



**HAL**  
open science

## Differential transcript usage across mammalian oocytes at the germinal vesicle and metaphase II stages

Thanida Sananmuang, Denis Puthier, Catherine Nguyen, Kaj Chokeshaiusaha

► **To cite this version:**

Thanida Sananmuang, Denis Puthier, Catherine Nguyen, Kaj Chokeshaiusaha. Differential transcript usage across mammalian oocytes at the germinal vesicle and metaphase II stages. *Theriogenology*, 2024, 215 (776), pp.1-9. 10.1016/j.theriogenology.2023.11.016 . hal-04886218

**HAL Id: hal-04886218**

**<https://amu.hal.science/hal-04886218v1>**

Submitted on 14 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Differential transcript usage across mammalian oocytes at the germinal vesicle and metaphase II stages**

2 Thanida Sananmuang<sup>1</sup>, Denis Puthier<sup>2</sup>, Catherine Nguyen<sup>2</sup>, and Kaj Chokeshaiusaha<sup>1\*</sup>

3

4 **\*Corresponding author: Kaj Chokeshaiusaha**

5 E-mail: [kaj.chk@gmail.com](mailto:kaj.chk@gmail.com)

6

7 <sup>1</sup>Rajamangala University of Technology Tawan-OK, Faculty of Veterinary Medicine, Chonburi, Thailand.

8 <sup>2</sup>Aix-Marseille Université, INSERM UMR 1090, TAGC, Marseille, France.

9

10 Running title: Comparative analysis of transcript utilization of oocytes across multiple species

11

12 **\*ORCID**

13 Kaj Chokeshaiusaha <https://orcid.org/0000-0002-2953-9101>

14 Thanida Sananmuang <https://orcid.org/0000-0002-8653-3558>

15 Denis Puthier <https://orcid.org/0000-0002-7240-5280>

16 Catherine Nguyen <https://orcid.org/0000-0001-9376-6360>

17 **Abstract**

18 Ongoing progress in mRNA-Sequencing technologies has significantly contributed to the refinement of  
19 assisted reproductive technologies. However, the prior investigations have predominantly concentrated on alterations in  
20 overall gene expression levels, thereby leaving a considerable gap in our understanding of the influence of transcript  
21 isoform expression on fundamental cellular mechanisms of oocytes. Given the efficacy of differential transcript usage  
22 (DTU) analysis to address such knowledge, we conducted comprehensive DTU analysis utilizing mRNA-Seq datasets  
23 of germinal vesicle (GV) and metaphase II (MII) oocytes across six mammalian species from the SRA database,  
24 including cow, donkey, horse, human, mouse, and pig. To further illuminate the roles of these genes, we also conducted  
25 a rigorous Gene Ontology (GO) term enrichment analysis. While the DTU analysis of each species exhibited several  
26 genes with alterations in their transcript isoform usage, referred to as DTU genes, this study focused on only ten cross-  
27 species DTU genes sharing among a minimum of five distinct species ( $FDR \leq 0.05$ ). These cross-species DTU genes  
28 were as follows: ABCF1, CDC6, CFAP36, CNOT10, DNM3, IWS1, NBN, NDEL1, RAD50 and ZCCHC17. GO term  
29 enrichment analysis unveiled the alignment of these cross-species DTU gene functions with RNA and cell-cycle control  
30 mechanisms across diverse mammalian species, thereby suggesting their vital roles during oocyte maturation. Further  
31 exploration of the transcript isoforms of these genes hence bore the potential to uncover novel transcript isoform  
32 markers for future reproductive technologies in both human and animal contexts.

33

34 **Keywords:** mRNA-Sequencing; Differential transcript usage; Gene Ontology; Transcript isoform markers

35

## 36 **Introduction**

37 High-quality oocytes are essential for the newborn generation in all mammalian species. An aberrant process  
38 during oocyte maturation—the final stages of oocyte development can lead to infertility or the incompetent  
39 development of early embryos [1,2]. Oocyte maturation encompasses the dynamic process initiated by the first meiotic  
40 resumption of the oocyte in the germinal vesicle (GV) stage through metaphase I (MI) and ultimately metaphase II  
41 (MII) stages, during which oocyte meiosis arrests once again. Until fertilization occurs, the delay in the second  
42 resumption of this arrest will persist [2,3].

43 Fine-tuning the progression of the oocyte from meiotic arrest to meiotic resumption is crucial for complete  
44 oocyte maturation. Disturbance during this process may fail oogenesis, fertilization, and consequent embryogenesis  
45 [1,3,4]. Therefore, understanding the regulatory network that governs the succession of oocyte meiosis not only clarifies  
46 the factors contributing to oocyte fertility; but also sheds light on potential biomarkers to determine its quality for  
47 further implementation. It is crucial to acknowledge that the current criteria used to assess oocyte quality for assisted  
48 reproductive treatment (ART), particularly *in vitro* fertilization (IVF), primarily rely on morphological observation.  
49 However, several oocytes with normal morphology fail to fertilize within the IVF program [2,5]. This observation  
50 highlights the significance of conducting oocyte gene expression profiling studies to identify improved biomarkers for  
51 such assessments.

52 To obtain a comprehensive overview of the cellular and molecular mechanisms during oocyte maturation,  
53 high-throughput sequencing of cDNA (RNA-seq) has become widespread to characterize the oocyte transcriptome  
54 [1,2,6]. Several transcriptomic studies endeavor to compare the transcriptomic abundance at the gene expression level,  
55 focusing on oocytes in the GV and MII stages. While the outcome of such studies has elucidated numerous genes  
56 related to the oocyte maturation's competency, they cannot provide insight into changes in transcript variants and  
57 isoform expressions. It is imperative to acknowledge that gene-expression level studies typically depend on the one  
58 gene, one polypeptide concept, which cannot adequately account for variations in polypeptide isoforms arising from  
59 different transcripts of the same gene [7–10].

60 Of note, when there is a shift in expression levels among two or more isoforms, the overall gene expression  
61 may not exhibit a noticeable change. However, this variation in transcript usage, known as differential transcript usage  
62 (DTU), holds significant biological relevance and importance. Numerous biological studies have consistently observed  
63 the robustness of DTU when comparing expression across different cell types, identifying several distinct patterns of  
64 cell-specific isoforms among the expressed genes [11]. DTU can lead to changes in the coding sequence of the  
65 transcript, resulting in functionally different gene products or even impacting the binding sites of RNA-binding proteins  
66 [12]. Unsurprisingly, specific patterns of DTU have been identified in various diseases and disorders [13]. Given its

67 significance, some recent studies have utilized mRNA sequencing data and other methodologies to gain insight into the  
68 DTU of oocytes. Nevertheless, these studies focused on mouse oocytes [14,15], limiting the extrapolation of the  
69 acquired knowledge to other mammalian species.

70 Undoubtedly, the DTU analysis across different species holds the potential to emerge as a novel field in  
71 theriogenology. Similar to cross-species differential gene expression (DGE) analysis, conducting cross-species  
72 differential transcript usage (DTU) analysis between oocytes in the germinal vesicle (GV) and metaphase II (MII)  
73 stages could yield valuable evidence regarding the universal factors governing the quality of oocyte maturation through  
74 transcript isoform switching. The significance of the outcomes derived from such an analysis could potentially elucidate  
75 new factors in conjunction with other crucial biomarkers previously identified by DGE approaches. Furthermore, it may  
76 open up a new research field about the epigenetic regulation associated with these DTU biomarkers in the future.

77 The objective of this study was to obtain a comprehensive understanding of how changes in transcript isoform  
78 expression impact universal cellular mechanisms during oocyte maturation. We performed differential transcript usage  
79 (DTU) and differential gene expression (DGE) meta-analyses on oocytes at the GV and MII stages across six  
80 mammalian species, utilizing human orthologous genes. Subsequently, Gene Ontology (GO) term Enrichment analysis  
81 was carried out on the identified DTU and DGE genes to elucidate the role of DTU in oocyte maturation across diverse  
82 mammalian species.

83

## 84 **Materials and Methods**

85

### 86 **Ethical approval**

87 The current study used the public secondary data available on NCBI SRA public site (<https://www.ncbi.nlm.nih.gov/sra>). This study did not involve the use of animals, and no animals or people are identifiable within this  
88 publication.  
89

90

### 91 **Sample datasets**

92 We obtained mRNA sequencing datasets of oocytes at the germinal vesicle (GV) and metaphase II (MII) stages  
93 from six different mammalian species: cow [6], donkey [16], horse [17], human [18], mouse [18], and pig [19] (Table  
94 1). These datasets were from the Sequence Read Archive (SRA) database (<https://www.ncbi.nlm.nih.gov/sra>). For brief  
95 details, the GV and MII oocyte datasets were obtained by isolating oocytes from Cumulus-oocyte complexes (COCs)  
96 and 24-hour *in vitro* matured COCs, respectively. Each species' respective references further documented specific  
97 variations in the protocols used for each species.

98

## 99 **Experimental design**

100 The experiment comprised three sections: 1) Data preparation, 2) Differential gene analyses, and 3) GO term  
101 enrichment analysis (Fig. 1). In the 'Data preparation' section, we pre-processed the mRNA-Seq oocyte datasets,  
102 importing and summarizing them into gene and transcript expression profiles. The 'Differential gene analyses' section  
103 utilized the prepared gene and transcript expression profiles to identify differentially expressed genes and genes with  
104 differential transcript usage through DGE and DTU analyses. Lastly, in the 'GO term enrichment analysis' section, we  
105 conducted GO term enrichment analysis using the obtained DGE and DTU gene lists and examined the relationship  
106 among significant terms using Igraph.

107

### 108 **Section 1 Data preparation**

#### 109 *Data pre-processing*

110 Pre-processing of raw mRNA-Seq oocyte datasets in this study utilized procedures similar to those previously  
111 described, with some modifications [20–22] We employed the 'Flexbar' program (<https://github.com/seqan/flexbar>) to  
112 eliminate any contaminated adapter sequences and disqualified sequences with a mean Phred score lower than 30.  
113 Subsequently, we utilized the 'STAR' software (<https://github.com/alexdobin/STAR>) to align the filtered sequences to  
114 the respective genome sequences corresponding to each species: ARS-UCD1.2 for cow, ASM1607732v2 for donkey,  
115 EquCab3.0 for horse, GRCh38 for human, GRCm39 for mouse, and Sscrofa11.1 for pig. This alignment process  
116 resulted in the generation of aligned BAM format files for each species.

117

#### 118 *Data import and summarization*

119 Using the 'RSEM' software (<https://github.com/deweylab/RSEM>), we quantified the transcript count reads of  
120 the sample datasets corresponding to their BAM files. Importing and summarizing these sample datasets at both the  
121 gene and transcript isoform levels were accomplished through the 'tximport' package  
122 (<https://github.com/thelovelab/tximport>) [8]. This process yielded gene expression libraries and transcript libraries for  
123 each data species. We would utilize the transcript-level expression library of each animal species for further differential  
124 transcript usage (DTU) analysis. To facilitate differential gene expression (DGE) analysis, we concatenate the gene-  
125 level expression libraries among species based on the human orthologous genes with batch correction through the  
126 'ComBatSeq' method (<https://github.com/web-mev/mev-combatseq>) [23].

127

### 128 **Section 2 Differential gene analyses**

129 *Differential gene expression (DGE) analysis*

130 With the cross-species gene-level expression profiles, we performed differential gene expression (DGE)  
131 analysis between the GV and MII oocyte datasets. We used the 'DESeq2' method  
132 (<https://github.com/the-lovelab/DESeq2>) for this analysis, with a significance threshold set at  $FDR \leq 0.01$  and  $\log_2$ -  
133 fold-change  $\geq 1.5$  (Negative binomial regression model with Likelihood ratio tests) [24]. This study used "DGE genes"  
134 to refer to the differentially expressed genes identified through the DGE analysis of each species. For the reliability of  
135 the DGE genes tested across all six mammalian species, we additionally employed the 'QUBIC' biclustering method  
136 (<https://github.com/cran/biclust>) for DGE analysis [25]. This biclustering approach enabled the identification of DGE  
137 genes that exhibited uniform co-expression patterns across oocyte datasets in the same stages across all species. This  
138 study considered the bicluster DGE genes identified as significant DGE genes in all species as "universal DGE genes".

139

140 *Differential transcript usage (DTU) analysis*

141 With the transcript-level expression library of each animal species, this study employed the DTU analysis  
142 protocol as previously guided by Love et al. [10] with some modifications. This study referred to the genes exhibiting  
143 significant differences in their isoform transcript usage as "DTU genes." After filtering out genes with only one  
144 transcript isoform or very low transcript abundance, we obtained the DTU genes between the GV and MII oocyte  
145 datasets in each species using the 'DEXSeq' model (<https://github.com/areyesq89/DEXSeq>) to identify differential  
146 transcript isoform using Negative binomial regression model with Wald test [7]. Subsequently, the 'stageR' package  
147 (<https://github.com/statOmics/stageR>) would concatenate and confirm the significant DTU genes, summarizing their  
148 transcript isoforms that exhibited statistical significance ( $FDR \leq 0.05$ , Kruskal–Wallis test).

149

150 *Virtualization*

151 We utilized the "ComplexHeatmap" package [26] to generate the UpSet plot and the heatmap. By employing a  
152 series of UpSet plots, we visualized the relationships among the DGE genes for each species. This study focused on  
153 DTU genes sharing among a minimum of five distinct species, referred to as "cross-species DTU genes". To exclusively  
154 present the transcript isoform expression levels of these cross-species DTU genes, we employed the 'ggplot2' package to  
155 create bar graphs. For universal DGE genes, we used the heatmap to illustrate the differential expression patterns of  
156 these genes between the GV and MII datasets across the six mammalian species.

157

158 **Section 3 GO term enrichment analysis**

159 *GO term enrichment analysis*

160 We evaluated the functional properties of the obtained universal DGE genes and the DTU genes specific to  
161 each species by conducting gene-annotation enrichment analysis using the 'goseq' package [27]. Significant Gene  
162 Ontology (GO) terms, categorized as Biological Process (BP), Cellular Component (CC), and Molecular Function  
163 (MF), were extracted from the GO consortium database based on available human orthologous genes ( $FDR \leq 0.05$ ).

164

#### 165 *Virtualization*

166 To visualize the network among significant GO terms generated by the universal DGE genes and the DTU  
167 genes specific to each species, we employed the Python "igraph" package [28] to plot the network depicting the  
168 relationships among them.

169

## 170 **Results**

171

### 172 **Relationships among DGE and DTU genes across six mammalian species**

173 As this study primarily focuses on cross-species analysis, the main emphasis of the DGE results lies on the  
174 universal DGE genes. However, we have provided the DGE genes obtained from each species analysis in the  
175 supplementary data. Through the DGE analysis process, we identified 195 universal DGE genes, demonstrating their  
176 differences in expression patterns between the GV and MII datasets across the six mammalian species. The expression  
177 heatmap (Fig 2) reveals that most universal DGE genes exhibit higher expression levels in GV-staged oocytes.

178 In addition, we have included the DTU genes obtained from the analysis of each species in the supplementary  
179 data. The UpSet plots depicted the relationships among the DTU genes across the six mammalian species (Fig 3).  
180 Despite the considerable number of DTU genes identified in each species (blue rectangle in Fig 3), we observed limited  
181 intersections among them. We observed the lowest number of DTU genes in pigs, the species with the fewest  
182 approximate transcript isoform numbers per gene (2.73). Conversely, the species with the highest approximate transcript  
183 isoform numbers per gene, the mouse (6.81), contributed to the highest number of DTU genes (Fig 3). We highlighted  
184 ten DTU genes shared among at least five animal species (red rectangle in Fig 3). These cross-species DTU genes were  
185 as follows: ABCF1, CDC6, CFAP36, CNOT10, DNMT3, IWS1, NBN, NDEL1, RAD50 and ZCCHC17. Table 2  
186 summarized details for these genes, including their complete names and corresponding Ensembl gene IDs for the human  
187 orthologs.

188 The discussion section further elucidated the significance and functions of these genes in detail. Two cross-  
189 species DTU genes—CDC6 and CFAP36 were also present in the universal DGE genes. Such notice implied  
190 remarkable changes in gene expression levels due to shifts in several transcript isoform usage of these genes. As



191 supportive evidence, the high expression levels of several transcript isoforms in GV-staged oocytes were eminent (Fig  
192 4). Regarding the remaining cross-species DTU genes (ABCF1, CNOT10, DNMT3, IWS1, NBN, NDEL1, RAD50, and  
193 ZCCHC17), variations in the expression levels of their respective transcript isoforms between GV-staged and MII-  
194 staged oocytes were evident (Supplementary Fig. 1-8).

195

#### 196 **Relationship among GO terms acquired from universal DGE and DTU genes across six mammalian species**

197 We utilized "Igraph" to visualize the network of significant GO terms produced by universal DGE and DTU  
198 genes for each species. The graphs categorized the terms into the Biological Process (BP) category (Fig 5A), Cellular  
199 Component (CC) category (Fig 5B), and Molecular Function (MF) category (Fig 5C). As this study mainly focused on  
200 cross-species analysis, we considered only those GO terms shared among at least five species as significant terms.

201 Upon observing the overall view of Igraph results, we found significant BP and MF terms mainly associated  
202 with RNA binding, RNA splicing, and protein binding mechanisms, which were strongly related to the cell division  
203 process (Fig 5A and 5C). Consistent with the BP and MF terms, the CC terms indicated the locations of these  
204 mechanisms within the nucleus and cytoplasm compartments (Fig 5B).

205

#### 206 **Discussion**

207 The present study demonstrated the comprehensive procedures for conducting DGE and DTU comparative  
208 analyses of mRNA-Seq datasets obtained from GV-staged and MII-staged oocytes across six mammalian species. In  
209 addition to identifying shared genes among species, we performed term enrichment analysis to elucidate how these  
210 cross-species DTU genes are involved in cell mechanisms responsible for the differences in oocyte nature between the  
211 GV and MII stages. These identified cross-species DTU genes, with their implied roles and functions, offered valuable  
212 insights into knowledge about the universal oocyte maturation process. Furthermore, the significant transcript isoforms  
213 of these cross-species DTU genes, identified across the various species, could introduce novel prognostic and  
214 therapeutic biological transcript isoform markers, thereby advancing assisted reproductive technologies in the future.

215 Despite yielding promising results, certain restrictions could limit the analytical power of this study. The main  
216 issue stemmed from the constrained availability of orthologous genes among the six mammalian species. Including six  
217 species in the analysis thus reduced the number of available genes for cross-species analysis. Furthermore, there are  
218 variations in the number of transcript isoforms per gene among the species, with pigs having the lowest number of  
219 approximate transcript isoforms per gene compared to other species. This disparity in isoform numbers among species  
220 results in some genes having several isoforms in one species but only one in another. Consequently, the inclusion of

221 more species in the comparison resulted in a lower number of DTU genes, which, in turn, contributed to the absence of  
222 DTU genes shared among all species in this study.

223 While the limited sample datasets collected from different species also likely contributed to a low number of  
224 cross-species DTU genes in this study, we expected its impact to be only partial. It is because we observed a significant  
225 influence of transcript isoform redundancy in each mammalian species on the identified DTU gene numbers. This  
226 redundancy was particularly low in cow and pig species, leading to their relatively small number of DTU genes  
227 compared to other species. Therefore, we expected more contribution of evolutionary factors in limiting the cross-  
228 species DTU genes identified in this study.

229 Evolutionary changes, such as mutations and genetic rearrangements, can alter genes, creating or eliminating  
230 alternative splice sites. These changes can lead to the generation of new transcript isoforms or the loss of existing ones  
231 [29]. Additionally, regulatory elements in the genome, such as enhancers and transcription factors, can evolve to  
232 modulate gene splicing patterns. This modulation can affect the inclusion or exclusion of specific exons, leading to  
233 diverse transcript isoforms of the gene [30]. These changes can be driven by natural selection, allowing organisms to  
234 adapt to their specific environments or lifestyles. In the case of oocytes, they might express different transcript isoforms  
235 of DTU genes during the maturation process in a manner unique to each species' physiological condition. In some  
236 species, such as humans and mice, the expression of the numerous repertoire of transcript isoforms may fine-tune  
237 particular gene functions, enabling more precise regulation of the complex maturation process of the species.

238 The presence of cross-species DTU genes thus implied the evolutionary conservation and functional  
239 constraints of the genes [31,32]. These factors suggest that these DTU genes have conserved fundamental functions that  
240 lead to similar transcript usage patterns, thereby limiting changes in their usage. A separate study to identify exon  
241 orthology among transcript isoforms [31–33] for these cross-species DTU genes across mammalian species is required  
242 to support such an idea. It is important to note that such knowledge would be valuable when applied to other closely  
243 related wildlife species, for which background information about their transcript isoform usage is limited.

244 We observed the lowest number of DTU genes in pigs, the species with the lowest approximate transcript  
245 isoform numbers per gene (2.73). Conversely, the species with the highest approximate transcript isoform numbers per  
246 gene, the mouse (6.81), contributed to the highest number of DTU genes (Fig 3). Among the resultant DTU genes, the  
247 plot successfully exhibited ten genes shared across at least five species (Table 2). While we could identify numerous  
248 significant transcript isoforms among these genes (Fig. 4, Supplementary Fig. 1-8), it is not appropriate to directly  
249 compare the transcript isoforms of the same orthologous gene across different species without prior knowledge of  
250 orthology among their exons. This study, therefore, primarily concentrated on the scientific evidence associated with the  
251 cross-species DTU genes, with a limited exploration of their transcript isoforms.

252 Before examining the relationship between these cross-species DTU genes and DGE genes, we identified  
253 cross-species DGE genes through multiple verification steps to obtain the most conserved universal DGE genes. We  
254 provided the supplementary data for DGE genes for those interested in exploring all DGEs across the species, which  
255 was beyond the scope of the current study. Based on this procedure, we identified two cross-species DTU genes—  
256 CDC6 and CFAP36—for which MII oocyte datasets expressed them lower than the GV ones. Changes in transcript  
257 isoform usage of these genes thus had conserved effects on the overall gene expression in MII-stage oocytes among  
258 mammalian species, leading to a decrease in their total expression. The CDC6 protein plays a crucial role in the DNA  
259 replication's initiation during the G1 phase of the cell cycle. Studies in mice [34] and frogs [35] have shown that its  
260 downregulation facilitates the GV to MII transition of oocytes, suggesting that it may play a similar role in other animal  
261 species. Among the cross-species DTU genes, CFAP36 and ZCCHC17 did not exhibit evidence of critical functions in  
262 the GV to MII transition. Such findings hence shed light on their roles' exploration in the future.

263 Notably, the ABCF1 gene encodes an ATP-binding protein implicating mRNA translation, RNA metabolism,  
264 and ribosome biogenesis in oocytes [36–38]. In line with this, the CNOT10 gene encodes a protein involved in mRNA  
265 deadenylation and degradation as part of the CCR4-NOT complex, which is the primary eukaryotic mRNA deadenylase  
266 that regulates meiotic cell cycle progression from the GV to MII stage [39]. Additionally, the DNMT3 gene encodes a  
267 protein belonging to the dynamin family, and we observed the redistribution of the DNMT3 protein to MII spindles  
268 during the transition from GV to MII oocytes [40]. Furthermore, the mRNA processing and export factor IWS1 play a  
269 crucial role in the embryonic genome activation process in mice [41].

270 Moreover, the RAD50 gene, along with the NBN gene, are essential components of the MRN complex  
271 (MRE11-RAD50-NBN), playing critical roles in DNA repair, recombination, and maintenance of genomic stability, as  
272 evidenced in mice [42] human [43], and swine [44]. During the oocyte maturation process from GV to MII stages,  
273 these genes likely play a significant role in safeguarding the integrity of the oocyte's genetic material, protecting it from  
274 DNA double-strand breaks (DSBs) that can occur during meiotic recombination. During this process, the microtubule-  
275 organizing protein encoded by the NDEL1 gene could regulate spindle positioning during this process [45].  
276 Interestingly, our study also identified changes in the isoform usage of the NDEL1 gene, supporting previous evidence  
277 in mice [14]. These collective findings offered valuable insights into the roles of these cross-species DTU genes'  
278 transcripts in regulating mRNA processing and cell-cycle progression across various mammalian species.

279 Supporting the functions of cross-species DTU genes, the related Biological Process (BP) and Molecular  
280 Function (MF) terms obtained from universal DGE and cross-species DTU gene enrichment also contributed to the  
281 control of mRNA processing and cell-cycle regulation. Additionally, the Cellular Component (CC) terms indicated the  
282 locations of these mechanisms within the nucleus and cytoplasm compartments (Fig 5). Remarkably, the cross-species

283 DTU terms identified in this study aligned with those found in a mouse study [14] and were related to GO terms  
284 acquired from DGE genes of several mammalian species in terms of mRNA processing and control [15,16,46,47]. This  
285 notification suggested the complementary roles of DTU and DGE genes in regulating various critical proteins'  
286 transcription and meiotic cell-cycle progression during oocyte maturation.

287

## 288 **Conclusion**

289 This knowledge acquired in this study yielded valuable insights into the involvement of changes in transcript  
290 usage during the oocyte maturation process across mammalian species. For future research, it would be worthwhile to  
291 investigate the entangled relationships between the identified DTU genes and specific cell processes—mRNA  
292 processing and cell-cycle progression in different mammalian species. Such exploration could provide a deeper  
293 understanding of how these DTU genes collectively regulate critical proteins, potentially leading to the discovery of  
294 novel transcript isoform markers for assisted reproductive technologies.

295

## 296 **Conflict of interest**

297 We certify no conflict of interest with any financial organization regarding the material and methods discussed  
298 in the manuscript.

299

## 300 **Acknowledgments**

301 The authors were thankful to the Office of Academic Resource and Information Technology, Rajamangala  
302 University of Technology, Tawan-OK, Thailand, for the contribution of the server computer for the analysis. The  
303 authors did not receive any funds for this study.

304

## 305 **References**

- [1] Awe JP, Byrne JA. Identifying candidate oocyte reprogramming factors using cross-species global transcriptional analysis. *Cell Reprogram* 2013;15:126–33. <https://doi.org/10.1089/cell.2012.0060>.
- [2] Virant-Klun I, Knez K, Tomazevic T, Skutella T. Gene Expression Profiling of Human Oocytes Developed and Matured *In Vivo* or *In Vitro*. *BioMed Research International* 2013;2013:e879489. <https://doi.org/10.1155/2013/879489>.
- [3] He M, Zhang T, Yang Y, Wang C. Mechanisms of Oocyte Maturation and Related Epigenetic Regulation. *Frontiers in Cell and Developmental Biology* 2021;9.

- [4] Pan B, Li J. The art of oocyte meiotic arrest regulation. *Reproductive Biology and Endocrinology* 2019;17:8. <https://doi.org/10.1186/s12958-018-0445-8>.
- [5] Lemseffer Y, Terret M-E, Campillo C, Labrune E. Methods for Assessing Oocyte Quality: A Review of Literature. *Biomedicines* 2022;10:2184. <https://doi.org/10.3390/biomedicines10092184>.
- [6] Reyes JM, Chitwood JL, Ross PJ. RNA-Seq profiling of single bovine oocyte transcript abundance and its modulation by cytoplasmic polyadenylation. *Mol Reprod Dev* 2015;82:103–14. <https://doi.org/10.1002/mrd.22445>.
- [7] Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. *Genome Res* 2012;22:2008–17. <https://doi.org/10.1101/gr.133744.111>.
- [8] Sonesson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences 2016. <https://doi.org/10.12688/f1000research.7563.2>.
- [9] Nowicka M, Robinson MD. DRIMSeq: a Dirichlet-multinomial framework for multivariate count outcomes in genomics. *F1000Res* 2016;5:1356. <https://doi.org/10.12688/f1000research.8900.2>.
- [10] Love MI, Sonesson C, Patro R. Swimming downstream: statistical analysis of differential transcript usage following Salmon quantification 2018. <https://doi.org/10.12688/f1000research.15398.3>.
- [11] THE GTEx CONSORTIUM. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* 2020;369:1318–30. <https://doi.org/10.1126/science.aaz1776>.
- [12] Reyes A, Huber W. Alternative start and termination sites of transcription drive most transcript isoform differences across human tissues. *Nucleic Acids Res* 2018;46:582–92. <https://doi.org/10.1093/nar/gkx1165>.
- [13] RNA mis-splicing in disease | *Nature Reviews Genetics* n.d. <https://www.nature.com/articles/nrg.2015.3> (accessed July 4, 2023).
- [14] He Y, Chen Q, Zhang J, Yu J, Xia M, Wang X. Pervasive 3'-UTR Isoform Switches During Mouse Oocyte Maturation. *Front Mol Biosci* 2021;8:727614. <https://doi.org/10.3389/fmolb.2021.727614>.
- [15] Qian Y, Cao Q, Liao J, Luk CS, Suen AHC, Lee AWT, et al. Simultaneous transcriptome and methylome profiles of single mouse oocytes provide novel insights on maturation and aging 2020:2020.08.22.260612. <https://doi.org/10.1101/2020.08.22.260612>.
- [16] Li Z, Song X, Yin S, Yan J, Lv P, Shan H, et al. Single-Cell RNA-Seq Revealed the Gene Expression Pattern during the In Vitro Maturation of Donkey Oocytes. *Genes* 2021;12:1640. <https://doi.org/10.3390/genes12101640>.
- [17] Angel-Velez D, Meese T, Hedia M, Fernandez-Montoro A, De Coster T, Pascottini OB, et al. Transcriptomics Reveal Molecular Differences in Equine Oocytes Vitrified before and after In Vitro Maturation. *International Journal of Molecular Sciences* 2023;24:6915. <https://doi.org/10.3390/ijms24086915>.

- [18] Hu W, Zeng H, Shi Y, Zhou C, Huang J, Jia L, et al. Single-cell transcriptome and translome dual-omics reveals potential mechanisms of human oocyte maturation. *Nat Commun* 2022;13:5114. <https://doi.org/10.1038/s41467-022-32791-2>.
- [19] Du Z-Q, Liang H, Liu X-M, Liu Y-H, Wang C, Yang C-X. Single cell RNA-seq reveals genes vital to in vitro fertilized embryos and parthenotes in pigs. *Sci Rep* 2021;11:14393. <https://doi.org/10.1038/s41598-021-93904-3>.
- [20] Chokeshaiusaha K, Sananmuang T, Puthier D, Nguyen C. Cross-species analysis of differential transcript usage in humans and chickens with fatty liver disease. *Vet World* 2023;16:1964–73. <https://doi.org/10.14202/vetworld.2023.1964-1973>.
- [21] Kuendee N, Puthier D, Nguyen C, Chokeshaiusaha K. Activation-induced cell death of NK cells in canine atopic dermatitis revealed by Droplet-Sequencing data of peripheral blood mononuclear cells *n.d.*;51:743–51.
- [22] Chokeshaiusaha K, Puthier D, Nguyen C, Sudjaidee P, Sananmuang T. Factor Analysis for Bicluster Acquisition (FABIA) revealed vincristine-sensitive transcript pattern of canine transmissible venereal tumors. *Heliyon* 2019;5:e01558. <https://doi.org/10.1016/j.heliyon.2019.e01558>.
- [23] Zhang Y, Parmigiani G, Johnson WE. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genomics and Bioinformatics* 2020;2. <https://doi.org/10.1093/nargab/lqaa078>.
- [24] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 2014;15. <https://doi.org/10.1186/s13059-014-0550-8>.
- [25] Zhang Y, Xie J, Yang J, Fennell A, Zhang C, Ma Q. QUBIC: a bioconductor package for qualitative biclustering analysis of gene co-expression data. *Bioinformatics* 2017;33:450–2. <https://doi.org/10.1093/bioinformatics/btw635>.
- [26] Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016;32:2847–9. <https://doi.org/10.1093/bioinformatics/btw313>.
- [27] Young MD, Wakefield MJ, Smyth GK. goseq : Gene Ontology testing for RNA-seq datasets Reading data. *Gene* 2010:1–21.
- [28] Csárdi G, Nepusz T. The igraph software package for complex network research. *InterJournal Complex Systems* 2006;1695:1–9. <https://doi.org/10.3724/SP.J.1087.2009.02191>.
- [29] Verta J-P, Jacobs A. The role of alternative splicing in adaptation and evolution. *Trends in Ecology & Evolution* 2022;37:299–308. <https://doi.org/10.1016/j.tree.2021.11.010>.
- [30] Rambout X, Dequiedt F, Maquat LE. Beyond Transcription: Roles of Transcription Factors in Pre-mRNA Splicing. *Chem Rev* 2018;118:4339–64. <https://doi.org/10.1021/acs.chemrev.7b00470>.

- [31] Márquez Y, Mantica F, Cozzuto L, Burguera D, Hermoso-Pulido A, Ponomarenko J, et al. ExOrthist: a tool to infer exon orthologies at any evolutionary distance. *Genome Biology* 2021;22:239. <https://doi.org/10.1186/s13059-021-02441-9>.
- [32] Zea DJ, Laskina S, Baudin A, Richard H, Laine E. Assessing conservation of alternative splicing with evolutionary splicing graphs. *Genome Res* 2021;31:1462–73. <https://doi.org/10.1101/gr.274696.120>.
- [33] Ma J, Wu JY, Zhu L. Detection of orthologous exons and isoforms using EGIO. *Bioinformatics* 2022;38:4474–80. <https://doi.org/10.1093/bioinformatics/btac548>.
- [34] Yi Z-Y, Meng T-G, Ma X-S, Li J, Zhang C-H, Ouyang Y-C, et al. CDC6 regulates both G2/M transition and metaphase-to-anaphase transition during the first meiosis of mouse oocytes. *Journal of Cellular Physiology* 2020;235:5541–54. <https://doi.org/10.1002/jcp.29469>.
- [35] Kinterová V, Kaňka J, Bartková A, Toralová T. SCF Ligases and Their Functions in Oogenesis and Embryogenesis—Summary of the Most Important Findings throughout the Animal Kingdom. *Cells* 2022;11:234. <https://doi.org/10.3390/cells11020234>.
- [36] Tyzack JK, Wang X, Belsham GJ, Proud CG. ABC50 interacts with eukaryotic initiation factor 2 and associates with the ribosome in an ATP-dependent manner. *J Biol Chem* 2000;275:34131–9. <https://doi.org/10.1074/jbc.M002868200>.
- [37] Stewart JD, Cowan JL, Perry LS, Coldwell MJ, Proud CG. ABC50 mutants modify translation start codon selection. *Biochem J* 2015;467:217–29. <https://doi.org/10.1042/BJ20141453>.
- [38] Paytubi S, Wang X, Lam YW, Izquierdo L, Hunter MJ, Jan E, et al. ABC50 promotes translation initiation in mammalian cells. *J Biol Chem* 2009;284:24061–73. <https://doi.org/10.1074/jbc.M109.031625>.
- [39] Sha Q-Q, Yu J-L, Guo J-X, Dai X-X, Jiang J-C, Zhang Y-L, et al. CNOT6L couples the selective degradation of maternal transcripts to meiotic cell cycle progression in mouse oocyte. *EMBO J* 2018;37:e99333. <https://doi.org/10.15252/emj.201899333>.
- [40] Mihalas BP, Redgrove KA, Bernstein IR, Robertson MJ, McCluskey A, Nixon B, et al. Dynamin 2-dependent endocytosis is essential for mouse oocyte development and fertility. *FASEB J* 2020;34:5162–77. <https://doi.org/10.1096/fj.201902184R>.
- [41] Oqani RK, Lin T, Lee JE, Kang JW, Shin HY, Il Jin D. Iws1 and Spt6 Regulate Trimethylation of Histone H3 on Lysine 36 through Akt Signaling and are Essential for Mouse Embryonic Genome Activation. *Sci Rep* 2019;9:3831. <https://doi.org/10.1038/s41598-019-40358-3>.
- [42] Mayer A, Baran V, Sakakibara Y, Brzakova A, Ferencova I, Motlik J, et al. DNA damage response during mouse oocyte maturation. *Cell Cycle* 2016;15:546–58. <https://doi.org/10.1080/15384101.2015.1128592>.

- [43] Suo L, Zhou Y xiao, Jia L ling, Wu H bo, Zheng J, Lyu Q feng, et al. Transcriptome profiling of human oocytes experiencing recurrent total fertilization failure. *Sci Rep* 2018;8:17890. <https://doi.org/10.1038/s41598-018-36275-6>.
- [44] Coutinho ARS, Bordignon V, Coutinho ARS, Bordignon V. 125 SWINE EMBRYOS WITH HIGH POTENTIAL TO DEVELOP IN VITRO HAVE LESS  $\gamma$ H2A.X AND MORE NBS1 PROTEINS. *Reprod Fertil Dev* 2009;22:221–221. <https://doi.org/10.1071/RDv22n1Ab125>.
- [45] Londoño-Vásquez D, Rodríguez-Lukey K, Behura SK, Balboula AZ. Microtubule organizing centers regulate spindle positioning in mouse oocytes. *Dev Cell* 2022;57:197-211.e3. <https://doi.org/10.1016/j.devcel.2021.12.011>.
- [46] Zhao Z-H, Meng T-G, Li A, Schatten H, Yang C, Sun Q-Y. RNA-Seq transcriptome reveals different molecular responses during human and mouse oocyte maturation and fertilization. *BMC Genomics* 2020;21. <https://doi.org/10.1186/s12864-020-06885-4>.
- [47] Zhang F-L, Li W-D, Zhang G, Zhang M, Liu Z-J, Zhu K-X, et al. Identification of unique transcriptomic signatures through integrated multispecies comparative analysis and WGCNA in bovine oocyte development. *BMC Genomics* 2023;24:265. <https://doi.org/10.1186/s12864-023-09362-w>.



308 **Figure legend**

309 **Fig. 1.** Experimental workflow. The workflow encompassed three sections: 1) Data preparation, 2) Differential gene  
310 analyses, and 3) GO term enrichment analysis. Each analytical process was represented by a rounded green square, with  
311 red eclipses denoting the input/output data types utilized. The yellow hexagons showed the final virtualizing step of the  
312 results. The analytical process involved various details, such as species, software, and plot types, all depicted in gray,  
313 rounded squares. The material and methods section contained comprehensive coverage of all analytical procedures.

314

315 **Fig. 2.** Scaled expression pattern of universal DGE genes. The heatmap displayed the expression patterns of  
316 differentially expressed genes between the germinal vesicle (GV) stage and metaphase II (M2) oocytes shared among  
317 six mammalian species (universal DGE genes). To group the genes, we applied hierarchical clustering using Pearson's  
318 correlation distance and the complete linkage method.

319

320 **Fig. 3.** DTU genes among species. The upset plot illustrates the numbers of DTU genes across six mammalian species-  
321 cow, donkey, horse, human, mouse, and pig. It displayed intersections among species of interest through colorized dots  
322 and interconnected lines. For instance, the leftmost column of the plot featured a series of orange connected dots  
323 representing cow, donkey, horse, human, and mouse species, all linked together. It meant that this column would show  
324 the intersected DTU genes acquired among cow, donkey, horse, human, and mouse species. The corresponding bar  
325 graph and the numerical annotation at the column's base depicted the five intersecting DTU genes among the grouped  
326 species. The red rectangle highlighted the DTU genes shared among at least five mammalian species, underscoring their  
327 cross-species significance. On the other hand, the blue rectangle highlighted the cumulative DTU genes of each species.

328

329 **Fig. 4.** Transcript isoform expressions of CDC6 and CFAP36 genes. The bar plots revealed the expression levels of  
330 transcript isoforms for two cross-species DTU genes, CDC6 and CFAP36, which were also in the universal DGE gene  
331 list. The DTU analysis identified the genes significantly differentially expressed in all species except pigs. The blue and  
332 pink bars represented the expression levels of each transcript isoform in GV and MII oocytes, respectively. The  
333 transcript isoforms specific to each mammalian species were in the Ensembl gene ID. The \* symbol noted the transcript  
334 isoforms that exhibited statistical differences.

335

336 **Fig. 5.** Networks of gene ontology (GO) terms of DTU and universal DGE genes. The figure displays the networks of  
337 gene ontology (GO) terms associated with DTU and universal DGE genes. The colorized areas demarcate the  
338 boundaries of GO terms (categorized into BP, CC, and MF terms) generated by each mammalian species and universal

339 DGE genes explained in the legend. Connected dots within each area represent the significant GO terms ( $FDR \leq 0.05$ ).  
340 Notably, only GO terms shared among at least five species are highlighted by colorized dots. The big yellow dots on the  
341 plot signified the GO terms shared among DTU genes of all species and universal DGE genes. Additionally, the big and  
342 small red dots correspond to GO terms shared among DTU genes of all species and those shared among five species,  
343 respectively. The tables adjacent to each graph provided detailed descriptions of the corresponding color-coded dots.