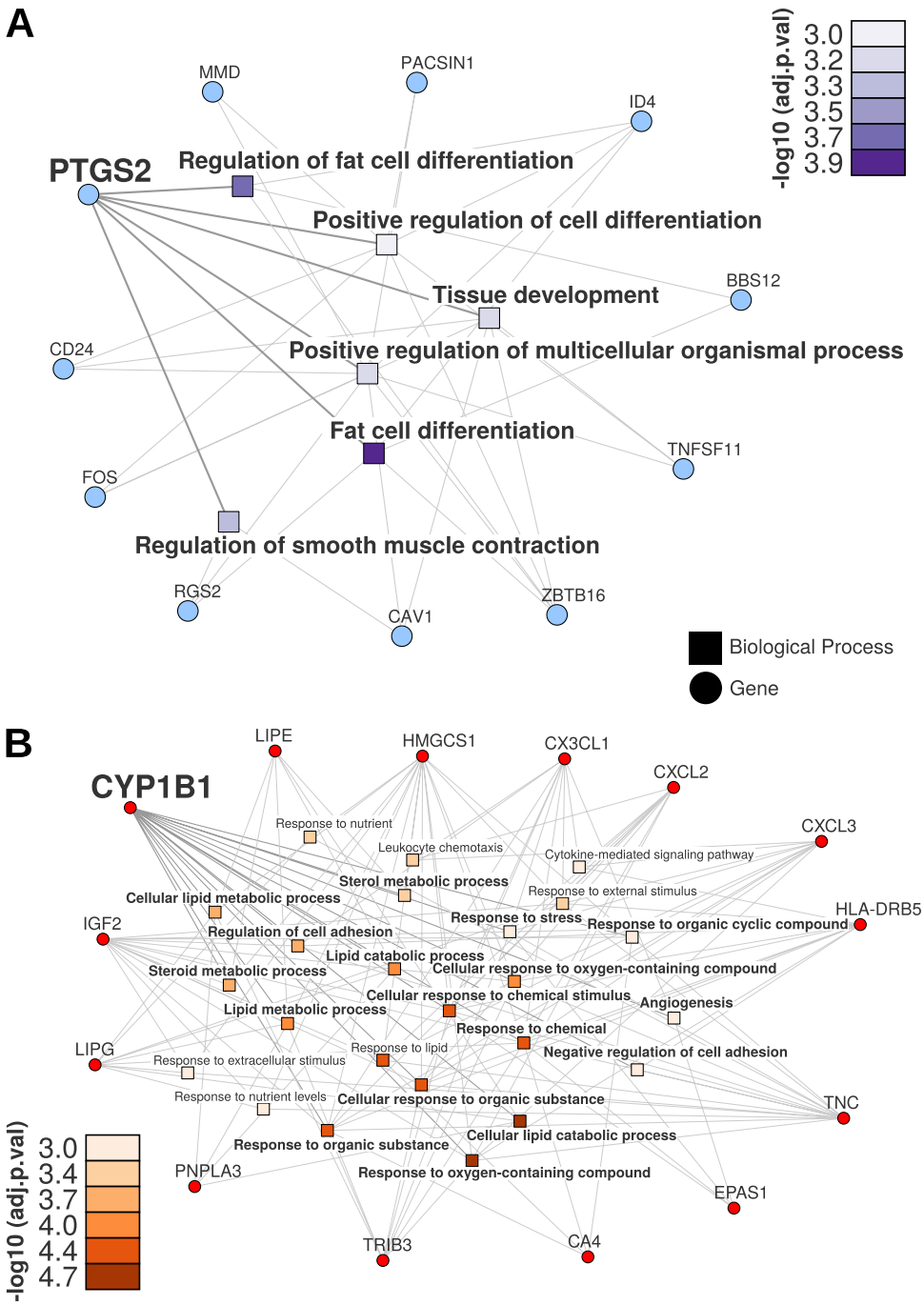

Fig. 3 Biological profile of *cumulus* cells associated with blastocyst formation (GSE37277). Microarray data were obtained from publicly available databanks. Probes were filtered and differentially expressed genes (DEGs) were assessed through log fold change (logFC) values. FGNet package was employed for biological processes. (**A**) Biological processes overexpressed in *cumulus* cells (CCs) of oocytes in the good-quality group ($n=40$). (**B**) Biological processes overexpressed in CC of oocytes in the poor-quality group that generated embryos with arrested development ($n=40$). In squares are represented biological processes whereas in circle are the genes

patient but with opposite outcomes reveals that the differential gene expression observed is directly related to oocyte quality and not a characteristic of the patient herself. During the time course of this study, only a few single embryo transfers were performed at the clinics. Thus, the biomarker potential of *GPX4* expression levels for embryo implantation was assessed in 11 individually collected CC samples related to oocytes that further generated blastocysts and were single transferred. *GPX4* levels were shown to be overexpressed in samples related to blastocysts that failed to implant ($P=0.0296$) (Fig. 4C), confirming the tendency observed for this proposed biomarker in the bioinformatic analysis and in pooled samples. Likewise, the biomarker potential of *PTGS2* expression levels for embryo implantation was assessed in 6 individually collected CC samples related to oocytes that further generated blastocysts and were single transferred. *PTGS2* levels were shown to be overexpressed in samples related to blastocysts that successfully implanted ($P=0.0207$) (Fig. 4D).

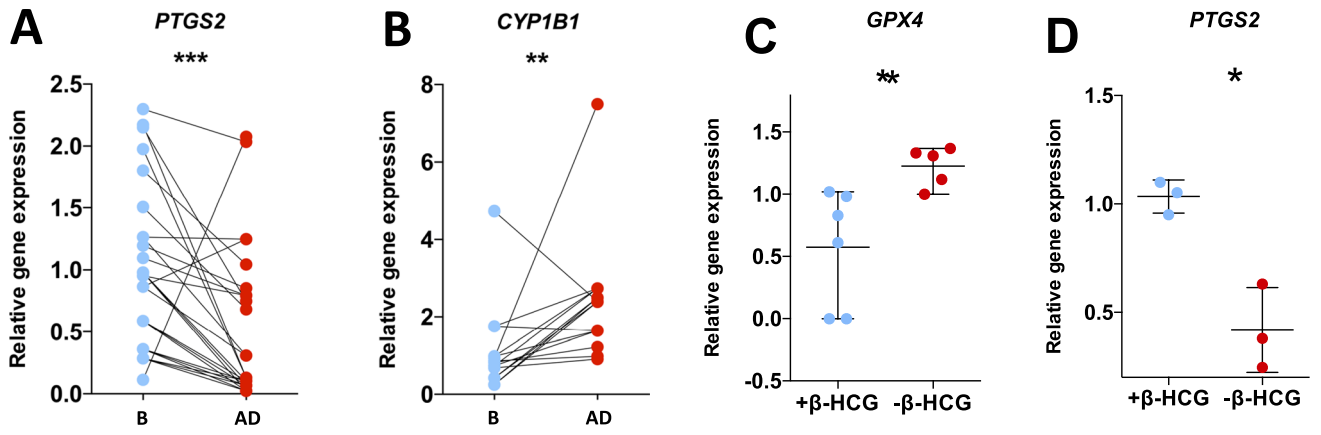

Fig. 4 Gene expression analysis on *cumulus* samples related to oocyte quality. Paired *cumulus* cells (CCs) samples from individual follicles related to embryos with arrested development (AD) (red) and blastocyst (B) embryos (blue), from 10 patients, were analyzed for **A** cytochrome P 450 1B1 (*CYP1B1*) ($n=15$ pairs) and **B** prostaglandin-endoperoxide synthase 2 (*PTGS2*) ($n=26$ pairs) expression, relative to housekeeping gene *HPRT1*. **C** CC samples from individual follicles related to embryos that implanted successfully after single embryo transfer (SET) (+βHCG, blue) and that failed to implant (−βHCG, red) were analyzed for *GPX4* expression levels, relative to housekeeping gene *HPRT1*. **D** CC samples from individual follicles related to embryos that implanted successfully after single-embryo transfer (SET) (+βHCG, blue) and that failed to implant (−βHCG, red) were analyzed for *PTGS2* expression levels, relative to housekeeping gene *HPRT1*. *$P=0.0296$ (*GPX4*) and $=0.0207$ (*PTGS2*); (unpaired t test) **$P=0.0084$; ***$P=0.0002$ (non-parametric paired Wilcoxon test)

Discussion

In spite of relevant advances in infertility treatments, success rates on in vitro fertilization techniques are still suboptimal. Tools that enable the selection of good-quality oocytes in clinical environment are greatly desired. Here, we proposed the use of bioinformatic tools such as functional enrichment analyses for identification of oocyte quality biomarkers and validated the approach considering patients' characteristics in the clinical scenario. CC expression profile may reveal existing alterations in corresponding oocytes, such as accumulation of oxidative damage, high expression of detoxification machinery, and aneuploidies [33]. Through the analysis of databases, bioinformatic tools are capable of revealing CC responses to the intra-follicular environment [41], such as signals from the oocyte. Through the study of CC gene expression data, our bioinformatic evaluation revealed a series of differentially expressed cellular biological processes related to embryo competence.

Using the same dataset (GSE37277), but selecting targets based on the most differentially expressed genes ($P<0.0001$), Feurestein [30] identified 3 different potential biomarkers. Our bioinformatic analysis aimed to firstly define a threshold for differentially expressed genes, and secondly, to identify the pattern of the cell processes those genes compound, to then select genes involved in multiple processes, despite of their expression levels, the rationale being that the contribution of genes to cellular processes are not necessarily correlated with the fold change of their expression, causing the analysis to highlight distinct biomarkers even when using the same dataset.

The two independent databases selected were analyzed under the same criteria (good- and bad-quality oocytes), but a much larger set of processes was revealed as significantly expressed between groups when fertilization potential was chosen as endpoint (Fig. 2). Since fertilization is the earliest event in embryo formation, it requires several basic biological processes to be properly functioning. Embryos that presented arrested development after this stage, on the other hand, also presented functional basic processes before arresting. Therefore, it is understandable that the earlier in the development timeline the endpoint is defined, the longer the list of processes that are differentially expressed between groups.

In CC obtained from oocytes that generated top-quality blastocysts after fertilization, we observed many cellular processes related to tissue development and cell differentiation, indicating an appropriate environment that supports cell growth and plasticity (Fig. 3A), whereas among the processes upregulated in samples related to oocytes that failed to fertilize and in samples related to oocytes that did not develop into blastocysts (Figs. 2B and 3B), one can find redox metabolism processes, steroid metabolism, response to external stimulus, and response to organic compounds, but also processes as response to oxygen-containing compounds, response to chemical stimulus, and response to stress that could be revealing a potentially toxic environment where these oocytes were exposed to. Based on this similarity between both groups with low-quality oocytes, *GPX4* and *GST1* were selected

as candidate biomarkers. This selection was also based on previous works that successfully correlated oxidative metabolism enzymes with oocyte quality [54, 64, 68].

In order to find representations of the processes observed, the genes involved in most of the processes in each of the four groups were investigated further (*CCL5*, *PTGS2*, *ANXA1*, and *CYP1B1*). *PTGS2* is involved in all six processes observed in the blastocyst sample group (Table S3 and Fig. 2A) and was previously associated with oocyte quality [55]. *PTGS2* has already been correlated with high fertilization rates, embryo development [7, 55, 74], and live birth rates [35]. While *PTGS2* expression positively correlates with blastocyst formation ($P = 0.0002$) (Fig. 4A) and pregnancy ($P = 0.0207$) (Fig. 4D), higher *CYP1B1* expression levels were detected in CC samples from oocytes that further generated embryos that presented arrested development ($P = 0.0084$) (Fig. 4B).

In this study, *GPX4* expression was significantly lower in CCs from good-quality oocytes, with better fertilizing and pregnancy rates. During maturation from germinal vesicle to MII state, oocytes do not synthesize mRNAs and therefore depend directly on stored mRNAs and CC mRNAs that are transferred through GAP junctions [53]. GPx mRNA has been detected in mature human oocytes but was shown to be absent in germinal vesicle oocytes, which implies it comes from other cells. The mRNA molecules suffer a specific "last minute" polyadenylation in mature oocytes, suggesting the gamete's recruitment of GPx mRNAs when needed [26]. Still, *GPX4* expression as a pregnancy potential biomarker is indicated here as a possibility for further investigations. Nevertheless, pregnancy is a highly complex event, and CC biomarkers should be used carefully when considering this endpoint. Based on this, a combined set of biomarkers is probably much more biologically relevant than a single marker.

Studying CC and its relationship with oocyte quality could provide valuable tools for improving ART rates. Nevertheless, it is paramount to also take into consideration clinical characteristics known to affect reproduction success in order to select markers that are applicable to the clinical environment [72]. Other promising biomarkers have presented contradictory results [7, 19, 76] [2, 29], and its oocyte quality biomarker potential was not verified when clinical variables were considered in the analysis [30]. The paired analysis applied combined with the ANOVA analysis in this study considers the influence of clinical characteristics (age, infertility diagnosis, and BMI) and pooled samples characteristics (percentage of mature oocytes, percentage of fertilized oocytes, percentage of discarded embryos) in gene expression levels and brings down to a minimum the contribution of possible confounding factors.

A large randomized prospective study composed by a diverse group of patients submitted to single embryo transfer is needed to further establish *PTGS2*, *CYP1B1*, and *GPX4* as biomarkers of oocyte quality, embryo development, and implantation potential. Using a combination of biomarkers that act in opposite directions allows a refined prediction model. By analyzing data from microarray sets from our group, it was stated that *PTGS2* and *CYP1B1* genes also follow the same expression pattern in CCs related to oocytes that were successful in generating a pregnancy, *PTGS2* being significantly overexpressed, while *CYP1B1* being down regulated in comparison to samples of embryos that failed to implant (unpublished data). This indicates that it is possible *CYP1B1* could also be related to pregnancy potential.

Thus, this study confirms that the bioinformatics approach is suitable for finding relevant biomarkers, and can be used in reproductive sciences, opening new venues in assisted reproduction. Also, a previous suggested biomarker is confirmed as blastocyst predictor and novel oocyte quality biomarkers are proposed, highlighting processes that are desirably or undesirably upregulated in follicles.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s43032-022-01046-2.

Acknowledgements We thank Publicase for the comments made during the preparation of the manuscript.

Author Contribution Lucia von Mengden developed the project, performed the experiments, analyzed the data, and prepared the manuscript. Marco Antônio De Bastiani participated in project development, performed the meta-analysis, analyzed the data, and participated in the discussion. Lucas Kich Grun participated in project development, performed the experiments, and participated in the discussion. Florencia Barbé-Tuana participated in project development and supervised the experiments. Tom Adriaenssens contributed to the data analysis and discussion. Johan Smitz contributed to the data analysis and discussion. Leticia Schmidt Arruda participated in project development, collected the samples and patient data, and participated in the discussion. Carlos Alberto Link participated in project development, collected the samples and patient data, and participated in the discussion. Fábio Klamt participated in project development, provided funding, contributed to manuscript development, and supervised the project.

Funding This study was supported by the Brazilian funds MCTI/CNPq INCT-TM/CAPES/FAPESP (465458/2014–9), PRONEX/FAPERGS (16/2551–0000499-4), and PDSE-CAPES 47/2017 (88881.188914/2018–01).

Data Availability The authors declare that data supporting the findings of this study are available within the paper and its supplementary information files.

Declarations

All authors consented to participate of the construction and development of this research. All authors consented to publish this research.

Ethics Approval and Trial Registration This study was approved by the Research Ethics Committee of the Federal University of Rio Grande do Sul (UFRGS) and by Brazilian's National Research Ethics Comitee (CONEP) under trial number (#68081017.2.0000.5347).

Competing Interests The authors declare no competing interests.

References

1. Adriaenssens T, Van Vaerenbergh I, Coucke W, Segers I, Verheyen G, Anckaert E, De Vos M, Smitz J. Cumulus-corona gene expression analysis combined with morphological embryo scoring in single embryo transfer cycles increases live birth after fresh transfer and decreases time to pregnancy. J Assist Reprod Genet. 2019;36:433–43. https://doi.org/10.1007/s10815-018-01398-2.
2. Adriaenssens T, Wathlet S, Segers I, Verheyen G, De Vos A, Van Der Elst J, Coucke W, Devroey P, Smitz J. Cumulus cell gene expression is associated with oocyte developmental quality and influenced by patient and treatment characteristics. Hum Reprod. 2010;25:1259–70. https://doi.org/10.1093/humrep/deq049.
3. Aibar S, Fontanillo C, Droste C, De Las RJ. Functional gene networks: R/Bioc package to generate and analyse gene networks derived from functional enrichment and clustering. Bioinformatics. 2015;31:1686–8. https://doi.org/10.1093/bioinformatics/btu864.
4. Akin N, von Mengden L, Herta A-C, Billooye K, Leersum J, Cava-Cami B, Saucedo-Cuevas L, Klamt F, Smitz J, Anckaert E. Glucose metabolism characterization during mouse in vitro maturation identifies alterations in cumulus cells. Biol Reprod. 2021:1–12. https://doi.org/10.1093/biolre/ioab008.
5. Albertini DF et al. Cellular basis for paracrine regulation of ovarian follicle development. Reproduction (Cambridge, England). (2001);121(5): 647–53. https://doi.org/10.1530/rep.0.1210647
6. Alexa A. Gene set enrichment analysis with topGO. 2007.
7. Anderson RA, Sciorio R, Kinnell H, Bayne RAL, Thong KJ, De Sousa PA, Pickering S. Cumulus gene expression as a predictor of human oocyte fertilisation, embryo development and competence to establish a pregnancy. Reproduction (Cambridge, England). 2009;138(4):629–37. https://doi.org/10.1530/REP-09-0144.
8. Assou S, Haouzi D, Dechaud H, Gala A, Ferrières A, Hamamah S. Comparative gene expression profiling in human cumulus cells according to ovarian gonadotropin treatments. Biomed Res Int. 2013. https://doi.org/10.1155/2013/354582.
9. Assou S, Haouzi D, Mahmoud K, Aouacheria A, Guillemin Y, Pantesco V, Rème T, Dechaud H, De Vos J, Hamamah S. A non-invasive test for assessing embryo potential by gene expression profiles of human cumulus cells: a proof of concept study. Mol Hum Reprod. 2008;14:711–9. https://doi.org/10.1093/molehr/gan067.
10. Balaban B, Urman B. Effect of oocyte morphology on embryo development and implantation. Reprod Biomed Online. 2006;12:608–15. https://doi.org/10.1016/S1472-6483(10)61187-X.
11. Balaban B, Urman B, Sertac A, Alatas C, Aksoy S, Mercan R. Oocyte morphology does not affect fertilization rate, embryo quality and implantation rate after intracytoplasmic sperm injection. Hum Reprod. 1998;13:3431–3. https://doi.org/10.1093/humrep/13.12.3431.
12. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A. NCBI GEO: archive for functional genomics data sets - Update. Nucleic Acids Res. 2013. https://doi.org/10.1093/nar/gks1193.
13. Becker M, De Bastiani MA, Parisi MM, Guma FTCR, Markoski MM, Castro MAA, Kaplan MH, Barbé-Tuana FM, Klamt F. Integrated transcriptomics establish macrophage polarization signatures and have potential applications for clinical health and disease. Sci Rep. 2015. https://doi.org/10.1038/srep13351.
14. Bellaver B, Rocha AS, Souza DG, Leffa DT, De Bastiani MA, Schu G, Lukasewicz Ferreira PC, Venturin GT, Greggio S, Ribeiro CT, da Costa JC, Moreira JCF, Gelain DP, Torres IL da S, Klamt F, Zimmer ER. Activated peripheral blood mononuclear cell mediators trigger astrocyte reactivity. Brain Behav Immun. 2019;80:879–88. https://doi.org/10.1016/j.bbi.2019.05.041.
15. Bouckenheimer J, Fauque P, Lecellier CH, Bruno C, Commes T, Lemaître JM, De Vos J, Assou S. Differential long non-coding RNA expression profiles in human oocytes and cumulus cells. Sci Rep. 2018;8:1–13. https://doi.org/10.1038/s41598-018-20727-0.
16. Burnik Papler T, Vrtacnik Bokal E, Lovrecic L, Kopitar AN, Maver A. No specific gene expression signature in human granulosa and cumulus cells for prediction of oocyte fertilisation and embryo implantation. PLoS ONE. 2015;10:e0115865. https://doi.org/10.1371/journal.pone.0115865.
17. Burnik Papler T, Vrtačnik Bokal E, Prosenc Zmrzljak U, Stimpfel M, Laganà AS, Ghezzi F, Jančar N. PGR and PTX3 gene expression in cumulus cells from obese and normal weighting women after administration of long-acting recombinant follicle-stimulating hormone for controlled ovarian stimulation. Arch Gynecol Obstet. 2019;299:863–71. https://doi.org/10.1007/s00404-018-5031-y.
18. Castro MAA, Wang X, Fletcher MNC, Meyer KB, Markowetz F. RedeR: R/Bioconductor package for representing modular structures, nested networks and multiple levels of hierarchical associations. Genome Biol. 2012. https://doi.org/10.1186/gb-2012-13-4-r29.
19. Cillo F, Brevini TAL, Antonini S, Paffoni A, Ragni G, Gandolfi F. Association between human oocyte developmental competence and expression levels of some cumulus genes. Reproduction. 2007. https://doi.org/10.1530/REP-07-0182.
20. Da Motta LL, Ledaki I, Purshouse K, Haider S, De Bastiani MA, Baban D, Morotti M, Steers G, Wigfield S, Bridges E, Li JL, Knapp S, Ebner D, Klamt F, Harris AL, McIntyre A. The BET inhibitor JQ1 selectively impairs tumour response to hypoxia and downregulates CA9 and angiogenesis in triple negative breast cancer. Oncogene. 2017;36:122–32. https://doi.org/10.1038/onc.2016.184.
21. De Bastiani MA, Pfaffenseller B, Klamt F. Master regulators connectivity map: a transcription factors-centered approach to drug repositioning. Front Pharmacol. 2018. https://doi.org/10.3389/fphar.2018.00697.
22. De Conto E, Matte U, Cunha-Filho JS. BMP-6 and SMAD4 gene expression is altered in cumulus cells from women with endometriosis-associated infertility. Acta Obstet Gynecol Scand. 2021;100:868–75. https://doi.org/10.1111/aogs.13931.
23. Devjak R, Burnik Papler T, Verdenik I, Fon Tacer K, Vrtačnik Bokal E. Embryo quality predictive models based on cumulus cells gene expression. Balk J Med Genet. 2016;19:5–12. https://doi.org/10.1515/bjmg-2016-0001.
24. Dibyajyoti S, Talha Bin E, Swati P. Bioinformatics:the effects on the cost of drug discovery. Gall Med J. 2013;18:44. https://doi.org/10.4038/gmj.v18i1.5511.
25. Donabela FC, Meola J, Padovan CC, De Paz CCP, Navarro PA. Higher SOD1 gene expression in cumulus cells from infertile women with moderate and severe endometriosis. Reprod Sci. 2015;22:1452–60. https://doi.org/10.1177/1933719115585146.
26. El Mouatassim S, Guérin P, Ménézo Y. Expression of genes encoding antioxidant enzymes in human and mouse oocytes during the final stages of maturation. Mol Hum Reprod. 1999;5:720–5. https://doi.org/10.1093/molehr/5.8.720.
27. Ekart J, McNatty K, Hutton J, Pitman J. Ranking and selection of MII oocytes in human ICSI cycles using gene expression levels from associated cumulus cells. Hum Reprod. 2013;28:2930–42. https://doi.org/10.1093/humrep/det357.
28. Faramarzi A, Khalili MA, Jahromi MG. Is there any correlation between apoptotic genes expression in cumulus cells with embryo morphokinetics? Mol Biol Rep. 2019;46:3663–70. https://doi.org/10.1007/s11033-019-04781-z.

29. Feuerstein P, Cadoret V, Guerif F, Bidault R, Royere D. Gene expression in human cumulus cells : one approach to oocyte competence. Hum Reprod. 2007;22:3069–77.
30. Feuerstein P, Puard V, Chevalier C, Teusan R, Cadoret V, Guerif F, Houlgatte R, Royere D. Genomic assessment of human cumulus cell marker genes as predictors of oocyte developmental competence: Impact of various experimental factors. PLoS One. 2012;7.https://doi.org/10.1371/journal.pone.0040449.
31. Fragouli E, Lalioti MD, Wells D. The transcriptome of follicular cells: biological insights and clinical implications for the treatment of infertility. Hum Reprod Update. 2014.
32. Fragouli E, Wells D. Transcriptomic analysis of follicular cells provides information on the chromosomal status and competence of unfertilized oocytes. Expert Rev Mol Diagn. 2012;12:1–4. https://doi.org/10.1586/erm.11.79.
33. Fragouli E, Wells D, Iager AE, Kayisli UA, Patrizio P. Alteration of gene expression in human cumulus cells as a potential indicator of oocyte aneuploidy. Hum Reprod. 2012;27:2559–68. https://doi.org/10.1093/humrep/des170.
34. Gardner DK, Schoolcraft WB. Culture and transfer of human blastocysts. Curr Opin Obstet Gynecol. 1999;11:307–11. https://doi.org/10.1097/00001703-199906000-00013.
35. Gebhardt KM, Feil DK, Dunning KR, Lane M, Russell DL. Human cumulus cell gene expression as a biomarker of pregnancy outcome after single embryo transfer. Fertil Steril. 2011;96:47-52.e2. https://doi.org/10.1016/j.fertnstert.2011.04.033.
36. Gentleman AR, Carey V, Huber W, Hahne F, Bioconductor M, Maintainer P. Genefilter: methods for filtering genes from high-throughput experiments. Version 1.63.0. 2018.
37. Green KA, Franasiak JM, Werner MD, Tao X, Landis JN, Scott RT, Treff NR. Cumulus cell transcriptome profiling is not predictive of live birth after in vitro fertilization: a paired analysis of euploid sibling blastocysts. Fertil Steril. 2018;109:460–6. https://doi.org/10.1016/j.fertnstert.2017.11.002.
38. Hammond ER, Stewart B, Peek JC, Shelling AN, Cree LM. Assessing embryo quality by combining non-invasive markers: early time-lapse parameters reflect gene expression in associated cumulus cells. Hum Reprod. 2015;30:1850–60. https://doi.org/10.1093/humrep/dev121.
39. Haouzi D, Assou S, Monzo C, Vincens C, Dechaud H, Hamamah S. Altered gene expression profile in cumulus cells of mature MII oocytes from patients with polycystic ovary syndrome. Hum Reprod. 2012;27:3523–30. https://doi.org/10.1093/humrep/des325.
40. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4:44–57. https://doi.org/10.1038/nprot.2008.211.
41. Huang Z, Wells D. The human oocyte and cumulus cells relationship: new insights from the cumulus cell transcriptome. Mol Hum Reprod. 2010;16:715–25. https://doi.org/10.1093/molehr/gaq031.
42. Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, Bravo HC, Davis S, Gatto L, Girke T, Gottardo R, Hahne F, Hansen KD, Irizarry RA, Lawrence M, Love MI, MaCdonald J, Obenchain V, Oleš AK, Pagès H, Reyes A, Shannon P, Smyth GK, Tenenbaum D, Waldron L, Morgan M. Orchestrating high-throughput genomic analysis with Bioconductor. Nat Methods. 2015.
43. Human Fertilisation & Embryology Authority. Fertility treatment 2018: trends and figures UK statistics for IVF and DI treatment, storage, and donation. 2020.
44. Hutt KJ, Albertini DF. An oocentric view of folliculogenesis and embryogenesis. Reprod Biomed Online. 2007.
45. Iager AE, Kocabas AM, Otu HH, Ruppel P, Langerveld A, Schnarr P, Suarez M, Jarrett JC, Conaghan J, Rosa GJM, Fernández E, Rawlins RG, Cibelli JB, Crosby JA. Identification of a novel gene set in human cumulus cells predictive of an oocyte's pregnancy potential. Fertil Steril. 2013;99. https://doi.org/10.1016/j.fertnstert.2012.10.041.
46. Imai H, Nakagawa Y. Biological significance of phospholipid hydroperoxide glutathione peroxidase (PHGPx, GPx4) in mammalian cells. Free Radic Biol Med. 2003.
47. Jin G, Wong STC. Toward better drug repositioning: prioritizing and integrating existing methods into efficient pipelines. Drug Discov Today. 2014;19:637–44. https://doi.org/10.1016/j.drudis.2013.11.005.
48. Kahraman S, Çetinkaya CP, Çetinkaya M, Tüfekçi MA, Ekmekçi CG, Montag M. Is there a correlation between follicle size and gene expression in cumulus cells and is gene expression an indicator of embryo development? Reprod Biol Endocrinol. 2018;16:1–10. https://doi.org/10.1186/s12958-018-0388-0.
49. Kenigsberg S, Bentov Y, Chalifa-caspi V, Potashnik G, Ofir R, Birk OS. Gene expression microarray profiles of cumulus cells in lean and overweight- obese polycystic ovary syndrome patients. Mol Hum Reprod. 2009;15:89–103. https://doi.org/10.1093/molehr/gan082.
50. Li SH, Lin MH, Hwu YM, Lu CH, Yeh LY, Chen YJ, Lee RKK. Correlation of cumulus gene expression of GJA1, PRSS35, PTX3, and SERPINE2 with oocyte maturation, fertilization, and embryo development. Reprod Biol Endocrinol. 2015;13:1–8. https://doi.org/10.1186/s12958-015-0091-3.
51. Liu Q, Zhang J, Wen H, Feng Y, Zhang X, Xiang H, Cao Y, Tong X, Ji Y, Xue Z. Analyzing the transcriptome profile of human cumulus cells related to embryo quality via RNA sequencing. Biomed Res Int. 2018. https://doi.org/10.1155/2018/9846274.
52. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. Methods. 2001;25:402–8. https://doi.org/10.1006/meth.2001.1262.
53. Macaulay AD, Gilbert I, Scantland S, Fournier E, Ashkar F, Bastien A, Saadi HAS, Gagné D, Sirard M-A, Khandjian ÉW, Richard FJ, Hyttel P, Robert C. Cumulus cell transcripts transit to the bovine oocyte in preparation for maturation1. Biol Reprod. 2016;94:1–11. https://doi.org/10.1095/biolreprod.114.127571.
54. Matos L, Stevenson D, Gomes F, Silva-carvalho JL, Almeida H. Superoxide dismutase expression in human cumulus oophorus cells. Mol Hum Reprod. 2009;15:411–9. https://doi.org/10.1093/molehr/gap034.
55. Mckenzie LJ, Pangas SA, Carson SA, Kovanci E, Cisneros P, Buster JE, Amato P, Matzuk MM. Human cumulus granulosa cell gene expression : a predictor of fertilization and embryo selection in women undergoing IVF. 2004;19:2869–2874.https://doi.org/10.1093/humrep/deh535.
56. Moor RM, Smith MW, Dawson RMC. Measurement of intercellular coupling between oocytes and cumulus cells using intracellular markers. Exp Cell Res. 1980;126:15–29. https://doi.org/10.1016/0014-4827(80)90466-8.
57. Papler TB, Bokal EV, Lovrecic L, Kopitar AN, Maver A. No specific gene expression signature in human granulosa and cumulus cells for prediction of oocyte fertilisation and embryo implantation. PLoS ONE. 2015;10:1–13. https://doi.org/10.1371/journal.pone.0115865.
58. Pfaffenseller B, da Silva MPV, De Bastiani MA, Castro MAA, Gallitano AL, Kapczinski F, Klamt F. Differential expression of transcriptional regulatory units in the prefrontal cortex of patients with bipolar disorder: potential role of early growth response gene 3. Transl Psychiatry. 2016. https://doi.org/10.1038/tp.2016.78.
59. R Core Team. R: a language and environment for statistical computing. R Found. Stat. Comput. Vienna, Austria. 2014.

60. Rienzi L, Balaban B, Ebner T, Mandelbaum J. The oocyte. Hum Reprod. 2012;27:2–21. https://doi.org/10.1093/humrep/des200.
61. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43:e47.
62. Salehi E, Aflatoonian R, Moeini A, Yamini N, Asadi E, Khosravizadeh Z, Tarzjani MD, Harat ZN, Abolhassani F. Apoptotic biomarkers in cumulus cells in relation to embryo quality in polycystic ovary syndrome. Arch Gynecol Obstet. 2017;296:1219–27. https://doi.org/10.1007/s00404-017-4523-5.
63. Sean D, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics. 2007;23:1846–7. https://doi.org/10.1093/bioinformatics/btm254.
64. Seino T. Eight-hydroxy-2′-deoxyguanosine in granulosa cells is correlated with the quality of oocytes and embryos in an in vitro fertilization-embryo transfer program. Fertil Steril. 2002;77:1184–90. https://doi.org/10.1016/S0015-0282(02)03103-5.
65. Shen Q, Chen M, Zhao X, Liu Y, Ren X, Zhang L. Versican expression level in cumulus cells is associated with human oocyte developmental competence. Syst Biol Reprod Med. 2020;66:176–84. https://doi.org/10.1080/19396368.2020.1725685.
66. Tabibnejad N, Sheikhha MH, Ghasemi N, Fesahat F, Soleimani M, Aflatoonian A. Association between early embryo morphokinetics plus cumulus cell gene expression and assisted reproduction outcomes in polycystic ovary syndrome women. Reprod Biomed Online. 2019;38:139–51. https://doi.org/10.1016/j.rbmo.2018.10.010.
67. Tanghe S, Van Soom A, Nauwynck H, Coryn M, De Kruif A. Minireview: Functions of the cumulus oophorus during oocyte maturation, ovulation, and fertilization. Mol Reprod Dev. 2002;61:414–24. https://doi.org/10.1002/mrd.10102.
68. Tatone C, Carbone MC, Falone S, Aimola P, Giardinelli A, Caserta D, Marci R, Pandolfi A, Ragnelli AM, Amicarelli F. Age-dependent changes in the expression of superoxide dismutases and catalase are associated with ultrastructural modifications in human granulosa cells. Mol Hum Reprod. 2006;12:655–60. https://doi.org/10.1093/molehr/gal080.
69. van Buuren S, Groothuis-Oudshoorn K. mice : multivariate imputation by chained equations in *R*. J Stat Softw. 2011. https://doi.org/10.18637/jss.v045.i03.
70. van Montfoort APA, Geraedts JPM, Dumoulin JCM, Stassen APM, Evers JLH, Ayoubi TAY. Differential gene expression in cumulus cells as a prognostic indicator of embryo viability: a microarray analysis. Mol Hum Reprod. 2008;14:157–68. https://doi.org/10.1093/molehr/gam088.
71. Vargas DM, De Bastiani MA, Zimmer ER, Klamt F. Alzheimer's disease master regulators analysis: search for potential molecular targets and drug repositioning candidates. Alzheimers Res Ther. 2018. https://doi.org/10.1186/s13195-018-0394-7.
72. von Mengden L, Klamt F, Smitz J. Redox biology of human cumulus cells: basic concepts, impact on oocyte quality, and potential clinical use. Antioxid Redox Signal. 2020;32:522–35. https://doi.org/10.1089/ars.2019.7984.
73. Wathlet S, Adriaenssens T, Segers I, Verheyen G, Janssens R, Coucke W, Devroey P, Smitz J. New candidate genes to predict pregnancy outcome in single embryo transfer cycles when using cumulus cell gene expression. Fertil Steril. 2012;98:432-439.e4. https://doi.org/10.1016/j.fertnstert.2012.05.007.
74. Wathlet S, Adriaenssens T, Segers I, Verheyen G, Van De VH, Coucke W, El RR, Devroey P, Smitz J. Cumulus cell gene expression predicts better cleavage-stage embryo or blastocyst development and pregnancy for ICSI patients. Hum Reprod. 2011;26:1035–51. https://doi.org/10.1093/humrep/der036.
75. Weymann D, Laskin J, Roscoe R, Schrader KA, Chia S, Yip S, Cheung WY, Gelmon KA, Karsan A, Renouf DJ, Marra M, Regier DA. The cost and cost trajectory of whole-genome analysis guiding treatment of patients with advanced cancers. Mol Genet Genomic Med. 2017;5:251–60. https://doi.org/10.1002/mgg3.281.
76. Zhang X, Jafari N, Barnes RB, Confino E, Milad M, Kazer RR. Studies of gene expression in human cumulus cells indicate pentraxin 3 as a possible marker for oocyte quality. Fertil Steril. 2005. https://doi.org/10.1016/j.fertnstert.2004.11.030.

Authors and Affiliations

Lucia von Mengden[1,2,3] · Marco Antônio De Bastiani[1,2,3] · Lucas Kich Grun[2,3,4,5] · Florencia Barbé-Tuana[3,4,5] · Tom Adriaenssens[6] · Johan Smitz[6] · Leticia Schmidt Arruda[7] · Carlos Alberto Link[7] · Fábio Klamt[1,2,3]

Marco Antônio De Bastiani
marco.bastiani@ufrgs.br

Lucas Kich Grun
lucaskgrun@gmail.com

Florencia Barbé-Tuana
florencia.tuana@pucrs.br

Tom Adriaenssens
tom.adriaenssens@uzbrussel.be

Johan Smitz
johan.smitz@uzbrussel.be

Leticia Schmidt Arruda
leticia@clinicaproser.com.br

Carlos Alberto Link
drlink@proser.com.br

Fábio Klamt
fabio.klamt@ufrgs.br

1 Laboratory of Cellular Biochemistry, Department of Biochemistry, ICBS/ Universidade Federal do Rio Grande do Sul (UFRGS), 90035-003 Rua Ramiro Barcelos 2600, Porto Alegre, RS 90035-003, Brazil

2 National Institutes of Science & Technology – Translational Medicine (INCT-TM), Porto Alegre, RS 90035-903, Brazil

3 Postgraduate Program: Biochemistry, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil

4 Laboratory of Molecular Biology and Bioinformatics, Department of Biochemistry, ICBS/UFRGS, Porto Alegre, RS 90035-003, Brazil

5 Pontifícia Universidade Católica Do Rio Grande Do Sul (PUCRS), Science School, Porto Alegre, RS 90619-900, Brazil

6 Laboratory of Follicle Biology, Vrije Universiteit Brussel, Brussels, Belgium

7 Clínica ProSer, Porto Alegre, RS 90420-160, Brazil