

6-21-2010

## Proteomics-based systems biology modeling of bovine germinal vesicle stage oocyte and cumulus cell interaction.

Divyaswetha Peddinti

Erdogan Memili

Shane C. Burgess

Follow this and additional works at: <https://scholarsjunction.msstate.edu/cvm-publications>

---

### Recommended Citation

Peddinti, Divyaswetha; Memili, Erdogan; and Burgess, Shane C., "Proteomics-based systems biology modeling of bovine germinal vesicle stage oocyte and cumulus cell interaction." (2010). *CVM Publications*. 7.

<https://scholarsjunction.msstate.edu/cvm-publications/7>

This Article is brought to you for free and open access by the College of Veterinary Medicine at Scholars Junction. It has been accepted for inclusion in CVM Publications by an authorized administrator of Scholars Junction. For more information, please contact [scholcomm@msstate.libanswers.com](mailto:scholcomm@msstate.libanswers.com).

# Proteomics-Based Systems Biology Modeling of Bovine Germinal Vesicle Stage Oocyte and Cumulus Cell Interaction

Divyaswetha Peddinti<sup>1,3\*</sup>, Erdogan Memili<sup>2,3</sup>, Shane C. Burgess<sup>1,3,4,5</sup>

**1** Department of Basic Sciences, College of Veterinary Medicine, Mississippi State University, Mississippi State, Mississippi, United States of America, **2** Department of Animal and Dairy Sciences, Mississippi State University, Mississippi State, Mississippi, United States of America, **3** Institute for Digital Biology, Mississippi State University, Mississippi State, Mississippi, United States of America, **4** Life Sciences and Biotechnology Institute, Mississippi State University, Mississippi State, Mississippi, United States of America, **5** Mississippi Agricultural and Forestry Experimental Station, Mississippi State University, Mississippi State, Mississippi, United States of America

## Abstract

**Background:** Oocytes are the female gametes which establish the program of life after fertilization. Interactions between oocyte and the surrounding cumulus cells at germinal vesicle (GV) stage are considered essential for proper maturation or 'programming' of oocytes, which is crucial for normal fertilization and embryonic development. However, despite its importance, little is known about the molecular events and pathways involved in this bidirectional communication.

**Methodology/Principal Findings:** We used differential detergent fractionation multidimensional protein identification technology (DDF-Mud PIT) on bovine GV oocyte and cumulus cells and identified 811 and 1247 proteins in GV oocyte and cumulus cells, respectively; 371 proteins were significantly differentially expressed between each cell type. Systems biology modeling, which included Gene Ontology (GO) and canonical genetic pathway analysis, showed that cumulus cells have higher expression of proteins involved in cell communication, generation of precursor metabolites and energy, as well as transport than GV oocytes. Our data also suggests a hypothesis that oocytes may depend on the presence of cumulus cells to generate specific cellular signals to coordinate their growth and maturation.

**Conclusions/Significance:** Systems biology modeling of bovine oocytes and cumulus cells in the context of GO and protein interaction networks identified the signaling pathways associated with the proteins involved in cell-to-cell signaling biological process that may have implications in oocyte competence and maturation. This first comprehensive systems biology modeling of bovine oocytes and cumulus cell proteomes not only provides a foundation for signaling and cell physiology at the GV stage of oocyte development, but are also valuable for comparative studies of other stages of oocyte development at the molecular level.

**Citation:** Peddinti D, Memili E, Burgess SC (2010) Proteomics-Based Systems Biology Modeling of Bovine Germinal Vesicle Stage Oocyte and Cumulus Cell Interaction. PLoS ONE 5(6): e11240. doi:10.1371/journal.pone.0011240

**Editor:** Jay M. Baltz, Ottawa Hospital Research Institute and University of Ottawa, Canada

**Received:** January 27, 2010; **Accepted:** May 28, 2010; **Published:** June 21, 2010

**Copyright:** © 2010 Peddinti et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This study was funded by the Life Sciences and Biology Institute and Mississippi Agricultural and Forestry Experiment Station, Mississippi State University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: dpeddinti@cvm.msstate.edu

## Introduction

Germinal vesicle (GV) breakdown is fundamental for maturation of fully grown, developmentally competent mammalian oocytes. Intercellular communication between oocytes and cumulus cells at GV stage is essential for proper maturation or 'programming' of oocytes, which is crucial for fertilization and embryonic development [1,2]. 'Gap junctions' in the regions of oocyte and cumulus cells association allow nutrient and paracrine factor transport between oocytes and cumulus cells [2,3,4]. Cumulus cell removal before maturation, or the obstruction of gap junctions, suppresses oocyte maturation [5,6,7,8,9]. Furthermore, cumulus cells are proposed to protect oocytes by preventing oxidative stress-induced cell death and DNA damage by increasing oocyte glutathione content [10] and thus functionally influence oocyte competence. In turn, via secreted factors, oocytes regulate folliculogenesis by promoting: granulosa cell proliferation, differ-

entiation, and gene expression as well as cumulus cell expansion [2,11]. Folliculogenesis fails in the absence of oocyte paracrine signaling, (whether due to genetic deficiency or experimental oocyte ablation) [2,12,13]. Although this oocyte and cumulus cell bidirectional communication is essential for competent oocyte development, the molecular details underlying this communication remain poorly defined. There is thus still a lack of reliable molecular markers and valid definition of a high quality oocyte have impeded the selection of optimal oocytes necessary for assisted reproductive techniques (ARTs) at a high efficiency in humans as well as farm animal species.

Published studies with mouse model show that cumulus cells play an important role in nutritional support of the developing oocyte in the form of pyruvate [14,15,16] and stimulation of this nutritional support of cumulus cells is in turn dependent upon the presence of paracrine factors secreted by the oocytes [17]. Although most basic reproductive biology work is done in the

mouse [18], significant species differences in oocyte biology exist between humans and mice [19,20]. The bovine is a relevant animal model for studies of oocyte and cumulus cell communication in human because oocyte biology, and many aspects of ovarian follicular dynamics, is similar between these two single ovulating species [14,15]. Bovine fertility is also important on its own merit; it has implications in agro-economics involving cattle industry worldwide. Evidences using both the bovine and porcine models show that attachment of cumulus cells to the oocyte during meiotic maturation and fertilization is critical for promoting subsequent embryo development [7,8,21,22].

Proteins primarily determine cell phenotypes and here we used a shotgun proteomics approach that allows us to relatively quantify which proteins are actually expressed in the cell compartments (as opposed to what might be or have the potential to be). This is especially important in oocytes, where there is no linear correlation between amounts of mRNA and the proteins they encode [23]. We previously analyzed the proteomes of bovine germinal vesicle (GV) stage oocytes and their surrounding cumulus cells using differential detergent fractionation two-dimensional liquid chromatography electrospray ionization tandem mass spectrometry (DDF 2D-LC ESI MS/MS) [19]. In the previous study, we reported the first descriptive map of bovine GV oocytes and their potentially-interacting cumulus cells with specific emphasis on membrane, nuclear proteins, receptor-ligand pairs, and transcription factors. Here, and in contrast, using separately-harvested bovine GV oocyte and cumulus cells, we did a separate proteomics experiment but this time using an updated bovine proteome, more stringent search criteria, and much better structural and functional annotation. This has allowed us to do more comprehensive quantitative computational systems biology modeling of oocyte and cumulus cell communication.

Biological systems utilize highly complex, interrelated networks and pathways to function and in contrast to our previous work [19], here we used two complementary computational systems biology modeling approaches: Gene Ontology (GO; [24]) -based and canonical genetic-network-based, to derive understanding from our large dataset. As part of this process, we functionally-annotated all oocyte and cumulus proteins we identified. This improved these protein's GO annotation quality score (GAQ) [25], and enhanced our ability to do GO quantitative modeling to identify the biological processes that are either agonistic or antagonistic to GV oocyte development and which may affect oocyte competence and maturation. Our data also provide a foundation for further structural and functional annotation of the bovine genome specifically-focused on identifying genes associated with developmental competency that could be used for selecting oocytes in manipulating mammalian reproduction. Complementing our biological process modeling, our canonical genetic network modeling identified signaling pathways likely to be involved in the bidirectional communication between oocytes and cumulus cells. Together our analyses can serve as a basis for further "omic" or reductionist research on oocyte and cumulus cell communication and related reproductive abnormalities.

## Results

### GV oocyte and cumulus cell proteomes

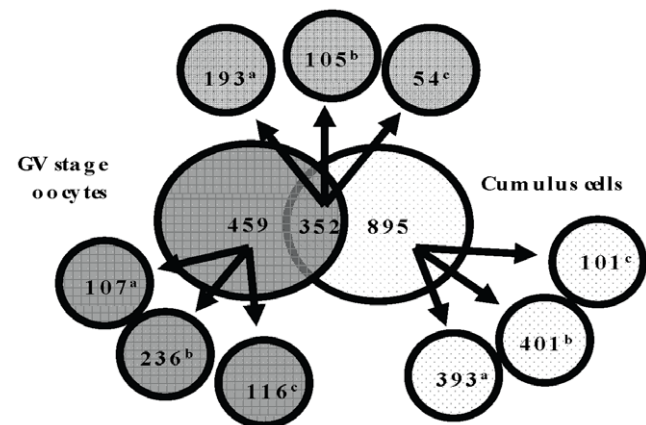
We identified 811 and 1,247 proteins in GV oocyte and cumulus cells, respectively (Table S1). All protein identifications and MS/MS data have been submitted to the PRoteomics IDentifications database (PRIDE [26]; accession #: 8691, 8692, 8693, 8694, 8695 and 8696, representing the 6 new datasets generated in this work). Of the total 2,058 proteins, 352 (20.4%)

were common to both cell types and 459, and 895 were unique to GV oocyte, and cumulus cells, respectively (Figure 1); 371 proteins were differentially-expressed between GV oocyte and cumulus cells (301 proteins had higher and 70 had lower expression in the cumulus cells, compared to the GV oocytes [Table S2]). Only 702 proteins (41.1%) were annotated as 'known' and their expression at protein level have been experimentally validated (23% cumulus cell specific, 6.3% GV oocyte specific, and 11.3% common to both [Figure 1]). Approximately 43% of the identified proteins were annotated as 'predicted' based on sequence homology and their expression at protein level has not previously been confirmed [27]. Our proteomics data has contributed to the bovine genome annotation by experimentally-confirming the *in vivo* expression of 742 electronically predicted proteins (Table S1). We also identified 6.5% and 5.7% of 'hypothetical' proteins (i.e. proteins predicted from nucleic acid sequences and that have not been shown to exist by experimental protein chemical evidence [28]) specific to GV oocyte and cumulus cells, respectively and 3.2% common to both cell types.

It is complicated to directly compare between the data set in present study with our previous study [19]. First this is because here we used more sophisticated proteomics methods. Secondly we used a bovine RefSeq database [08/28/2007; 25,078 entries] with very many different protein identification numbers (GI numbers) and protein names; previously *Bos taurus*, had no RefSeq database and so we had to derive a non-redundant protein database [NRPD; National Center for Biotechnology Institute. (NCBI); 07/20/2005; 39,963 entries] using search terms 'Bos' and 'taurus'. Together this means that any estimate of overlap based on GI or name will be an underestimate. Here we BLAST searched the protein sequences of the smaller current dataset against the larger previous dataset and found an overlap of 439 (26%) (Figure S1).

### GO Modeling

Predicted and hypothetical proteins do not have any functional annotation associated with them and they represent ~60% of total proteins in our oocyte and cumulus datasets. To compensate for this lack of functional annotation, and to provide the best foundation for biological modeling, we GO-annotated all proteins in our GV oocyte and cumulus data sets. We annotated all the



**Figure 1. Comparison of proteins identified in germinal vesicle (GV) oocytes and cumulus cells.** Distribution of predicted, known, and hypothetical proteins in GV oocytes and cumulus cells is shown. Superscript a, b and c=known, predicted and hypothetical proteins, respectively.

doi:10.1371/journal.pone.0011240.g001

proteins identified in GV oocyte and cumulus cells using GO and obtained annotation for 765 and 1159 proteins of GV oocyte and cumulus cell, respectively. Compared with our previous work, the GAQ score, which is measure of GO annotation quality [25] was almost doubled from 49.2 to 75 (Figure 2). This allowed us to do comprehensive modeling of bovine GV oocyte and cumulus cell communication. Grouping of biological process annotations into more generalized GO categories using a generic GO slim revealed 19 functional categories were represented in GV oocyte and cumulus cell proteomic datasets. Functional categories and percentages of proteins in each category from oocyte and cumulus are shown in Figure 3. The largest GO category represented in GV oocyte was related to cell communication with translation, transport, RNA metabolism, reproduction, and cytoskeletal organization and biogenesis also well represented (42.7% of the oocyte proteome). The largest represented GO category in cumulus cells was transport with cell communication, generation of precursor metabolites and energy, transcription, translation, and protein modification also substantially represented (40.6% of the cumulus cell proteome). Membrane and nuclear proteins are fundamental for inter- and intracellular signaling and are thus fundamental for modeling cell–cell interactions. From the cellular component GO, we identified 404 membrane proteins (24% of the total proteins): 248 unique to CC, 84 unique to oocytes, and 72 in both cell types. Using GO associations, we also identified 273 nuclear proteins: 172 unique to CC, 46 unique to oocytes, and 55 in both cell types (Table S3).

We next focused on the 371 proteins differentially expressed between GV oocyte and cumulus cells. Application of the AgBase generic GO Slim [29] revealed that 7 functional categories were represented in these differentially expressed proteins. In comparison to oocytes, cumulus cells had significantly higher expression of proteins involved in three biological processes: generation of precursor metabolites and energy, transport, and cell communication (Figure 4).

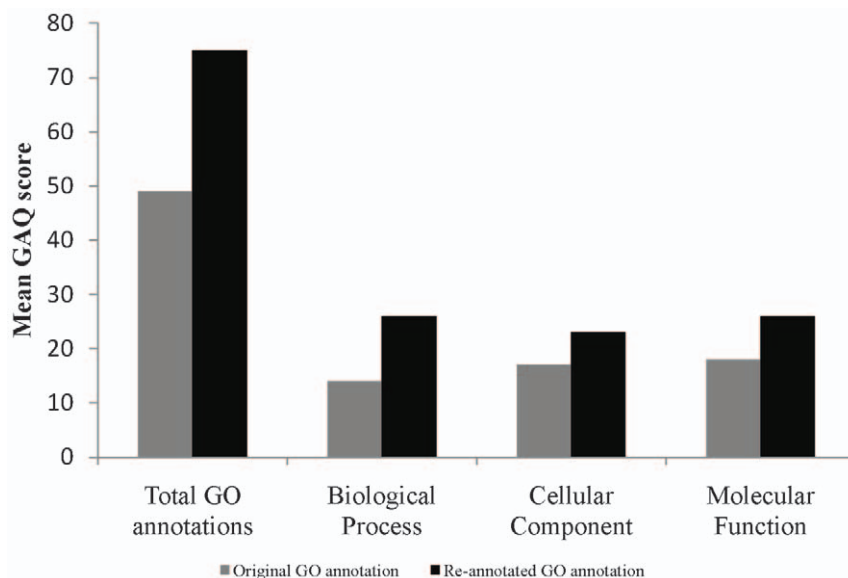
Our GO based quantitative modeling also showed that GV stage oocytes are biased towards biological processes such as cell cycle regulation, signal transduction, DNA transcription, protein

metabolism and modification, generation of precursor metabolites and energy, response to oxidative stress, protein amino acid phosphorylation, and cytoskeletal organization and biogenesis but biased against apoptosis (Figure 5).

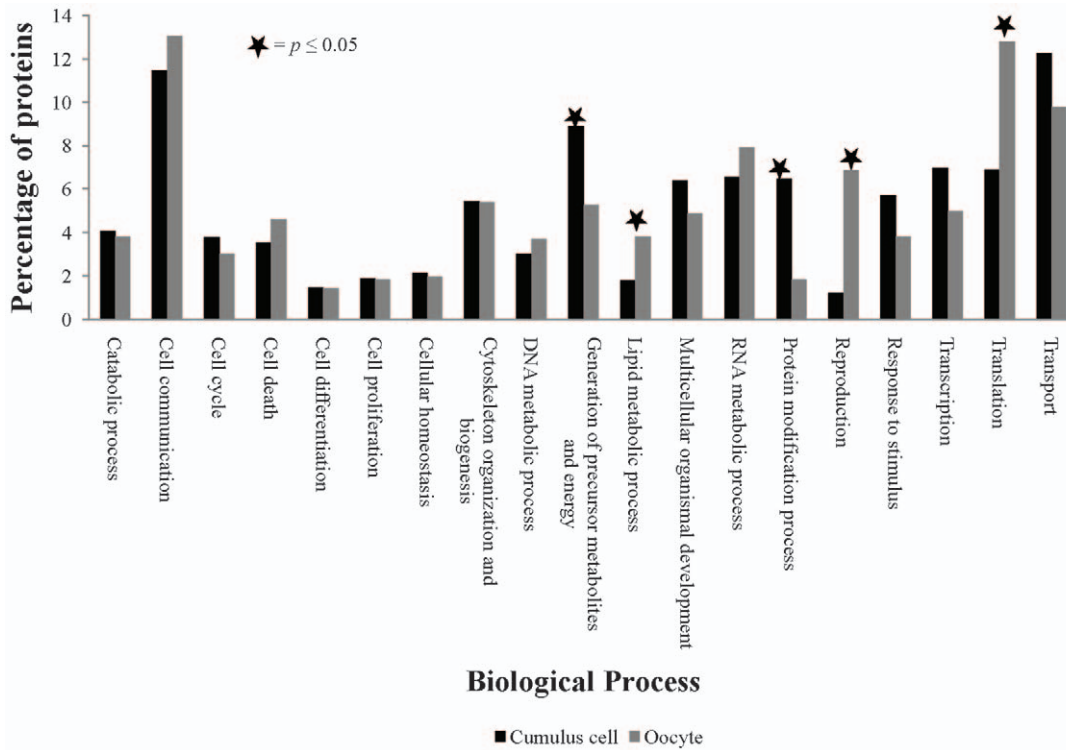
### Canonical genetic network analysis

Analysis of protein-protein interactions as part of complexes, pathways and biological networks is complimentary to analysis of functional annotations and here we used canonical pathway analysis. Among the 811 proteins identified in GV oocytes, 727 proteins had information about their contribution in canonical networks and functions/pathways. We identified 30 canonical networks and 49 functions. Functions of proteins involved in the top five networks were related to protein synthesis, DNA replication, recombination and repair, cell-to-cell signaling and interaction, molecular transport, amino acid metabolism, reproductive system development and function, small molecule biochemistry, and cellular function and maintenance. Of the 1247 proteins identified in cumulus cells, 1114 had information about their contribution in canonical genetic networks and functions/pathways, respectively. We identified 46 networks and 50 functions. Functions of proteins involved in top five networks were related to cell-to-cell signaling and interaction, molecular transport, protein synthesis, nucleic acid metabolism, cellular function and maintenance, small molecule biochemistry, molecular transport, RNA trafficking and post translational modification. The top five networks (ranked based on statistical significance), and their associated proteins are shown in Table 1 and 2 for GV oocyte and cumulus cell, respectively.

Twenty-three and 39 canonical pathways were significantly represented in GV oocytes and cumulus cells, respectively. Oxidative phosphorylation is the top represented canonical pathway significant only in cumulus cells. In this pathway, expression of 21 proteins was significantly altered and all these proteins have higher expression in cumulus cells compared to GV oocyte. These proteins include 9 ATP (adenosine triphosphate) synthases (ATP5A1, ATP5B, ATP5C1, ATP5F1, ATP5H, ATP5J, ATP5L, ATP5O, and ATP6V1E1), 3 cytochrome-c



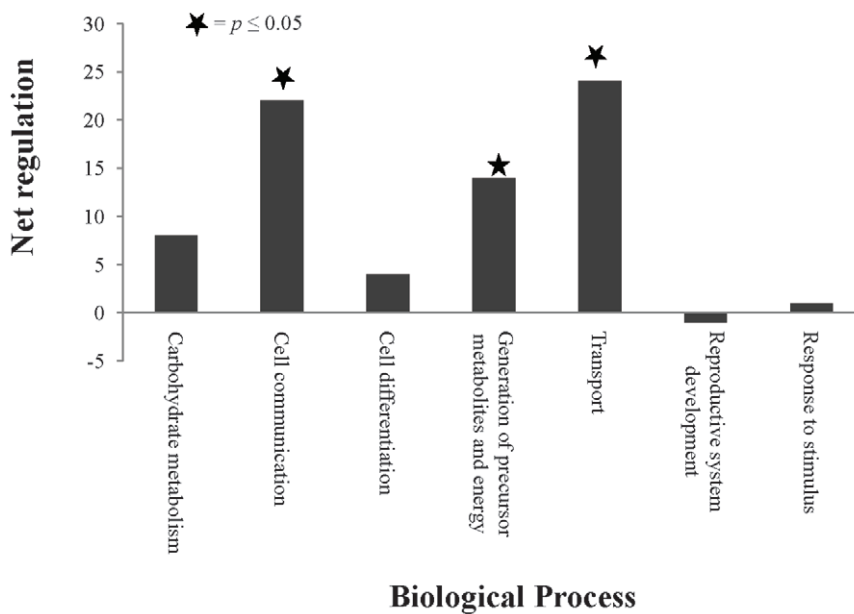
**Figure 2. Mean Gene ontology Annotation Quality (GAQ) scores.** Mean Gene ontology Annotation Quality (GAQ) score of original and improved Gene ontology annotations of germinal vesicle (GV) oocyte and cumulus cell proteome data sets before and after reannotation. doi:10.1371/journal.pone.0011240.g002



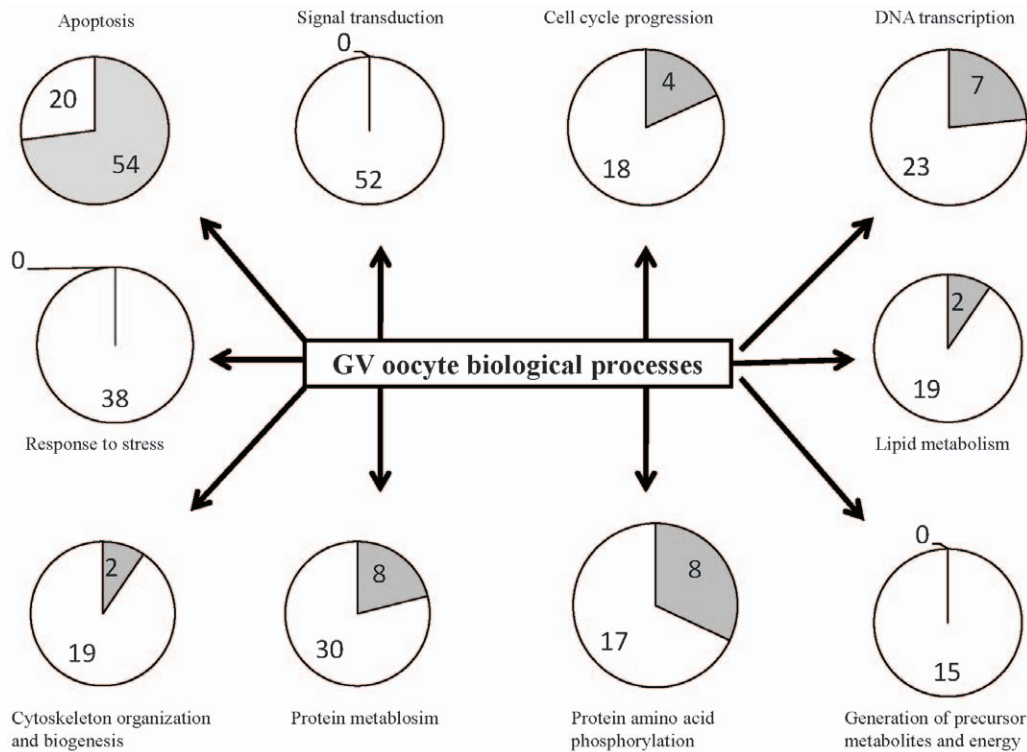
**Figure 3. Gene Ontology (GO) modeling of germinal vesicle (GV) oocyte and cumulus cell proteomes.** Distribution of percentages of GV stage oocyte and cumulus cell proteins involved in various biological processes. Significant differences in percentage of proteins involved in various identified GO categories in between GV oocyte and cumulus cells were evaluated by student's t-test. doi:10.1371/journal.pone.0011240.g003

oxidases (COX2, COX17, and COX5A), cytochrome C-1(CYC1), 6 NADH (reduced nicotinamide adenine dinucleotide) dehydrogenase (ubiquinone) complexes (NDUFA5, NDUFA9,

NDUFB, NDUFC2, NDUFS2, and NDUFV2) and 2 ubiquinol cytochrome c reductase core proteins (UQCRC1 and UQCRC2) (Table S2). Glycolysis and pyruvate metabolism pathways were



**Figure 4. Overall effects in GO Slims of differentially expressed proteins of germinal vesicle (GV) oocyte and cumulus cells.** Biological process GO annotations of all differentially-expressed proteins between GV oocyte and cumulus cells were used to generate GO Slims. For each GO Slim the difference in the numbers of proteins with higher expression and the number of proteins with lower expression in cumulus cells (relative to GV oocyte) was calculated to estimate the net regulatory effect. doi:10.1371/journal.pone.0011240.g004



**Figure 5. Germinal vesicle (GV) oocyte biological processes.** Numbers of proteins agonistic (white) or antagonistic (grey) for each biological process including signal transduction, cell cycle regulation, DNA transcription, apoptosis regulation, protein metabolism and modification, generation of precursor metabolites and energy, cytoskeleton organization and biogenesis, and response to stress and calculated the net effect. doi:10.1371/journal.pone.0011240.g005

significantly represented in both GV oocyte and cumulus cell proteomes. Expression of 12 key enzymes involved in glycolysis/pyruvate metabolism, including triosephosphate isomerase (TPIS), enolase 1 and 2 (ENO1 and ENO2), pyruvate kinase muscle (PKM), fructose-bisphosphate aldolase A (ALDOA), fructose-bisphosphate aldolase C (ALDOC), lactate dehydrogenase A (LDHA), phosphoglycerate mutase 1 (PGAM1), phosphoglycerate kinase 1 and 2 (PGK1 and PGK2), glucose phosphate isomerase (GPI), and aldehyde dehydrogenase 3 family member A2 (ALDH3A2) was higher in cumulus cells compared to oocyte (Table S2).

### Cell-to-cell signaling network

Bidirectional communication between oocytes and cumulus cells is essential for the development and function of both cell types [1,11,15]. We focused on the 91 identified proteins GO-annotated as involved in cell-to-cell signaling; 29 of these proteins were differentially-expressed between GV oocyte and cumulus cells (23 were greater and 6 lower in the cumulus cells compared to GV oocytes [Table S4]). The canonical network generated by IPA using proteins involved in cell-to-cell signaling is shown in Figure 6. We next overlaid canonical pathway information on to this network to identify important signaling pathways involved in the intricate cross talk between the GV oocyte and its surrounding cumulus cells. Major pathways associated with proteins involved in this cell-to-cell signaling network are related to integrin signaling, actin cytoskeleton signaling, mitogen-activated protein kinase (MAPK) signaling, phosphoinositide 3-kinase (PI3K) signaling, and ephrin receptor signaling (Table 3). Expression of six proteins—actin, beta (ACTB), actinin, alpha 4 (ACTN4), integrin alpha 2 (ITGA2), integrin alpha V (ITGAV), integrin beta 1

(ITGB1), and talin 1 (TLN1)—involved in integrin signaling and actin cytoskeleton signaling was higher in cumulus cells compared to GV oocyte (Figure 7) (Table S4). Expression of six proteins—zona pellucida glycoprotein 2 [ZP2], zona pellucida 3 [ZP3], periredoxin 2 [PRDX2], complement component 3 [C3], milk fat globule-EGF factor 8 protein [MFG8] and vitronectin [VTN]—involved in cell-to-cell signaling was significantly higher in GV oocyte compared to cumulus cells (Figure 6; Table S4).

### Discussion

Here we provide a foundational, proteomics-based descriptive computational systems-biology modeling of oocyte and cumulus cell interaction at the GV stage. Bidirectional communication between oocytes and cumulus cells is considered essential for proper maturation or ‘programming’ of oocytes, which is crucial for normal fertilization and embryonic development [15]. Cumulus cells are unique in that they are differentiated somatic cells essential for development of a competent oocyte. In turn, oocytes through the secretion of secreted factors (OSF’s), regulate a multitude of key cumulus cell functions, which may in turn produce positive regulatory factors that pass back to the oocyte, improving subsequent development [17]. However, despite their importance, little is known about the molecular events and pathways involved in the bidirectional communication between oocyte and cumulus cells. The proteins identified in this study in combination with the functional modeling using GO and IPA can serve as a basis for future hypothesis-driven research on follicle development, oocyte and cumulus cell communication, oocyte maturation, and related reproductive abnormalities.

In this study, we identified 811, and 1247 proteins of oocyte, and cumulus cells respectively, which is a significant increase in

**Table 1.** The top five biological networks in bovine GV oocyte proteome.

Network IDs	Proteins in Network <sup>a)</sup>	Score <sup>b)</sup>	Focus Proteins	Top Functions
1	<b>Akt</b> , <b>ANGPTL3</b> , <b>C17ORF61</b> , <b>C7ORF20</b> , <b>DDB1</b> , <b>EEF2</b> , <b>EEF1A1</b> , <b>EEF1A2</b> , <b>EEF1B2</b> , <b>EEF1D</b> , <b>EEF1G</b> , <b>ENO1</b> , <b>ENO3</b> , <i>Enolase</i> , <b>FAM62A</b> , <b>GPI</b> , <b>HARS</b> , <b>HRNR</b> , <b>HYOU1</b> , <b>ILF2</b> (includes EG:3608), <b>KRT10</b> , <b>PDCD5</b> , <b>PHIP</b> , <b>PLA2G1B</b> , <b>PPT1</b> , <i>Protein-synthesizing GTPase</i> , <b>RAD23B</b> , <b>RAI14</b> , <b>RPS7</b> , <b>SRGAP1</b> , <b>TSFM</b> (includes EG:10102), <b>TUBB4</b> , <b>TUFM</b> , <b>WARS</b> , <b>YWHAZ</b>	52	32	Protein synthesis, Lipid metabolism, Small-molecule biochemistry
2	<b>AHCY</b> , <b>AK2</b> , <b>AURKA</b> , <b>CACYBP</b> , <b>DDX6</b> , <b>EIF4ENIF1</b> , <b>EIF5A</b> , <b>ELAVL1</b> , <b>G3BP1</b> , <b>H2AFX</b> , <b>HIST3H2A</b> , <b>HMGB2</b> , <b>HNRNPK</b> , <i>Importin alpha/beta</i> , <i>Importin beta</i> , <b>NME2</b> , <b>NUP50</b> , <b>PCBP2</b> , <b>PCBP1</b> (includes EG:5093), <b>PDIA6</b> , <i>Pkc(s)</i> , <b>RAG1</b> , <b>RALY</b> , <b>RAN</b> , <b>RBMX</b> , <b>RPL23</b> , <b>RPS2</b> , <b>RPS20</b> , <b>TAGLN</b> , <b>TERF1</b> , <b>TGM1</b> , <b>TLN1</b> , <b>UBE2N</b> , <b>UGP2</b> , <b>UQCRC2</b>	52	32	Molecular transport, Protein trafficking, Amino acid metabolism
3	<b>ACR</b> , <i>Adenosine-tetraphosphatase</i> , <b>ALDH2</b> , <i>ATP synthase</i> , <b>ATP5A1</b> , <b>ATP5B</b> , <b>ATP5C1</b> , <b>ATP5G2</b> , <b>ATP5O</b> , <i>ATPase</i> , <b>BAT1</b> , <b>ERK</b> , <b>ETFA</b> (includes EG:2108), <b>ETFB</b> (includes EG:2109), <b>FAU</b> , <b>FHIT</b> , <b>GOT</b> , <b>GOT2</b> , <i>H<sub>+</sub>-transporting two-sector ATPase</i> , <b>HSPD1</b> , <b>HSPE1</b> , <b>M6PRBP1</b> , <b>MAPKAPK3</b> , <b>MFGE8</b> , <b>MVP</b> , <b>NSF</b> , <b>PJA1</b> , <b>RAB1A</b> , <b>RAB2A</b> , <b>RAB7A</b> , <b>RABEP1</b> , <i>Rsk</i> , <b>STX12</b> , <b>ZP2</b> , <b>ZP3</b>	42	28	Cell-to-cell signaling and interaction, Reproductive system development and function, DNA replication, recombination, and repair
4	<i>26s Proteasome</i> , <b>AGA</b> , <b>COX5B</b> , <b>DNAJA1</b> , <b>DNMT1</b> , <b>EPHA2</b> , <b>FKBP4</b> , <b>HSP</b> , <i>Hsp70</i> , <i>Hsp90</i> , <b>HSP90AA1</b> , <b>HSP90AB1</b> , <b>HSPA2</b> , <b>HSPA5</b> , <b>HSPA6</b> , <b>HSPA8</b> , <b>HSPA9</b> , <b>HSPB1</b> , <b>HSPH1</b> , <b>IARS2</b> , <b>MDH1</b> , <b>MTHFD1</b> , <i>Nos</i> , <b>PARK7</b> , <b>PARP1</b> , <b>PFN1</b> , <b>PGK1</b> , <i>PI3K</i> , <b>PPA1</b> , <b>PSAP</b> , <b>PTGES3</b> (includes EG:10728), <b>RPL7</b> , <b>RPS3A</b> , <i>Shc</i> , <b>STIP1</b>	42	28	Cellular function and maintenance, Cellular compromise
5	<b>AICDA</b> , <b>Ant</b> , <b>CYCS</b> (includes EG:54205), <i>Cytochrome c oxidase</i> , <b>DDOST</b> , <i>Dolichyl-diphosphooligosaccharide-protein glycotransferase</i> , <i>Glutathione peroxidase</i> , <b>GPX1</b> , <b>IL1F5</b> , <b>INPP1</b> , <i>Ldh</i> , <b>LDHA</b> , <b>LDHB</b> , <i>NF-kappaB (family)</i> , <i>NFkB (complex)</i> , <b>PECAM1</b> , <i>peroxidase (miscellaneous)</i> , <b>PRDX1</b> , <b>PRDX2</b> , <b>PRDX3</b> , <b>PRDX4</b> , <b>PRDX5</b> , <b>PRDX6</b> , <b>RAB3C</b> , <b>RPN1</b> , <b>RPN2</b> , <b>SLC25A3</b> , <b>SLC25A4</b> , <b>SLC25A10</b> , <b>SLC25A13</b> , <i>Sod</i> , <b>SOD2</b> , <b>TRPM8</b> , <b>TXN</b> , <b>VDAC1</b>	38	26	Small molecule biochemistry, Molecular transport, Cellular function and maintenance

a) The focus proteins are indicated with gene names and shown in bold letters.

b) A score of >2 is considered statistically significant.

doi:10.1371/journal.pone.0011240.t001

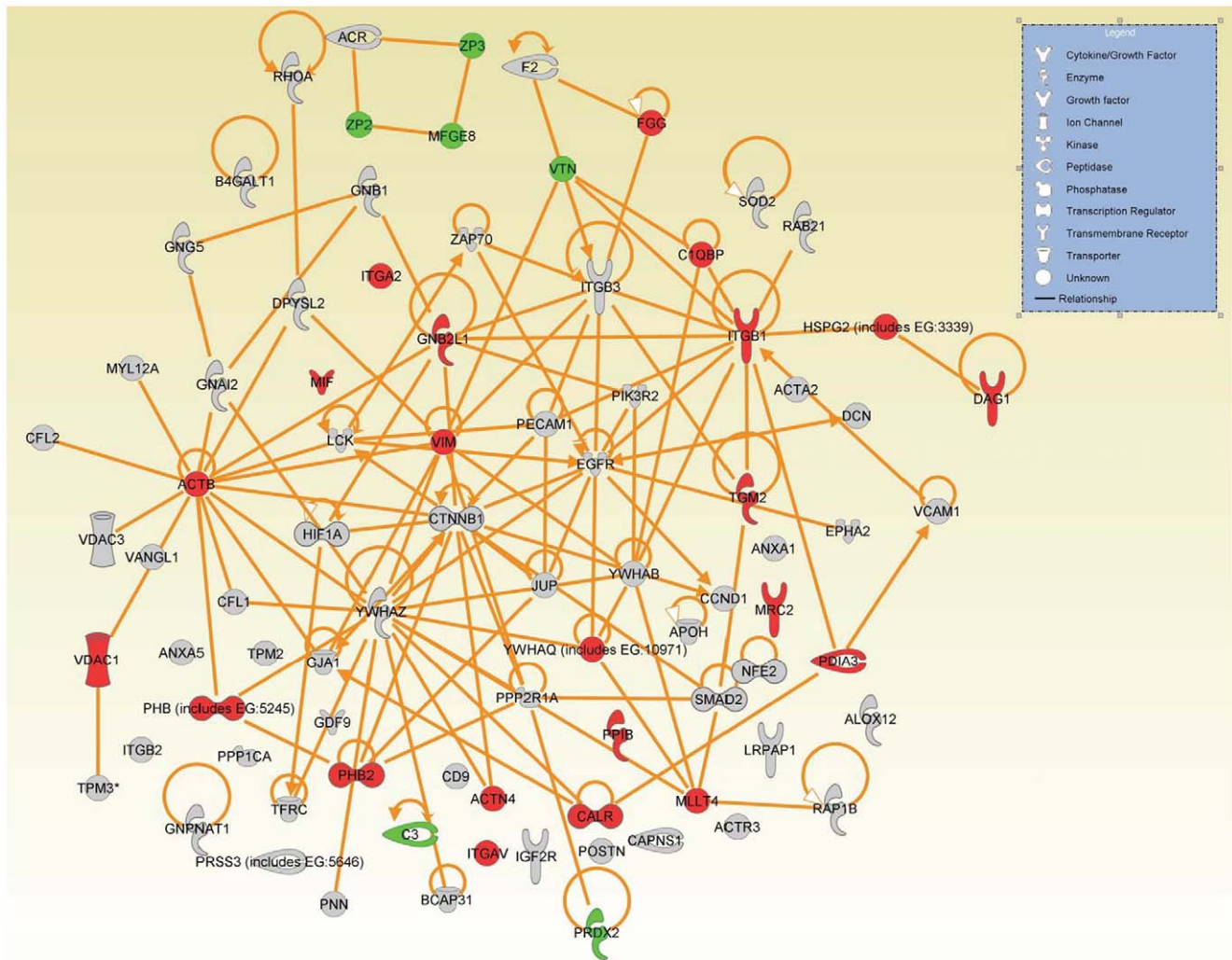
**Table 2.** The top five biological networks in bovine cumulus cell proteome.

Network IDs	Proteins in Network <sup>a)</sup>	Score <sup>b)</sup>	Focus Proteins	Top Functions
1	<b>ACADM</b> , <b>ATP1A2</b> , <b>BCKDHA</b> , <b>C7ORF20</b> , <b>CAPNS1</b> , <b>CSPG4</b> , <b>DPYSL2</b> , <i>Enolase</i> , <b>ERK</b> , <b>ETFA</b> (includes EG:2108), <b>ETFB</b> (includes EG:2109), <b>FAU</b> , <b>FHIT</b> , <b>GAK</b> , <b>HSPE1</b> , <b>IPO7</b> , <b>LGI1</b> , <b>MAP1B</b> , <b>MFGE8</b> , <b>MPZL1</b> , <b>MVP</b> , <b>NUMA1</b> , <b>PELP1</b> , <b>POSTN</b> , <b>PRKCSH</b> , <i>Rab11</i> , <b>RAB11B</b> , <b>RAB1A</b> , <b>TUBA3C</b> , <b>TUBA4A</b> , <b>TUBB2C</b> , <i>Tubulin</i> , <b>ULBP3</b> , <b>ZP2</b> , <b>ZP3</b>	43	31	Cell-To-Cell Signaling and Interaction, Reproductive system development and function
2	<b>ACTR6</b> , <b>AK2</b> , <b>ANXA4</b> , <b>AP3D1</b> , <b>BRP44</b> , <b>CALU</b> , <b>EEF2</b> , <b>EEF1A1</b> , <b>EEF1A2</b> , <b>EEF1B2</b> , <b>EEF1D</b> , <b>FOS</b> , <i>Immunoproteasome Pa28/20s</i> , <b>NIPSNAP1</b> , <b>OPTN</b> , <i>Proteasome PA700/20s</i> , <i>Protein-synthesizing GTPase</i> , <b>PSMA</b> , <b>PSMA1</b> , <b>PSMA2</b> , <b>PSMA4</b> , <b>PSMA6</b> , <b>PSMA7</b> , <b>PSMB2</b> , <b>PSMB4</b> , <b>PSME1</b> , <b>PSME2</b> , <b>PTRPN</b> , <b>RPLP1</b> , <b>RPLP2</b> , <b>RPLP0</b> (includes EG:6175), <b>RPS18</b> , <b>SEC23B</b> , <b>TKT</b> , <b>TUFM</b>	41	31	Protein synthesis, Molecular transport, Nucleic acid metabolism
3	<i>Adaptor protein 2</i> , <b>Ap1</b> , <b>AP1G1</b> , <b>BLVRA</b> , <b>BTF3</b> , <b>CD58</b> , <b>Ck2</b> , <i>Clathrin</i> , <b>CLTC</b> , <b>CSNK2A1</b> , <b>EIF5A</b> , <b>FKBP3</b> , <b>FKBP10</b> , <b>GTF2F1</b> , <b>HNRNPA2B1</b> , <b>HNRNPL</b> , <b>IGF2R</b> , <b>NUCKS1</b> , <i>Peptidylprolyl isomerase</i> , <i>Phosphatidylinositol4,5 kinase</i> , <b>PIN1</b> , <b>PIP4K2A</b> , <b>POLR3C</b> , <b>PIIB</b> , <b>PSMA3</b> , <b>RGS19</b> , <b>RPL5</b> (includes EG:6125), <b>SSB</b> (includes EG:6741), <b>TF</b> , <b>TFR2</b> , <b>TFRC</b> , <b>TOP1</b> , <b>TUBB4</b> , <b>VPS35</b> , <b>WARS</b>	38	29	Cellular function and maintenance, Small molecule biochemistry, Molecular transport
4	<i>3-hydroxyacyl-CoA dehydrogenase</i> , <b>ACAT1</b> , <i>Acetyl-CoA C-acetyltransferase</i> , <b>ATIC</b> , <b>B4GALT1</b> , <b>CYLD</b> , <b>DAD1</b> , <b>DDOST</b> , <i>Dolichyl-diphosphooligosaccharide-protein glycotransferase</i> , <b>ECH1</b> , <b>ECHS1</b> , <i>Enoyl-CoA hydratase</i> , <b>ENPP1</b> , <b>GOT</b> , <b>GOT2</b> , <b>HADHA</b> , <b>HADHB</b> , <b>HSD17B4</b> , <b>HSD17B10</b> , <i>NFkB (complex)</i> , <b>NME2</b> , <b>OSTC</b> , <b>PDIA6</b> , <b>PGRMC1</b> , <i>PPAR<math>\alpha</math>-RXR<math>\alpha</math></i> , <b>PRDX4</b> , <b>RPN1</b> , <b>RPN2</b> , <b>SFXN1</b> , <b>SLC27A1</b> , <b>STK10</b> , <b>STT3A</b> , <b>TOMM20</b> , <b>TOMM40</b> , <b>TOMM70A</b>	36	28	Lipid metabolism, Small molecule biochemistry, Post-translational modification
5	<b>ALPL</b> , <b>ANP32B</b> , <b>C21ORF33</b> , <b>CBX2</b> , <b>CCND1</b> , <b>CCT8</b> , <b>CDC37</b> , <b>DDX3X</b> , <b>DHR512</b> , <b>EIF3D</b> , <b>EIF4A2</b> , <b>EIF4ENIF1</b> , <b>ELAVL1</b> , <b>ENO3</b> , <b>H1FO</b> , <b>H3F3B</b> , <i>Hdac</i> , <i>Histone h3</i> , <i>Histone h4</i> , <b>HNRNPM</b> , <i>Importin beta</i> , <b>INHA</b> , <b>KPNB1</b> , <i>Mi2</i> , <b>NCAPD2</b> , <b>PNO1</b> , <b>RAN</b> , <b>SF3B1</b> , <b>SF3B3</b> , <i>SWI-SNF</i> , <i>Tcf/lef</i> , <b>THRAP3</b> , <b>TNPO1</b> , <b>VRK1</b> , <b>XPO1</b>	36	28	Molecular Transport, RNA Trafficking

a) The focus proteins are indicated with gene names and shown in bold letters.

b) A score of >2 is considered statistically significant.

doi:10.1371/journal.pone.0011240.t002



**Figure 6. Cell-to-cell signaling and interaction network.** Network generated with proteins involved in cell-to-cell signaling biological process using Ingenuity Pathway Analysis (IPA) as described in materials and methods. Each node represents a protein; proteins in shaded nodes were found in either GV oocyte or cumulus cell or both (see Table 3). Proteins in red and green nodes were higher and lower, respectively, in cumulus cells compared with GV oocytes.

doi:10.1371/journal.pone.0011240.g006

information compared to other female gametal proteomic studies [30,31,32,33,34,35]. Like all proteomics studies, we identified relatively small number of proteins compared to the numbers of mRNAs that can be identified in transcriptome studies. However, mRNAs are at least one step removed from phenotype; proteins are the functional molecules of cells and there is, in general, a very low correlation between amounts of mRNAs and amounts of the encoded protein [36,37,38,39]. Together this means that mRNA is a much less accurate predictor of phenotype than is protein. Even though the coverage of genetic pathways may be lower from proteomics experiments than transcriptomics experiments, because of the statistics that underlie the pathway analysis we are still able to identify relevant pathways that are important to the phenotype.

This current work is based on a separate previous proteomics study [19] which was smaller and generated less than one third of the data than this study did and did not use tissue replicates. It also uses a newer version of the bovine proteome. Though both studies follow accepted standards in the field, there are some significant technical differences between this and the previous

study which exemplify the maturation of the field of proteomics. In this present study, protein identification is more stringent and we used the probabilistic approach of decoy database searching [40]. Decoy database searching is based on the real data and uses a decoy database derived from this real database; as this maintains the same amino acid composition and the search space characteristics. Not only does this method allow a probability estimation for every peptide identification is a false positive (and thus protein) but also, because it uses mass spectra generated in the experiment, it does not rely on arbitrary cut off values derived from unrelated experiments [40]. Here we have more confident protein identifications than in the previous work, though we are likely to have increased our type II error. Even though we analyzed almost three times as many mass spectra in this new study, we ended up “confidently” identifying fewer proteins compared to previous study. In the previous study, we identified 5253 proteins were expressed by the cumulus cells and 1950 proteins were expressed by GV oocytes; in this study, we identified 1247 proteins in cumulus cells, and 811 proteins in GV oocytes.



**Table 3.** Top five signaling pathways associated with cell to cell signaling network.

Pathway name <sup>A</sup>	# of focus proteins	Focus proteins <sup>B</sup>	P-value
Integrin signaling	15	ACTA2, ↑ ACTB, ↑ ACTN4, ACTR3, CAPNS1, ↑ ITGA2, ITGB2, ↑ ITGAV, ↑ ITGB1, ITGB3, MYL12A, PIK3R2, PPP1CA, RAP1B, RHOA, ↑ TLN	1.01E-12
Actin cytoskeleton signaling	13	ACTA2, ↑ ACTB, ↑ ACTN4, ACTR3, CFL1, CFL2, F2, ↑ ITGA2, ↑ ITGB1, MYL12A, PIK3R2, PPP1CA, RHOA	5.48E-10
Ephrin receptor signaling	12	ACTR3, CFL1, CFL2, EPHA2, GNAI2, GNB1, ↑ GNB2L1, GNG5, ↑ ITGA2, ↑ ITGB1, RAP1B, RHOA,	1.11E-09
PI3K signaling	9	CCND1, CTNNB1, ↑ ITGA2, ↑ ITGB1, PIK3R2 PPP2RIA, YWHAB, ↑ YWHAQ, YWHAZ	5.44E-08
MAPK signaling	9	↑ ITGA2, ↑ ITGB1, PIK3R2, PPP2RIA, PPP1CA, RAP1B, YWHAB, ↑ YWHAQ, YWHAZ	2.09E-06

<sup>A</sup>MAPK : Mitogen activated protein kinase; PI3K ; Phosphoinositide 3-kinase.

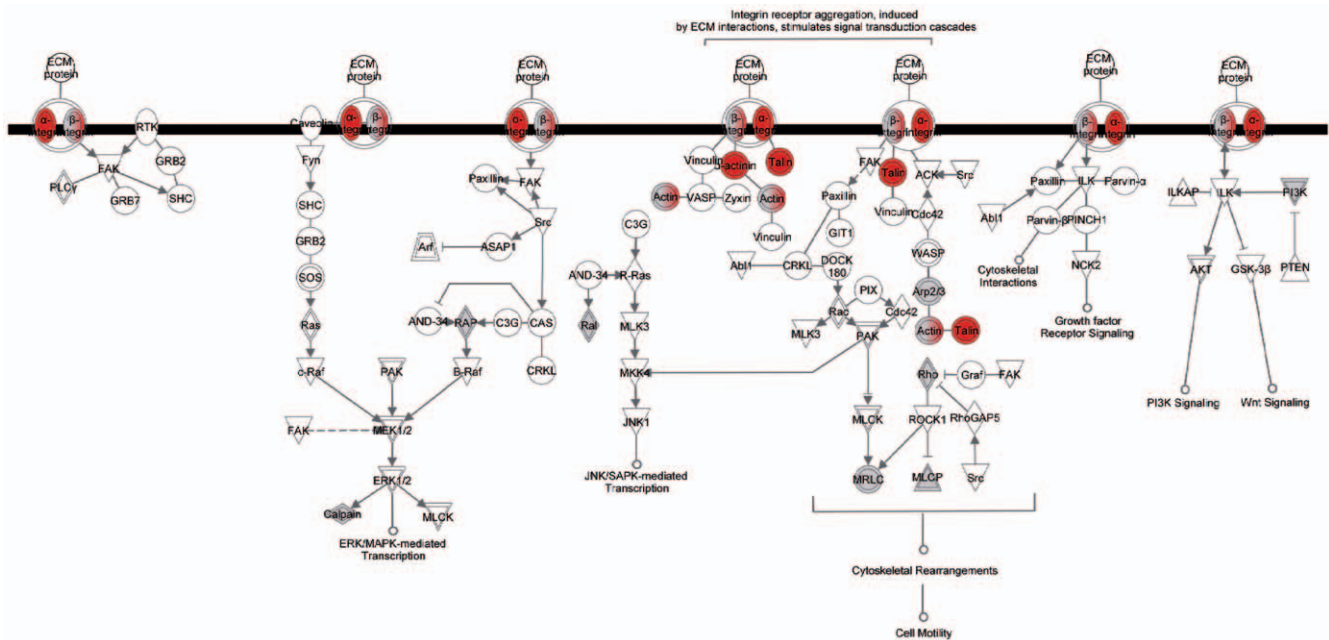
<sup>B</sup> ↑ higher expression in cumulus cells compared to GV oocyte.

doi:10.1371/journal.pone.0011240.t003

Maternal effect proteins are the proteins that are present in the oocyte, remain expressed throughout oocyte maturation and early embryonic development until maternal-to-embryonic transition (MET) and they may also be required for successful activation of the embryonic genome. We identified 11 maternal effect proteins in bovine oocytes: Maternal Antigen that Embryos Require (Mater; a.k.a NACHT, leucine rich repeat and PYD containing 5, *Nalp5*), zona pellucid proteins-2, 3 and 4 (ZP-2, -3, and -4), growth

differentiation factor 9 (GDF 9), beta-actin (ACTB), heat shock protein-70 (HSP-70), peroxiredoxins 1 and 2 (PRDX-1, -2), DNA (cytosine-5) methyl transferase one (DNMT1), and fibroblast growth factor-8 (FGF-8) all of which are important to oogenesis and early embryonic development [41,42,43,44,45,46].

Our previous publication described an initial GO-based functional analysis of bovine GV oocyte and cumulus cells using ‘known proteins’ and focusing on the cellular component Gene



**Figure 7. Integrin signaling pathway** Integrin signaling pathway generated by the Ingenuity Pathway Analysis (IPA) software. Integrin and actin cytoskeleton signaling pathways were the top two pathways associated with cell-to-cell signaling. Each node represents a protein; the proteins in shaded nodes in the pathway are identified/relate to an identified protein in the proteomic analysis. While proteins in clear nodes are part of the pathway but have not been identified in the GV oocyte or cumulus datasets. Proteins in red nodes were shown higher expression in cumulus cells compared to GV oocyte. doi:10.1371/journal.pone.0011240.g007

Ontology to identify membrane and nuclear proteins. Here, we take a much more comprehensive approach to include not only ‘known’ but also ‘predicted’ and ‘hypothetical’ proteins. By definition, predicted and hypothetical proteins do not have any functional annotation associated with them and they represent ~60% of total proteins in our oocyte and cumulus datasets. Hence we manually annotated all these predicted and hypothetical proteins based on sequence similarity and orthology from human, mouse and rat proteins. Overall our approach, while much more laborious and requiring recognized GO biocuration skills, provides more comprehensive information on the biological processes associated with bovine GV oocyte and cumulus cell function. Not only does this improved data quality allow a much better modeling, but we almost doubled the Gene ontology Annotation Quality (GAQ) score [25] (a quantitative measure of the GO annotation) of the proteins in our dataset and this will be valuable for others (the annotations are available at the AgBase database ([29]; [www.agbase.msstate.edu](http://www.agbase.msstate.edu)). By manual literature biocuration, we also captured the quantitative positive and negative regulatory roles that these GV oocyte proteins play in the various biological processes. Overall we could identify and quantify proteins agonistic or antagonistic to biological processes including signal transduction, cell cycle regulation, DNA transcription, apoptosis regulation, protein metabolism and modification, generation of precursor metabolites and energy, cytoskeleton organization and biogenesis, and response to stress in GV oocyte; all of which may have putative roles in oocyte competence and maturation. This type of quantitative modeling is not possible with the available pathway analysis tools. Furthermore, we have analyzed the protein datasets identified in this experiment in the context of protein interaction networks and pathways using Ingenuity Pathway Analysis. This pathway and network modeling also reveals signaling pathways associated with the proteins involved in cell-to-cell signaling biological process which may have implications in various reproductive processes such as oocyte development, and maturation.

Our GO-based modeling identified that the percentage of proteins involved in translation in GV oocytes is twice as high as in cumulus cells and this is consistent with other’s data showing that GV oocytes are transcriptionally and translationally active and also that proteins synthesized in this stage might be crucial to achieve maturation and successful subsequent development [19,47]. We identified a comparative up regulation of three biological processes—generation of precursor metabolites and energy, transport, and cell communication—in cumulus cells compared to oocytes. Up regulation of generation of precursor metabolites and energy in cumulus cells is consistent with a model that compensates for oocytes inability to metabolize glucose in which surrounding cumulus cells may absorb and metabolize glucose to provide products that can be utilized by oocytes for energy metabolism [48,49]. In support, cumulus cells nutritionally-support the developing oocytes by providing pyruvate [14,15,16]. Furthermore, stimulation of this nutritional support is in turn dependent on the presence of paracrine factors secreted by the oocyte (oocyte secreted factors (OSFs)); OSFs stimulate the glycolytic activity in cumulus cells by promoting the expression of genes involved in glycolysis [17] in mice, and our bovine model also suggests the glycolytic support of cumulus cells by showing higher expression of twelve proteins involved in glycolysis in cumulus cells associated with GV oocyte. Glycolysis is the metabolic pathway that converts glucose into pyruvate and the free energy released in this process is used to form the high energy compounds, ATP and NADH. We also identified oxidative phosphorylation as the most prominent metabolic pathway

significantly represented only in cumulus cells, but not in oocyte suggesting a hypothesis that cumulus cells may also provide energy to oocytes in the form of ATP. Expression of core proteins (ATP synthases, cytochrome oxidases, NADH dehydrogenases, and ubiquinol cytochrome c reductase) involved in oxidative phosphorylation was higher in cumulus cells compared to oocytes. NADH dehydrogenases and cytochrome C oxidases are, respectively, the first and last enzymes of the mitochondrial electron transport chain helping to establish a transmembrane difference of proton electrochemical potential that the ATP synthase then uses to synthesize ATP [50].

Cumulus cells provide a network of gap junction transmembrane channels facilitating two way communications for nutrient or paracrine factor exchange between oocytes and cumulus cells [1,3]. Several small metabolites (such as some energy substrates, nucleotides, and amino acids) are absorbed by oocytes mostly, or entirely, from surrounding cumulus cells via gap junctions [15,51]. Our bovine model also demonstrated an increase in transport in cumulus cells and this is in agreement with previous studies that cumulus cells have key role in amino acid and energy substrate uptake and transport of these substances to the oocyte.

Cumulus cells are essential for protecting oocytes from oxidative stress-induced apoptosis and DNA damage and their communication with oocytes is essential for development of an oocyte competent to undergo fertilization and embryogenesis [5,52]. We identified 91 proteins involved in cell-to-cell signaling from oocytes and cumulus cells. To effect intercellular communication several signaling pathways are necessary. We identified the integrin and actin cytoskeleton signaling pathways as the top two associated pathways with cell-to-cell signaling. Integrin signaling converges on cell cycle regulation, directing cells to live or die, to proliferate or to exit the cell cycle and differentiate [53]. Expression of ACTB, ACTN4, ITGAV, ITGA2, ITGB1, and TLN1, involved in integrin signaling and actin cytoskeleton signaling was higher in cumulus cells compared to oocytes. These proteins are the cell surface receptors present in the follicular basement membrane and around follicular cells and participate in cell attachment to matrix and mediate mechanical and chemical signals from it [54]. These signals in turn regulate the activities of cytoplasmic kinases, growth factor receptors, and ion channels that control the organization of the intracellular actin cytoskeleton [53,54]. The actin cytoskeleton of the human oocyte plays a key role in cell surface morphology as well processes associated with oocyte maturation [55]. Together this data suggests a hypothesis that oocytes depend on the presence of cumulus cells to generate specific cellular signals to coordinate their growth and maturation.

In addition to integrin and actin cytoskeleton signaling pathways, we also identified MAPK signaling, PI3K signaling, and ephrin receptor signaling pathways may involved in bidirectional oocyte-cumulus cell communication. MAPK and PI3K signaling are involved in activation of several membrane signaling molecules followed by sequential activation of protein kinases which involves highly regulated and modulated cascades of phosphorylation events [56]. Protein phosphorylation during this signaling plays a major role in the oocyte meiotic maturation and many protein kinases activated during oocyte maturation are also involved in cumulus cell proliferation and differentiation [57]. Expression of three proteins ITGA2, ITGB, and YWHAQ involved in MAPK and PI3K signaling was higher in cumulus cells compared to oocytes. YWHAQ, a protein of tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein family, modulate/complement intracellular events involving

phosphorylation dependent switching or protein modification in mammalian oocytes. The YWHAQ protein has been shown to protect phosphorylated proteins from inopportune dephosphorylation [58]. This is suggesting a hypothesis that cumulus cells have important role in oocyte maturation by regulating the protein phosphorylation, an important event during mammalian oocyte maturation.

We also observed expression of peroxiredoxins (PRDX's) involved in cell-to-cell signaling was significantly higher in oocytes compared to cumulus cells. PRDX's are peroxidases involved in antioxidant defence and intracellular signaling. Cumulus cells induce PRDX up regulation in oocytes [59] and our work also suggests that cumulus cells may have a key role in protecting oocytes from oxidative stress induced apoptosis and DNA damage via inducing oocytes to produce peroxiredoxins.

In conclusion, our systems biology modeling of bovine oocyte and cumulus proteomes, in the context of gene ontology and canonical protein interaction networks identified ninety-one proteins involved in the cell-to-cell signaling biological process that may have role in bidirectional communication between oocytes and cumulus cells. Most of the proteins involved in the cell-to-cell signaling biological process are the components of integrin, actin cytoskeleton, MAPK and PI3K signaling pathways that may have implications in various reproductive processes such as oocyte development, and maturation. In addition, GO-based analysis of differentially expressed proteins facilitated the identification that compared to GV oocytes, cumulus cells have higher expression of proteins involved in the generation of precursor metabolites and energy, transport, and cell communication. This systems biology model of bovine oocytes and cumulus cell potential interacting proteomes not only provides a foundation for understanding signaling and cell physiology at the GV stage of oocyte development, but is also valuable for comparative studies of other stages of oocyte development at the molecular level. The proteomes and biological process identified in this study can also serve as reference points for further comparative studies on immature and abnormal oocyte and cumulus cells elucidating the underlying molecular mechanisms involved in normal and pathological oocyte-cumulus cell communication. Furthermore, some of the proteins involved in cell-to-cell signaling may have value as molecular biomarkers which could be useful for assessing oocyte quality.

## Materials and Methods

### GV oocytes and cumulus cells

Separate from our previous work [19], ovaries from Holstein cows were collected by Biomed Inc. (Madison, WI, USA) from a local slaughter house. Cumulus-oocyte complexes (COC) were aspirated only from the follicles with a diameter of 2–8 mm using an 18-gauge needle attached to a vacuum system [60]. Only oocytes with intact cumulus cell layers and homogeneously granulated cytoplasm were selected, washed three times in TL-HEPES supplemented with polyvinylpyrrolidone (3 mg/ml polyvinyl-pyrrolidone-40; Sigma), Na-pyruvate (0.2 mM) and gentamycin (25 µg/ml). To obtain oocytes free of cumulus cells, COCs were vortexed in TL-HEPES (3 min), oocytes were collected under a stereomicroscope, further vortexed with hyaluronidase to remove adhering cumulus cells completely (3 min), washed thrice in saline and stored in a cell lysis buffer (4°C) until use. The cumulus cells separated from the oocytes after the first vortex were centrifuged, washed twice with saline, and the resulted pellets were resuspended in the lysis buffer and stored (4°C) until use.

### Protein extraction using differential detergent fractionation (DDF)

DDF sequentially extracts proteins from different cellular compartments using a series of detergents and provides information on cell location as well as increasing proteome coverage. Three replicates of five hundred GV oocytes and their surrounding CC were each subjected to DDF exactly as previously described [61]. Briefly, cytosolic proteins were extracted by six sequential incubations in a buffer containing digitonin (10 min each); next a fraction containing predominantly membrane proteins was isolated by incubating the cells in 10% Triton X-100 buffer for 30 min and then removing the soluble protein. Nuclear DDF buffer containing deoxycholate (DOC) was then added to the remaining pellet and subjected to freeze-thawing to disrupt the intact nuclear membrane. Nuclear proteins were collected from the resulting soluble fraction and the sample was then aspirated through an 18-gauge needle and treated with a mixture of DNase I (50U, Invitrogen, Carlsbad, CA) and RNase A (50 mg, Sigma-Aldrich, St Louis, MO) at 37°C for 1 h to digest nucleic acids. The remaining undissolved pellet was then treated with a buffer containing 5% SDS.

### Proteomics

Protein quantification and trypsin digestion of DDF fractions were done exactly as described [19]. Briefly, proteins were precipitated with 25% trichloroacetic acid to remove salts and detergents, then resuspended in 0.1 M ammonium bicarbonate with 5% HPLC grade acetonitrile (ACN), reduced (Dithiothreitol, 5mM, 65°C, 5 min), alkylated (iodoacetamide, 10 mM, 30°C, 30 min) and then trypsin digested until there was no visible pellet (sequencing grade modified trypsin, Promega; 1:50 w/w 37°C, 16 h). Peptides were desalted using a peptide microtrap (Michrom BioResources, Inc., Auburn, CA) and eluted using a 0.1% trifluoroacetic acid, 95% ACN solution. Desalted peptides were dried in a vacuum centrifuge and resuspended in 20µL of 0.1% formic acid and 5% ACN.

Two-dimensional liquid chromatography (LC) analysis was accomplished by strong cation exchange (SCX) followed by reverse phase (RP) coupled directly in line with an electrospray ionization (ESI) ion trap tandem mass spectrometer (LCQ; ThermoElectron Corp., San Jose, Calif, USA) essentially as described in [19]. The salt gradient applied in this study was different from our previous published method and was applied in steps of 0, 10, 15, 20, 25, 30, 35, 40, 45, 50, 57, 64, 90, and 700 mM ammonium acetate in 5% acetonitrile (ACN) and 0.1% formic acid. The reverse phase gradient used 0.1% formic acid in ACN and increased the ACN concentration in a linear gradient from 5% to 30% in 20 minutes and then 30% to 95% in 7 minutes, followed by 5% for 10 minutes for 0, 10, 15, 25, 30, 45, 64, 90, and 700 mM salt gradient steps. For 20, 35, 40, 50, and 57 mM salt gradient steps, ACN concentration was increased in a linear gradient from 5% to 40% in 65 minutes, 95% for 15 minutes, and 5% for 20 minutes.

Mass spectra and tandem mass spectra were searched against an *in silico* trypsin-digested bovine database using TurboSE-QUEST™ (Bioworks Browser 3.2; ThermoElectron). The bovine database used in this study was different from the previous publication and here we used database of bovine Reference Sequence (RefSeq) proteins downloaded from the National Center for Biotechnology Institute [NCBI; 08/28/2007; 25,078 entries]. The Reference Sequence (RefSeq) collection aims to provide a comprehensive, integrated, non-redundant, and well-annotated set of sequences of proteins [62]. Trypsin digestion including mass changes due to cysteine carbamidomethylation (C, 57.02 Da) and

methionine mono- and di-oxidation (15.99 Da and 32 Da), were included in the search criteria. The peptide (MS precursor ion) mass tolerance was set to 1.5 Da and the fragment ion ( $MS^2$ ) mass tolerance was set to 1.0 Da. Rsp Value less than 5.

As a primary filter we first limited our Sequest search output to include only peptides  $\geq 6$  amino acids long, with  $\Delta Cn \geq 0.1$  and Sequest cross correlation (Xcorr) scores of 1.9, 2.2, and 3.7 for +1, +2, and +3 charge states, respectively. We next used a decoy database search strategy [40] (using the same search criteria as the real database search) to calculate p-values for peptide identifications as this allows us to assign the probability of a false identification based on the real data from the experiment itself [40,63,64]. Since the accuracy of peptide identification depends on the charge state we calculated P-values for +1, +2, and +3 charge states separately. The probability that peptide identification from the original database is a random match (P-value) is estimated based on the probability that a match against the decoy database will achieve the same Xcorr [65,66]. Protein probabilities were calculated exactly as described [67,68] and we used only proteins identified by peptides with a  $p < 0.05$  for further analysis. All protein identifications and MS/MS data have been submitted to the PRoteomics IDentifications database (PRIDE [26]; accession numbers 8691, 8692, 8693, 8694, 8695 and 8696). PRIDE submission requirements are based on the proposed guidelines by proteomics standards initiative [69] and include all the peptides identified for each protein with their sequence, charge state, Xcorr, and delta cn.

We used an isotope-free quantification method [70] and a custom program *ProtQuant* [71] to identify differences in protein expression between oocyte and cumulus cell datasets. *ProtQuant* is a java based tool for label-free quantification that uses a spectral counting method with increased specificity (and thus decreased false positive i.e. type I errors). This increased specificity is achieved by incorporating the quantitative aspects of the Sequest cross correlation (XCorr) into the spectral counting method. *ProtQuant* also computes the statistical significance of differential expression by one-way ANOVA ( $\alpha \leq 0.05$ ).

### Gene Ontology (GO) Annotation

To identify the biological processes of all proteins in our datasets we used Gene Ontology annotations and the GO resources and tools available at AgBase [29]. We had to overcome the limitation that most literature for any species is not yet curated and so functional annotations from this literature are not yet available at GO databases. First, we used *GORetriever* to obtain all existing GO annotations available for known proteins in our datasets [72]. To obtain GO annotations for proteins without existing annotation, but between 70–90% amino acid sequence identity to presumptive orthologs with GO annotation were annotated using *GOanna* by manually checking the similarity between our proteins and orthologs [72]. GO biological process annotations for proteins were grouped into more generalized categories using *GOSlim viewer* [72]. Significant differences in percentage of proteins involved in various identified GO categories in between GV oocyte and cumulus cells were evaluated by *student's t-test*. Differences at  $p < 0.05$  were considered statistically significant.

GO annotation quality (*GAQ*) scores were calculated exactly as described [25] to quantify the improvement in GO annotation quality of GV oocyte and cumulus proteins due to re-annotation in this study compared to existing annotations. The *GAQ* score is a quantitative measure of the GO annotation of a set of gene products based on the number of GO annotations available, the level of detail of the annotation and the types of evidence used to make these GO annotations.

### GO based quantitative modeling

Our GO based quantitative modeling of GV stage oocyte was based on specific hypothesis framed in GO biological process (GOBP) terms defining the phenotype of oocyte competence and maturation. Although GOBP terms exists for gene products that effect and affect oocyte competence and maturation, and there is functional literature on these genes, the literature and GO are unconnected [42]. To connect the data with the GO, we manually annotated the literature of GV oocyte proteins to compare the number of proteins that were either agonistic or antagonistic for each biological process including signal transduction, cell cycle regulation, DNA transcription, apoptosis regulation, protein metabolism and modification, generation of precursor metabolites and energy, cytoskeleton organization and biogenesis, and response to stress and calculated the net effect of each process in GV oocyte; all of which may have a putative role in oocyte competence and maturation.

### Modeling using Ingenuity pathway analysis

To gain insights into the biological pathways and networks that are statistically significantly represented in our proteomic datasets we used Ingenuity Pathways Analysis (IPA; Ingenuity Systems, California). We imported protein accessions from our GV oocyte and cumulus cell datasets into IPA. IPA selects “focus proteins” to be used for generating biological networks. Focus proteins are the proteins from our datasets that are mapped to corresponding gene objects in the Ingenuity Pathways Knowledgebase (IPKB) and are known to interact with other proteins based on published, peer reviewed content in the IPKB. Based on these interactions IPA builds networks with a size of no more than 35 genes or proteins. A p-value for each network and canonical pathway is calculated according to the fit of the user's set of significant genes/proteins. IPA computes a score for each network from the p-value that indicates the likelihood of the focus proteins in a network being found together due to chance. We selected only networks scoring  $\geq 2$ , which have  $>99\%$  confidence of not being generated by chance [73] [74]. Biological functions are assigned to each network by using annotations from scientific literature and stored in the IPKB. A Fisher exact test is used to calculate the p-value determining the probability of each biological function/disease or pathway being assigned by chance. We used  $P \leq 0.05$  to select highly significant biological functions and pathways represented in our proteomic datasets [73].

### Supporting Information

**Figure S1** Comparison of proteins identified in present study with previous published study by Memili et al., 2007.

Found at: doi:10.1371/journal.pone.0011240.s001 (0.01 MB PDF)

**Table S1** List of proteins identified in germinal vesicle (GV) stage oocyte and cumulus cells. Proteins identified by DDF-MudPIT and their distribution in GV oocyte and cumulus cells. GI numbers of the identified proteins (column A) and corresponding protein names (assigned by NCBI; column B). Protein distribution in GV oocyte or cumulus cells or common to both (O: Oocyte; CC: Cumulus Cells; C: common to both; column C). Numbers of peptides and Sequest cross correlation scores ( $\Sigma Xcorr$ ) in columns D, F and E, G for oocyte and cumulus cells respectively.

Found at: doi:10.1371/journal.pone.0011240.s002 (0.29 MB XLS)

**Table S2** List of differentially expressed proteins in cumulus cells compared with GV oocytes. Gene name, numbers of peptides, Sequest cross correlation ( $\Sigma X_{corr}$ ) score and P value for of each differentially expressed protein in GV oocyte and cumulus cells. Found at: doi:10.1371/journal.pone.0011240.s003 (0.10 MB XLS)

**Table S3** Distribution of GV oocyte and cumulus cell proteins by subcellular compartments. The classification of the identified GV oocyte and cumulus proteins among subcellular compartments was performed using cellular component GO annotations. Found at: doi:10.1371/journal.pone.0011240.s004 (0.02 MB XLS)

**Table S4** Proteins involved in cell-to-cell signaling biological process. Found at: doi:10.1371/journal.pone.0011240.s005 (0.04 MB XLS)

## References

- Eppig JJ (1991) Intercommunication between mammalian oocytes and companion somatic cells. *Bioessays* 13: 569–574.
- Gilchrist RB, Ritter LJ, Armstrong DT (2004) Oocyte-somatic cell interactions during follicle development in mammals. *Anim Reprod Sci* 82–83: 431–446.
- Grazul-Bilska AT, Reynolds LP, Redmer DA (1997) Gap junctions in the ovaries. *Biol Reprod* 57: 947–957.
- Tanghe S, Van Soom A, Nauwynck H, Coryn M, de Kruif A (2002) Minireview: Functions of the cumulus oophorus during oocyte maturation, ovulation, and fertilization. *Mol Reprod Dev* 61: 414–424.
- Assidi M, Dufort I, Ali A, Hamel M, Algriany O, et al. (2008) Identification of Potential Markers of Oocyte Competence Expressed in Bovine Cumulus Cells Matured with Follicle-Stimulating Hormone and/or Phorbol Myristate Acetate In Vitro. *Biol Reprod*.
- Vozzi C, Formenton A, Chanson A, Senn A, Sahli R, et al. (2001) Involvement of connexin 43 in meiotic maturation of bovine oocytes. *Reproduction* 122: 619–628.
- Fatehi AN, Zeinstra EC, Kooij RV, Colenbrander B, Bevers MM (2002) Effect of cumulus cell removal of in vitro matured bovine oocytes prior to in vitro fertilization on subsequent cleavage rate. *Theriogenology* 57: 1347–1355.
- Mori T, Amano T, Shimizu H (2000) Roles of gap junctional communication of cumulus cells in cytoplasmic maturation of porcine oocytes cultured in vitro. *Biol Reprod* 62: 913–919.
- Wongsrikeao P, Kaneshige Y, Ooki R, Taniguchi M, Agung B, et al. (2005) Effect of the removal of cumulus cells on the nuclear maturation, fertilization and development of porcine oocytes. *Reprod Domest Anim* 40: 166–170.
- de Matos DG, Furnus CC, Moses DF (1997) Glutathione synthesis during in vitro maturation of bovine oocytes: role of cumulus cells. *Biol Reprod* 57: 1420–1425.
- Matzuk MM, Burns KH, Viveiros MM, Eppig JJ (2002) Intercellular communication in the mammalian ovary: oocytes carry the conversation. *Science* 296: 2178–2180.
- Su YQ, Sugiura K, Wigglesworth K, O'Brien MJ, Affourtit JP, et al. (2008) Oocyte regulation of metabolic cooperativity between mouse cumulus cells and oocytes: BMP15 and GDF9 control cholesterol biosynthesis in cumulus cells. *Development* 135: 111–121.
- Sugiura K, Su YQ, Diaz FJ, Pangas SA, Sharma S, et al. (2007) Oocyte-derived BMP15 and FGFs cooperate to promote glycolysis in cumulus cells. *Development* 134: 2593–2603.
- Biggers JD, Whittingham DG, Donahue RP (1967) The pattern of energy metabolism in the mouse oocyte and zygote. *Proc Natl Acad Sci U S A* 58: 560–567.
- Buccione R, Schroeder AC, Eppig JJ (1990) Interactions between somatic cells and germ cells throughout mammalian oogenesis. *Biol Reprod* 43: 543–547.
- Kim H, Schuetz AW (1991) Regulation of nuclear membrane assembly and maintenance during in vitro maturation of mouse oocytes: role of pyruvate and protein synthesis. *Cell Tissue Res* 265: 105–112.
- Sugiura K, Pendola FL, Eppig JJ (2005) Oocyte control of metabolic cooperativity between oocytes and companion granulosa cells: energy metabolism. *Dev Biol* 279: 20–30.
- Eppig JJ, Wigglesworth K, Chesnel F (1993) Secretion of cumulus expansion enabling factor by mouse oocytes: relationship to oocyte growth and competence to resume meiosis. *Dev Biol* 158: 400–409.
- Memili E, Peddinti D, Shack LA, Nanduri B, McCarthy F, et al. (2007) Bovine germinal vesicle oocyte and cumulus cell proteomics. *Reproduction* 133: 1107–1120.
- Sutton ML, Cetica PD, Beconi MT, Kind KL, Gilchrist RB, et al. (2003) Influence of oocyte-secreted factors and culture duration on the metabolic activity of bovine cumulus cell complexes. *Reproduction* 126: 27–34.
- Amano T, Mori T, Matsumoto K, Iritani A, Watanabe T (2005) Role of cumulus cells during maturation of porcine oocytes in the rise in intracellular Ca<sup>2+</sup> induced by inositol 1,4,5-trisphosphate. *Theriogenology* 64: 261–274.
- Zhang L, Jiang S, Wozniak PJ, Yang X, Godke RA (1995) Cumulus cell function during bovine oocyte maturation, fertilization, and embryo development in vitro. *Mol Reprod Dev* 40: 338–344.
- Richter JD (1991) Translational control during early development. *Bioessays* 13: 179–183.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: tool for the unification of biology. *The Gene Ontology Consortium. Nat Genet* 25: 25–29.
- Buza TJ, McCarthy FM, Wang N, Bridges SM, Burgess SC (2008) Gene Ontology annotation quality analysis in model eukaryotes. *Nucleic Acids Res* 36: e12.
- Martens L, Hermjakob H, Jones P, Adamski M, Taylor C, et al. (2005) PRIDE: the proteomics identifications database. *Proteomics* 5: 3537–3545.
- McCarthy FM, Cooksey AM, Wang N, Bridges SM, Pharr GT, et al. (2006) Modeling a whole organ using proteomics: the avian bursa of Fabricius. *Proteomics* 6: 2759–2771.
- Lubec G, Afjehi-Sadat L, Yang JW, John JP (2005) Searching for hypothetical proteins: theory and practice based upon original data and literature. *Prog Neurobiol* 77: 90–127.
- McCarthy FM, Bridges SM, Wang N, Magee GB, Williams WP, et al. (2007) AgBase: a unified resource for functional analysis in agriculture. *Nucleic Acids Res* 35: D599–603.
- Bhojwani M, Rudolph E, Kanitz W, Zuehlke H, Schneider F, et al. (2006) Molecular analysis of maturation processes by protein and phosphoprotein profiling during in vitro maturation of bovine oocytes: a proteomic approach. *Cloning Stem Cells* 8: 259–274.
- Coonrod SA, Wright PW, Herr JC (2002) Oolemmal proteomics. *J Reprod Immunol* 53: 55–65.
- Ellederova Z, Halada P, Man P, Kubelka M, Motlik J, et al. (2004) Protein patterns of pig oocytes during in vitro maturation. *Biol Reprod* 71: 1533–1539.
- Hamamah S, Matha V, Berthet C, Anahory T, Loup V, et al. (2006) Comparative protein expression profiling in human cumulus cells in relation to oocyte fertilization and ovarian stimulation protocol. *Reprod Biomed Online* 13: 807–814.
- Ma M, Guo X, Wang F, Zhao C, Liu Z, et al. (2008) Protein expression profile of the mouse metaphase-II oocyte. *J Proteome Res* 7: 4821–4830.
- Meng Y, Liu XH, Ma X, Shen Y, Fan L, et al. (2007) The protein profile of mouse mature cumulus-oocyte complex. *Biochim Biophys Acta* 1774: 1477–1490.
- Chen G, Gharib TG, Huang CC, Taylor JM, Misk DE, et al. (2002) Discordant protein and mRNA expression in lung adenocarcinomas. *Mol Cell Proteomics* 1: 304–313.
- Cox B, Kislinger T, Emili A (2005) Integrating gene and protein expression data: pattern analysis and profile mining. *Methods* 35: 303–314.
- Gygi SP, Rochon Y, Franza BR, Aebersold R (1999) Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol* 19: 1720–1730.
- Nie L, Wu G, Culley DE, Scholten JC, Zhang W (2007) Integrative analysis of transcriptomic and proteomic data: challenges, solutions and applications. *Crit Rev Biotechnol* 27: 63–75.
- Elias JE, Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* 4: 207–214.
- Dalbies-Tran R, Papillier P, Penneret S, Uzbekova S, Monget P (2005) Bovine mater-like NALP9 is an oocyte marker gene. *Mol Reprod Dev* 71: 414–421.

## Acknowledgments

We would like to acknowledge Dr. Susan M Bridges, Dr. Bindu Nanduri and Bryce Magee for helping us with Biomarker discovery tool, Dr. Teresia Buza for GAQ score and Dr. Tibor Pechan for running the mass spectrometer.

## Author Contributions

Conceived and designed the experiments: DP EM SB. Performed the experiments: DP. Analyzed the data: DP. Contributed reagents/materials/analysis tools: SB. Wrote the paper: DP. Performed the proteomics sample preparation, data generation, analyzed and interpreted proteomic data, computational systems biology modeling and analysis and wrote the draft of the manuscript: DP. Facilitated sample collection, contributed to design of the study, provided expert knowledge and interpretation in reproductive biology: EM. Helped analyze and interpret the statistical analysis of the proteomics data and helped to draft the manuscript: SCB. Reviewed manuscript draft: SBC.

42. Fair T, Carter F, Park S, Evans AC, Lonergan P (2007) Global gene expression analysis during bovine oocyte in vitro maturation. *Theriogenology* 68 Suppl 1: S91–97.
43. Penmetier S, Perreau C, Uzbekova S, Thelie A, Delaleu B, et al. (2006) MATER protein expression and intracellular localization throughout folliculogenesis and preimplantation embryo development in the bovine. *BMC Dev Biol* 6: 26.
44. Pfeffer PL, Sisco B, Donnison M, Somers J, Smith C (2007) Isolation of genes associated with developmental competency of bovine oocytes. *Theriogenology* 68 Suppl 1: S84–90.
45. Racedo SE, Wrenzycki C, Lepikhov K, Salamone D, Walter J, et al. (2009) Epigenetic modifications and related mRNA expression during bovine oocyte in vitro maturation. *Reprod Fertil Dev* 21: 738–748.
46. Thelie A, Papillier P, Penmetier S, Perreau C, Traverso JM, et al. (2007) Differential regulation of abundance and deadenylation of maternal transcripts during bovine oocyte maturation in vitro and in vivo. *BMC Dev Biol* 7: 125.
47. Memili E, First NL (1999) Control of gene expression at the onset of bovine embryonic development. *Biol Reprod* 61: 1198–1207.
48. Donahue RP, Stern S (1968) Follicular cell support of oocyte maturation: production of pyruvate in vitro. *J Reprod Fertil* 17: 395–398.
49. Leese HJ, Barton AM (1985) Production of pyruvate by isolated mouse cumulus cells. *J Exp Zool* 234: 231–236.
50. Weber J, Senior AE (2003) ATP synthesis driven by proton transport in F1F0-ATP synthase. *FEBS Lett* 545: 61–70.
51. Sugiura K, Eppig JJ (2005) Society for Reproductive Biology Founders' Lecture 2005. Control of metabolic cooperativity between oocytes and their companion granulosa cells by mouse oocytes. *Reprod Fertil Dev* 17: 667–674.
52. Bettgowda A, Patel OV, Lee KB, Park KE, Salem M, et al. (2008) Identification of Novel Bovine Cumulus Cell Molecular Markers Predictive of Oocyte Competence: Functional and Diagnostic Implications. *Biol Reprod*.
53. Giancotti FG, Ruoslahti E (1999) Integrin signaling. *Science* 285: 1028–1032.
54. Monniaux D, Huet-Calderwood C, Le Bellego F, Fabre S, Monget P, et al. (2006) Integrins in the ovary. *Semin Reprod Med* 24: 251–261.
55. Veselska RRJ (2001) Cortical actin cytoskeleton in human oocytes: a comparison with mouse oocytes. *Scripta medica* 74: 265–274.
56. Seger R, Krebs EG (1995) The MAPK signaling cascade. *Faseb J* 9: 726–735.
57. Su YQ, Wigglesworth K, Pendola FL, O'Brien MJ, Eppig JJ (2002) Mitogen-activated protein kinase activity in cumulus cells is essential for gonadotropin-induced oocyte meiotic resumption and cumulus expansion in the mouse. *Endocrinology* 143: 2221–2232.
58. Snow AJ, Puri P, Acker-Palmer A, Bouwmeester T, Vijayaraghavan S, et al. (2008) Phosphorylation-dependent interaction of tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein (YWHA) with PADI6 following oocyte maturation in mice. *Biol Reprod* 79: 337–347.
59. Leyens G, Verhaeghe B, Landtmeters M, Marchandise J, Knoop B, et al. (2004) Peroxiredoxin 6 is upregulated in bovine oocytes and cumulus cells during in vitro maturation: role of intercellular communication. *Biol Reprod* 71: 1646–1651.
60. Sagirkaya H, Misirlioglu M, Kaya A, First NL, Parrish JJ, et al. (2006) Developmental and molecular correlates of bovine preimplantation embryos. *Reproduction* 131: 895–904.
61. Peddinti D, Nanduri B, Kaya A, Feugang JM, Burgess SC, et al. (2008) Comprehensive proteomic analysis of bovine spermatozoa of varying fertility rates and identification of biomarkers associated with fertility. *BMC Syst Biol* 2: 19.
62. Pruitt KD, Tatusova T, Maglott DR (2007) NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* 35: D61–65.
63. Elias JE, Gibbons FD, King OD, Roth FP, Gygi SP (2004) Intensity-based protein identification by machine learning from a library of tandem mass spectra. *Nat Biotechnol* 22: 214–219.
64. Peng J, Elias JE, Thoreen CC, Licklider LJ, Gygi SP (2003) Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J Proteome Res* 2: 43–50.
65. Park GW, Kwon KH, Kim JY, Lee JH, Yun SH, et al. (2006) Human plasma proteome analysis by reversed sequence database search and molecular weight correlation based on a bacterial proteome analysis. *Proteomics* 6: 1121–1132.
66. Qian WJ, Liu T, Monroe ME, Strittmatter EF, Jacobs JM, et al. (2005) Probability-based evaluation of peptide and protein identifications from tandem mass spectrometry and SEQUEST analysis: the human proteome. *J Proteome Res* 4: 53–62.
67. Nesvizhskii AI, Keller A, Kolker E, Aebersold R (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* 75: 4646–4658.
68. MacCoss MJ, Wu CC, Yates JR, 3rd (2002) Probability-based validation of protein identifications using a modified SEQUEST algorithm. *Anal Chem* 74: 5593–5599.
69. Orchard S, Jones AR, Stephan C, Binz PA (2007) The HUPO pre-congress Proteomics Standards Initiative workshop. HUPO 5th annual World Congress. Long Beach, CA, USA 28 October–1 November 2006. *Proteomics* 7: 1006–1008.
70. Nanduri B, Lawrence ML, Vanguri S, Pechan T, Burgess SC (2005) Proteomic analysis using an unfinished bacterial genome: the effects of subminimum inhibitory concentrations of antibiotics on *Mannheimia haemolytica* virulence factor expression. *Proteomics* 5: 4852–4863.
71. Bridges SM, Bryce GB, Wang N, Williams WP, Burgess SC, et al. (2007) ProtQuant: a tool for the label-free quantification of mudPIT proteomics data. *BMC Bioinformatics* 8.
72. McCarthy FM, Wang N, Magee GB, Nanduri B, Lawrence ML, et al. (2006) AgBase: a functional genomics resource for agriculture. *BMC Genomics* 7: 229.
73. Gerling IC, Singh S, Lenchik NI, Marshall DR, Wu J (2006) New data analysis and mining approaches identify unique proteome and transcriptome markers of susceptibility to autoimmune diabetes. *Mol Cell Proteomics* 5: 293–305.
74. Huang Y, Yan J, Lubet R, Kensler TW, Sutter TR (2006) Identification of novel transcriptional networks in response to treatment with the anticarcinogen 3H-1,2-dithiole-3-thione. *Physiol Genomics* 24: 144–153.