



DATA NOTE

# -A curated transcriptomic dataset collection relevant to embryonic development associated with *in vitro* fertilization in healthy individuals and patients with polycystic ovary syndrome [version 1; referees: 1 approved, 2 approved with reservations]

Rafah Mackeh <sup>1</sup>, Sabri Boughorbel<sup>2</sup>, Damien Chaussabel<sup>2</sup>, Tomoshige Kino<sup>1</sup>

<sup>1</sup>Department of Human Genetics, Division of Translational Medicine, Sidra Medical and Research Center, Doha, 26999, Qatar

<sup>2</sup>Department of Systems Biology, Division of Translational Medicine, Sidra Medical and Research Center, Doha, 26999, Qatar

**v1** First published: 23 Feb 2017, 6:181 (doi: [10.12688/f1000research.10877.1](https://doi.org/10.12688/f1000research.10877.1))  
 Latest published: 23 Feb 2017, 6:181 (doi: [10.12688/f1000research.10877.1](https://doi.org/10.12688/f1000research.10877.1))

**Abstract**

The collection of large-scale datasets available in public repositories is rapidly growing and providing opportunities to identify and fill gaps in different fields of biomedical research. However, users of these datasets should be able to selectively browse datasets related to their field of interest. Here we made available a collection of transcriptome datasets related to human follicular cells from normal individuals or patients with polycystic ovary syndrome, in the process of their development, during *in vitro* fertilization. After RNA-seq dataset exclusion and careful selection based on study description and sample information, 12 datasets, encompassing a total of 85 unique transcriptome profiles, were identified in NCBI Gene Expression Omnibus and uploaded to the Gene Expression Browser (GXB), a web application specifically designed for interactive query and visualization of integrated large-scale data. Once annotated in GXB, multiple sample grouping has been made in order to create rank lists to allow easy data interpretation and comparison. The GXB tool also allows the users to browse a single gene across multiple projects to evaluate its expression profiles in multiple biological systems/conditions in a web-based customized graphical views. The curated dataset is accessible at the following link: <http://ivf.gxbsidra.org/dm3/landing.gsp>.

**Open Peer Review**

Referee Status:

	Invited Referees		
	1	2	3
<b>version 1</b> published 23 Feb 2017	 report	 report	 report

- 1 **Rawad Hodeify**, Weill Cornell Medical College in Qatar (WCM-Q) Qatar
- 2 **François J. Richard**, Université Laval Canada
- 3 **Rita Singh**, University of Delhi India

**Discuss this article**

Comments (0)



This article is included in the **Molecular Profiling of Pregnancy** channel.



This article is included in the **Collective Data Access** channel.



This article is included in the **Sidra Research** channel.

**Corresponding author:** Rafah Mackeh ([rmackeh@sidra.org](mailto:rmackeh@sidra.org))

**How to cite this article:** Mackeh R, Boughorbel S, Chaussabel D and Kino T. -**A curated transcriptomic dataset collection relevant to embryonic development associated with *in vitro* fertilization in healthy individuals and patients with polycystic ovary syndrome [version 1; referees: 1 approved, 2 approved with reservations]** *F1000Research* 2017, 6:181 (doi: [10.12688/f1000research.10877.1](https://doi.org/10.12688/f1000research.10877.1))

**Copyright:** © 2017 Mackeh R *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

**Grant information:** This study was supported by the Intramural Grant of the Sidra Medical and Research Center.

**Competing interests:** No competing interests were disclosed.

**First published:** 23 Feb 2017, 6:181 (doi: [10.12688/f1000research.10877.1](https://doi.org/10.12688/f1000research.10877.1))

## Introduction

Oocytes are maternal germ cells developed in ovaries during the fetal phase and kept throughout the female reproductive ages for monthly maturation and subsequent ovulation following the endocrinological regulation associated with menstrual cycles<sup>1</sup>. Oocyte maturation starts with the monthly resumption of the first meiotic process of one primary oocyte arrested in prophase I (characterized by the germinal vesicle, also classified as immature or metaphase I (MI) stage)<sup>1</sup>. After extrusion of the first polar body, the primary oocyte progresses to metaphase II of the second meiosis and becomes the secondary oocyte, which is competent to fertilization by a sperm.

Such oocyte growth/maturation occurs inside the ovarian follicle, which is also concomitantly under a process called folliculogenesis. Folliculogenesis consists of follicular cell proliferation, development and differentiation<sup>1</sup>. Primordial follicles containing primary oocytes grow into the mature Graafian follicle with the coordinated progression of the holding germ cells to the secondary oocytes<sup>2</sup>. Ovulation then occurs under the regulation of gonadotropins and sex steroids, resulting in the release of an oocyte into the peritoneal cavity. Upon fertilization by a sperm, the liberated oocyte resumes its second meiotic division to become the zygote, which further goes into a form of embryo called morula through several mitotic divisions and compaction of component cells. Continuous cell division further transforms morula to blastocyst, which has a fluid-filled cavity and is ready for implanting to the uterine endometrium<sup>3</sup>.

The oocyte in the ovarian follicle is a primary regulator of follicular cell differentiation and function, whereas metabolic cooperation occurs between oocytes and follicular cells to ensure substrate supply necessary for oocyte growth/maturation<sup>4</sup>. The follicular cells consist of two types of cell groups, theca cells (also known as stromal cells) and granulosa cells. Theca cells form the outer layer of the ovarian follicle, while inner granulosa cells make a direct contact with the oocyte. These cells also produce steroid hormones, such as progestins and estrogens, under the control of pituitary gonadotropins, which is important for priming uterine endometrium and other reproductive tissues for supporting expected implantation and pregnancy<sup>5</sup>. During folliculogenesis, granulosa cells continuously proliferate to form the follicular antrum, a fluid-filled cavity formed among the granulosa cell cluster. Upon formation of the antrum, two populations of granulosa cells become identifiable: one cell group known as **cumulus cells** (CCs), which surround the oocyte and remain associated with it even after ovulation, and the other group called **mural granulosa cells**, which form an inner layer of the follicle. The oocyte and CCs form the cumulus-oocyte-complex in which these cells directly communicate with each other through the gap-junctions created between them. This cellular communication plays a central role in the regulation of folliculogenesis and oocyte maturation by enabling the nutritional transfer and traffic of macromolecules between them<sup>6</sup>.

*In vitro* fertilization (IVF) is one type of assisted reproductive technology developed for the treatment of infertility<sup>7</sup>. It is a procedure consisting of (1) harvesting oocytes from the peritoneal cavity of the women artificially stimulated for their ovulation, (2) fertilization of the oocytes by mixing with sperms *in vitro*, and (3)

implantation of fertilized oocytes into the uterine cavity. Before implantation, fertilized oocytes are regularly cultured for 2–6 days in a growth medium allowing its cell division and multiplication. Although a lot of improvements have been added to IVF, its success rate for successful live birth is still less than 50% even in younger women, and the main challenge remains the risk of multiple pregnancies, which is directly associated with increased incidence of fetal morbidity and infant mortality during maternal, perinatal and neonatal periods<sup>8</sup>. To prevent multiple IVF-associated pregnancies, single-embryo transfer is considered, for which selection of the most viable and healthy embryo is critical. Morphological inspection of embryos is employed for selecting high quality embryos<sup>9,10</sup>, but it is not sufficient to predict the developmental potential of embryos. Therefore, studies have been performed during the last several years to develop better methods of embryo selection by examining proteomics or metabolomics of embryos<sup>11–13</sup>. Recently, emergence of microarray technology has introduced a new approach to study the genetic aspects of fertility. Primarily, studies employing this new technique focused on the role surrounding follicular cells for evaluating the quality of carrying oocytes, and estimated its usefulness by comparing and correlating the data from stromal cells with the quality of embryos and with a positive or negative IVF outcome<sup>14–19</sup>. Such studies also included samples obtained from healthy or diseased women, for example women with polycystic ovary syndrome (PCOS), for whom the IVF success rate is known to be reduced compared with healthy subjects<sup>20</sup>.

To help identify knowledge gaps in the field of IVF, ovarian function and/or the influence of reproductive diseases, we provide here a resource enabling mainstream researchers in this field to browse transcriptomic datasets relevant to the oocyte and surrounding stromal cells obtained from healthy subjects or those with PCOS, in association with IVF outcome. Such a resource offers a unique opportunity to identify the genes that play key roles in oocyte maturation, embryonic development and crosstalk between oocytes and granulosa cells, eventually contributing to the future improvement of the IVF procedure.

## Methods

In order to identify datasets relevant to IVF, we developed queries in a way to include the conditions, such as oocytes, CCs or granulosa cells in humans. Queries were employed on NCBI (<https://www.ncbi.nlm.nih.gov/>) and are as follows:

- *Homo sapiens*[organism] AND (oocyte OR oocytes) AND (“Expression profiling by array”[gdsType] OR “Expression profiling by high throughput sequencing”[gdsType]).
- *Homo sapiens*[organism] AND cumulus cells AND (“Expression profiling by array”[gdsType] OR “Expression profiling by high throughput sequencing”[gdsType]).
- *Homo sapiens*[organism] AND Granulosa cells AND (“Expression profiling by array”[gdsType] OR “Expression profiling by high throughput sequencing”[gdsType]).
- *Homo sapiens*[organism] AND (in vitro fertilization OR in vitro fertilization OR in vitro fecundation) AND (“Expression profiling by array”[gdsType] OR “Expression profiling by high throughput sequencing”[gdsType]).

This query retrieved 85 datasets. After excluding RNA-seq datasets from the collection and examining each dataset carefully based on study description and list of samples and their annotations to verify their direct relevance to the theme of this data compendium, a total number of 23 datasets were selected. In total, 12 were successfully uploaded into the data browser. Details of these datasets are recapitulated in [Table 1](#).

After curation, each dataset was downloaded from the Gene Expression Omnibus of the National Center for Biotechnology Information website (NCBI GEO) using the SOFT file format, and was then uploaded, along with its study information and samples available, to the Gene Expression Browser, version 1.2 (GXB; <http://ivf.gxbsidra.org/dm3/geneBrowser/list>), an interactive web-based application developed at the Benaroya Research Institute (Seattle, WA, USA), hosted on the Amazon Web Services cloud (<https://github.com/BenaroyaResearch/gxbrowser>) (<https://aws.amazon.com>)<sup>21</sup>. In GXB, we grouped the samples according to the expected future interpretation and comparison of study results. Each group contains samples of biological replicates, such as samples from control patients, and is compared to another

group of samples. For example, Control group vs PCOS group, or Blastocysts group vs embryos of poor quality. Finally, computed ranking lists were created based on each grouping, using the rank list option provided in the GXB software. Therefore, GXB provides the users with a means to easily navigate and filter our uploaded and processed dataset collections, which are available at <http://ivf.gxbsidra.org/dm3/landing.gsp>.

A web tutorial for GXB is available online: <http://ivf.gxbsidra.org/dm3/tutorials.gsp#gxbtut> and is briefly reproduced here so that readers can use this article as a standalone resource<sup>21,22</sup>: “datasets of interest can be quickly identified either by filtering criteria from pre-defined lists shown on the left side of the GXB dataset navigation window, or by entering a query term in the search box located at its left top portion. Clicking on one of the studies listed in the dataset navigation window opens a viewer, which is designed to provide interactive browsing and graphic representations of the large-scale data in an interpretable format. This interface is intended to navigate ranked gene lists and displays transcriptomic results graphically in a context-rich environment. Selecting a gene from the rank-ordered list on the left side of the data-viewing window

**Table 1. Datasets\* included in our collection.**

GEO ID	Title	Platform	Number of samples	Genes expression used for dataset validation	Reference
GSE34526	Differential gene expression in granulosa cells from polycystic ovary syndrome patients with and without insulin resistance: Identification of susceptibility gene sets through network analysis	Affymetrix	10	XIST FIGLA	27
GSE37277	Differentiating factors of cumulus cells related to quality of the human oocyte	Agilent 014850 v1	92	XIST	16
GSE37110	Differentiating factors of cumulus cells related to quality of the human oocyte	Agilent 014850 v1	32	XIST	16
GSE37117	Differentiating factors of cumulus cells related to quality of the human oocyte	Agilent 014850 v1	36	XIST	16
GSE37116	Differentiating factors of cumulus cells related to quality of the human oocyte	Agilent 014850 v1	24	XIST	16
GSE9526	Expression data from cumulus cells that surround oocytes resulting in early or late cleaving embryos	Affymetrix	16	XIST	17
GSE40400	Expression data from human cumulus cells isolated from oocytes at MI and MII staged in polycystic ovary syndrome (PCOS) patients	Affymetrix	8	XIST FIGLA	14
GSE10946	Gene expression microarray profiles of cumulus cells in lean and overweight-obese polycystic ovary syndrome patients	Affymetrix	23	XIST	15
GSE31681	Human cumulus cells	Affymetrix	24	XIST	28
GSE5850	Microarray analysis on NL and PCOS oocytes	Affymetrix	12	XIST	18
GSE43684	Modified natural and stimulated <i>in vitro</i> fertilization cycle: Cumulus cells	Affymetrix	8	FIGLA	19
GSE12034	The transcriptome of human oocytes	Affymetrix	6	FIGLA, BMP15, XIST, Zp1, ZP2, ZP3	29

\*: available at <http://ivf.gxbsidra.org/dm3/geneBrowser/list>.

displays its expression values graphically. The drop-down menus directly above the graphical display give the users the following options: a) Change how the gene list is ranked, which allows the user to change the method used to rank the genes, or to include only the genes that are selected based on his/her specific biological interest; b) Change sample grouping (Group Set button), so that in some datasets, a user can switch between groups, based on, for example, the cell types and the diseases of interest; c) Sort individual samples within a group based on the associated categorical or continuous variables (e.g., gender and age); d) Toggle between the histogram and a box-plot plot with expression values, which are demonstrated as a single point for each sample in the graph; e) Paste color legends for sample groups; f) Select categorical information that is to be overlaid at the bottom of the graph. For example, the user can display gender or smoking status using this function; g) Provide a color legend for the categorical information overlaid at the bottom of the graph; h) Download the graph in a jpeg format. Generally, raw data of the measurements *per se* shown in graphs have no intrinsic utility in the absence of their contextual information. It is therefore important to display such information together with the data shown in the graphs, so that viewers are able to interpret demonstrated data and gain new insights from it. In the datasets provided, the contextual information has been organized under different tabs directly above the graphical display. The tabs can be hidden to make more room for displaying the data plots, or revealed by clicking on the blue “Show Info Panel” button in the top right corner of the display window. Information for the gene, which is selected from the list and is shown in the left side of the display, is available under the “Gene” tab. The study information is also available under the “Study” tab. Further, information on individual samples is provided under the “Sample” tab. Rolling the mouse cursor over a histogram bar while displaying the “Sample” tab enables viewing of any clinical, demographic, or laboratory information provided for the selected sample. Finally, the “Downloads” tab allows advanced users to retrieve the original datasets for their future analysis to be performed outside GXB. It also provides all available sample annotation data together with the expression data.”

### Dataset validation

Quality checks for the datasets uploaded to GXB were performed by validating the specific expression of the Xist transcript (X-inactive specific transcript), which is a non-protein-coding RNA that inactivates one of the diploid X chromosomes existing in the female cells of mammals<sup>23,24</sup>. Since all uploaded datasets comprised

samples obtained from women, Xist was expected to be present and expressed at high levels in all samples, except one dataset which comprises oocyte transcriptomic data, as haploid oocytes do not bear chromosome X inactivation. Expectedly, when microarrays provided probes for Xist, its expression was present in all datasets comprising cumulus or granulosa cells. While Xist expression was absent in oocyte samples of the GSE12034, it was highly expressed in the non-ovarian diploid tissue samples of the same dataset. Additional validation of our datasets was performed by examining the expression of some ovarian-specific genes, such as those specific to the zona pellucida protein (ZP1, ZP2 and ZP3), *FIGLA* (folliculogenesis-specific basic helix-loop-helix gene, also known as factor in the germline  $\alpha$ ), which encodes a transcription factor regulating the expression of multiple oocyte-specific genes<sup>25</sup>, and *BMP15* (bone morphogenetic protein 15), which is functional in the folliculogenesis<sup>26</sup>. *FIGLA* was selectively expressed in oocyte samples in the GSE12034 dataset, but not in non-ovarian control tissues. The same expression pattern was also confirmed for *ZP1*, *ZP2*, *ZP3*, and *BMP15*.

### Data availability

All datasets were cited in our manuscript. They are designated by their GEO accession numbers (e.g. GSE34526), and can also be accessed using this identifier via the NCBI GEO website (<https://www.ncbi.nlm.nih.gov/gds/?term=>). User can download all uploaded dataset files and associated sample information through the GXB tool: “Downloads” tab.

### Author contributions

RM and TK conceived the theme of this dataset collection, SB contributed to the loading and curation of datasets. RM prepared the first draft of the manuscript, TK and DC edited it.

### Competing interests

No competing interests were disclosed.

### Grant information

This study was supported by the Intramural Grant of the Sidra Medical and Research Center.

### Acknowledgements

We thank all the investigators who deposited their datasets used in this study to NCBI GEO.

## References

- Hutt KJ, Albertini DF: **An oocentric view of folliculogenesis and embryogenesis.** *Reprod Biomed Online.* 2007; **14**(6): 758–64.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- van den Hurk R, Zhao J: **Formation of mammalian oocytes and their growth, differentiation and maturation within ovarian follicles.** *Theriogenology.* 2005; **63**(6): 1717–51.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Vaillancourt C, Lafond J: **Human embryogenesis: overview.** *Methods Mol Biol.* 2009; **550**: 3–7.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Li R, Albertini DF: **The road to maturation: somatic cell interaction and self-organization of the mammalian oocyte.** *Nat Rev Mol Cell Biol.* 2013; **14**(3): 141–52.  
[PubMed Abstract](#) | [Publisher Full Text](#)

5. Rider V, Kimler BF, Justice WM: **Progesterone-growth factor interactions in uterine stromal cells.** *Biol Reprod.* 1998; **59**(3): 464–9.  
[PubMed Abstract](#) | [Publisher Full Text](#)
6. Rossello RA, H D: **Cell communication and tissue engineering.** *Commun Integr Biol.* 2010; **3**(1): 53–6.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
7. Wood C, Trounson A: **Clinical *in vitro* fertilization.** Second edition, Springer-Verlag, 1989.  
[Publisher Full Text](#)
8. Land JA, Evers JL: **Risks and complications in assisted reproduction techniques: Report of an ESHRE consensus meeting.** *Hum Reprod.* 2003; **18**(2): 455–7.  
[PubMed Abstract](#) | [Publisher Full Text](#)
9. Borini A, Lagalla C, Cattoli M, *et al.*: **Predictive factors for embryo implantation potential.** *Reprod Biomed Online.* 2005; **10**(5): 653–68.  
[PubMed Abstract](#) | [Publisher Full Text](#)
10. Coticchio G, Bonu MA, Bianchi V, *et al.*: **Criteria to assess human oocyte quality after cryopreservation.** *Reprod Biomed Online.* 2005; **11**(4): 421–7.  
[PubMed Abstract](#) | [Publisher Full Text](#)
11. Katz-Jaffe MG, Gardner DK, Schoolcraft WB: **Proteomic analysis of individual human embryos to identify novel biomarkers of development and viability.** *Fertil Steril.* 2006; **85**(1): 101–7.  
[PubMed Abstract](#) | [Publisher Full Text](#)
12. Katz-Jaffe MG, McReynolds S, Gardner DK, *et al.*: **The role of proteomics in defining the human embryonic secretome.** *Mol Hum Reprod.* 2009; **15**(5): 271–7.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. Singh R, Sinclair KD: **Metabolomics: approaches to assessing oocyte and embryo quality.** *Theriogenology.* 2007; **68**(Suppl 1): S56–62.  
[PubMed Abstract](#) | [Publisher Full Text](#)
14. Huang X, Hao C, Shen X, *et al.*: **Differences in the transcriptional profiles of human cumulus cells isolated from MI and MII oocytes of patients with polycystic ovary syndrome.** *Reproduction.* 2013; **145**(6): 597–608.  
[PubMed Abstract](#) | [Publisher Full Text](#)
15. Kenigsberg S, Bentov Y, Chalifa-Caspi V, *et al.*: **Gene expression microarray profiles of cumulus cells in lean and overweight-obese polycystic ovary syndrome patients.** *Mol Hum Reprod.* 2009; **15**(2): 89–103.  
[PubMed Abstract](#) | [Publisher Full Text](#)
16. Feuerstein P, Puard V, Chevalier C, *et al.*: **Genomic assessment of human cumulus cell marker genes as predictors of oocyte developmental competence: impact of various experimental factors.** *PLoS One.* 2012; **7**(7): e40449.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
17. van Montfoort AP, Geraedts JP, Dumoulin JC, *et al.*: **Differential gene expression in cumulus cells as a prognostic indicator of embryo viability: a microarray analysis.** *Mol Hum Reprod.* 2008; **14**(3): 157–68.  
[PubMed Abstract](#) | [Publisher Full Text](#)
18. Wood JR, Dumesic DA, Abbott DH, *et al.*: **Molecular abnormalities in oocytes from women with polycystic ovary syndrome revealed by microarray analysis.** *J Clin Endocrinol Metab.* 2007; **92**(2): 705–13.  
[PubMed Abstract](#) | [Publisher Full Text](#)
19. Papler TB, Bokal EV, Tacer KF, *et al.*: **Differences in cumulus cells gene expression between modified natural and stimulated *in vitro* fertilization cycles.** *J Assist Reprod Genet.* 2014; **31**(1): 79–88.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
20. Palomba S, Daolio J, La Sala GB: **Oocyte Competence in Women with Polycystic Ovary Syndrome.** *Trends Endocrinol Metab.* 2016; pii: S1043-2760(16)30168-0.  
[PubMed Abstract](#) | [Publisher Full Text](#)
21. Speake C, Presnell S, Domico K, *et al.*: **An interactive web application for the dissemination of human systems immunology data.** *J Transl Med.* 2015; **13**: 196.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
22. Rinchai D, Boughorbel S, Presnell S, *et al.*: **A curated compendium of monocyte transcriptome datasets of relevance to human monocyte immunobiology research [version 2; referees: 2 approved].** *F1000Res.* 2016; **5**: 291.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
23. Hendrich BD, Plenge RM, Willard HF: **Identification and characterization of the human *XIST* gene promoter: implications for models of X chromosome inactivation.** *Nucleic Acids Res.* 1997; **25**(13): 2661–71.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
24. Brown CJ, Ballabio A, Rupert JL, *et al.*: **A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome.** *Nature.* 1991; **349**(6304): 38–44.  
[PubMed Abstract](#) | [Publisher Full Text](#)
25. Huntriss J, Gosden R, Hinkins M, *et al.*: **Isolation, characterization and expression of the human Factor In the Germline alpha (*FIGLA*) gene in ovarian follicles and oocytes.** *Mol Hum Reprod.* 2002; **8**(12): 1087–95.  
[PubMed Abstract](#) | [Publisher Full Text](#)
26. Persani L, Rossetti R, Di Pasquale E, *et al.*: **The fundamental role of *bone morphogenetic protein 15* in ovarian function and its involvement in female fertility disorders.** *Hum Reprod Update.* 2014; **20**(6): 869–83.  
[PubMed Abstract](#) | [Publisher Full Text](#)
27. Kaur S, Archer KJ, Devi MG, *et al.*: **Differential gene expression in granulosa cells from polycystic ovary syndrome patients with and without insulin resistance: identification of susceptibility gene sets through network analysis.** *J Clin Endocrinol Metab.* 2012; **97**(10): E2016–21.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
28. Ouandaogo ZG, Haouzi D, Assou S, *et al.*: **Human cumulus cells molecular signature in relation to oocyte nuclear maturity stage.** *PLoS One.* 2011; **6**(11): e27179.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
29. Gonzalez-Muñoz E, Arboleda-Estudillo Y, Otu HH, *et al.*: **Cell reprogramming. Histone chaperone *ASF1A* is required for maintenance of pluripotency and cellular reprogramming.** *Science.* 2014; **345**(6198): 822–5.  
[PubMed Abstract](#) | [Publisher Full Text](#)

# Open Peer Review

Current Referee Status:



Version 1

Referee Report 20 March 2017

doi:10.5256/f1000research.11728.r20492



**Rita Singh**

Lab of Molecular Reproduction, Department of Zoology, University of Delhi, Delhi, India

The manuscript by Mackeh et al. is a collection of the gene expression datasets of oocyte, cumulus cells, and granulosa cells of normal and PCOS patients undergoing IVF. It is a good compilation of related datasets already published in public repositories.

However, there are some concerns that are not addressed here.

How were the expression values, as shown in the graphs, obtained from the raw data files? The details of the methodology used to analyse the raw data and to generate the ranked gene lists should be given. In the present form, it is difficult to make use of the data for any meaningful scientific analysis.

The purpose of this study is to browse a single gene across multiple projects to evaluate its expression profiles in multiple biological systems/conditions in web-based customized graphical views. However, the gene expression data is shown as expression values for some datasets and as  $\text{Log}_2$  expression values for the others.

There is a typographical mistake in the spellings of granulosa cells.

The Pubmed articles linked to the data sets are not available.

Although putting together these data is helpful for the analysis of transcriptome data from normal and PCOS patients undergoing IVF, it would be meaningful but not mandatory to include the data available from similar platforms for theca cells.

It is a good effort done by the authors to put together several studies.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

**Competing Interests:** No competing interests were disclosed.

Referee Report 17 March 2017

doi:10.5256/f1000research.11728.r20579



**François J. Richard**

Département des Sciences Animales, Université Laval, Ville de Québec, QC, Canada

Although this manuscript is showing a web platform to look at genes involved in PCOS patients in oocyte, cumulus cells and granulosa cells, some concerns should also be considered and is not addressed here.

1. Bias can be obtained because of the method of RNA isolation, purification and RNA amplification that may be different between published papers.
2. The datasets were validated using cellular specific gene expression but nothing is mentioned about cellular contamination, reference genes (housekeeping genes), ...

Because the data are classified by study using raw values for specific gene for each sample, it becomes highly difficult to grasp meaningful information. It would have been helpful to further analyze the data and not only showing the raw data of each sample.

In conclusion, it seems to be a good platform to run a quick analysis looking at several studies.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

**Competing Interests:** No competing interests were disclosed.

Referee Report 01 March 2017

doi:10.5256/f1000research.11728.r20488



**Rawad Hodeify**

Department of Physiology and Biophysics, Weill Cornell Medical College in Qatar (WCM-Q), Doha, Qatar

The manuscript by Mackeh et al. presents a very interesting and novel approach to identify genes that are potentially linked to embryonic development. The authors introducing a valuable resource collecting gene expression profiling datasets from oocytes and surrounding stromal cells of healthy subjects or those with polycystic ovary syndrome, in correlation with IVF outcome. This resource is quite beneficial by providing a catalogue of genes that show altered expression in negative IVF outcome.

The transcriptomic datasets are presented in an easy-to-use interactive web application that enables users, including those who are not experts in gene expression profiling, to identify altered gene expression in oocytes and associated cells in normal and diseased situations. Overall, I would give the manuscript in its current form a high priority to be indexed. I have minor comment and suggestion.

1. In page 3, the first paragraph (line 10) of the introduction, the secondary oocyte is also commonly known as 'egg'. I would suggest that both terms are mentioned.
2. It will be interesting for the authors to check whether adding the term (Egg) in the queries will yield extra datasets.



I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

***Competing Interests:*** No competing interests were disclosed.

---