



**HAL**  
open science

## Novel classifier orthologs of bovine and human oocytes matured in different melatonin environments

Thanida Sananmuang, Denis Puthier, Catherine Nguyen, Kaj Chokeshaiusaha

### ► To cite this version:

Thanida Sananmuang, Denis Puthier, Catherine Nguyen, Kaj Chokeshaiusaha. Novel classifier orthologs of bovine and human oocytes matured in different melatonin environments. *Theriogenology*, 2020, 156, pp.82-89. 10.1016/j.theriogenology.2020.06.029 . hal-03015699

**HAL Id: hal-03015699**

**<https://amu.hal.science/hal-03015699>**

Submitted on 20 Nov 2020

**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 **Novel Classifier Orthologs Of Bovine And Human Oocytes Matured In Different Melatonin Environments**

2

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

4

5 **\*Corresponding author: Kaj Chokeshaiusaha**

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

7

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

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

10

11 Running title: Novel classifier orthologs for different melatonin environment of oocytes

12

13 Research funding: Thailand Research Fund (TRF) through New Research Scholar Programme (Grant No.

14 TRG5880003)

15 **Title of the manuscript:** Novel classifier orthologs of bovine and human oocytes matured in different melatonin  
16 environments

17

18 **Abstract**

19 It has been demonstrated that melatonin influences the developmental competence of both *in vivo* and *in*  
20 *vitro* matured oocytes. It modulates oocyte-specific gene expression patterns among mammalian species. Due to  
21 differences among study systems, the identification of the classifier orthologs—the homologous genes related  
22 among mammals that could universally categorize oocytes matured in environments with varied melatonin levels  
23 is still limitedly studied. To gain insight into such orthologs, cross-species transcription profiling meta-analysis  
24 of *in vitro* matured bovine oocytes and *in vivo* matured human oocytes in low and high melatonin environments  
25 was demonstrated in the current study. RNA-Seq data of bovine and human oocytes were retrieved from the  
26 Sequence Read Archive database and pre-processed. The used datasets of bovine oocytes obtained from culturing  
27 in the absence of melatonin and human oocytes from old patients were regarded as oocytes in the low melatonin  
28 environment (Low). Datasets from bovine oocytes cultured in  $10^{-9}$  M melatonin and human oocytes from young  
29 patients were considered as oocytes in the high melatonin environment (High). Candidate orthologs differentially  
30 expressed between Low and High melatonin environments were selected by a linear model, and were further  
31 verified by Zero-inflated regression analysis. Support Vector Machine (SVM) was applied to determine the  
32 potentials of the verified orthologs as classifiers of melatonin environments. According to the acquired results,  
33 linear model analysis identified 284 candidate orthologs differentially expressed between Low and High  
34 melatonin environments. Among them, only 15 candidate orthologs were verified by Zero-inflated regression  
35 analysis ( $FDR \leq 0.05$ ). Utilization of the verified orthologs as classifiers in SVM resulted in the precise  
36 classification of oocyte learning datasets according to their melatonin environments (Misclassification rates  $<$   
37  $0.18$ , area under curves  $> 0.9$ ). In conclusion, the cross-species RNA-Seq meta-analysis to identify novel  
38 classifier orthologs of matured oocytes under different melatonin environments was successfully demonstrated in  
39 this study- delivering candidate orthologs for future studies at biological levels. Such verified orthologs might  
40 provide valuable evidence about melatonin sufficiency in target oocytes- by which, the decision on melatonin  
41 supplementation could be implied.

42 **Keywords:** Classifier orthologs; Melatonin; Oocytes; Cross-species RNA-Seq meta-analysis

## 43 1. Introduction

44 *in vitro* embryo production (IVP) is regarded as a well-established assisted reproductive technology  
45 (ART) in both humans and animal species [1]. *in vitro* oocyte maturation (IVM) is one of the multiple steps of  
46 IVP that the oocytes are recovered directly from the follicles and are cultured in a specific condition. Despite its  
47 advancement, incomparable *in vitro* and *in vivo* environment is still a major problem contributing to the success  
48 of IVM performance [2,3]. During *in vivo* oocyte development, several delicate substances and hormones are  
49 locally secreted in follicular fluid by ovarian and granulosa cells to nourish oocytes. This includes melatonin  
50 which was widely studied for its contribution to oocyte quality [2–5]. Melatonin was most recognized for its  
51 scavenging function on reactive oxygen species (ROS). Since ROS play a major role in impairing oocyte  
52 development [6,7], the lack of melatonin in conventional *in vitro* culture environments has contributed to ROS-  
53 induced poor oocyte quality in several mammals [3,8,9].

54 Several factors also affect *in vivo* oocyte development. This includes reproductive aging—one of the  
55 most common causes of female infertility. It has been reported that reproductive aging induces a variety of  
56 physiological aberrations withering oocyte quality [9–12]. One of them is melatonin deprivation, caused by  
57 reduced melatonin production from aging granulosa and ovarian cells. Melatonin deprivation was hereby blamed  
58 as a major contributor to aging oocyte's incompetency [2,9]. Aging also results in other diverse hormonal  
59 imbalances and assorted physiological changes [9,13]. All these factors represent complex negative effects  
60 during *in vivo* oocyte aging. Therefore, it has been difficult to exclusively determine the melatonin deprivation  
61 effects on the oocyte during *in vivo* development in such circumstances. By means of this, it was difficult to  
62 exclusively determine the melatonin deprivation effect on *in vivo* oocyte in such circumstances.

63 Interestingly, similar aging aberrance also occurred with *in vitro* cultured oocytes in a time-dependent  
64 manner. *in vitro*-aging phenomenon also included ROS-induced quality deterioration due to the absence of  
65 melatonin in the conventional culture environment [2–5,9,11]. In agreement with the protective effects of *in*  
66 *vivo* melatonin produced by ovarian follicles, *in vitro* melatonin supplement was evidenced to restore the oocyte  
67 developmental competency in several species [3–5,9]. Unlike *in vivo*-aging condition divided effects of *in vitro*  
68 melatonin deprivation should be conveniently observed due to the simplicity of the culture environment.  
69 Through this, analysis of crossed effects between *in vitro* and *in vivo* melatonin deprivation due to aging should  
70 hereby contribute a novel insight into the universal melatonin effects independent from other factors.

71 With comparable reproductive physiology and embryogenesis, cattle are regarded as excellent study  
72 models for human IVP [10,14]. Employing this, novel transcriptomic technologies including RNA sequencing  
73 (RNA-Seq) have been continuously applied with oocyte samples obtained from both species for studying their

74 universal transcriptome's regulation. However, the application of standard RNA-Seq or bulk RNA-Seq with  
75 mammalian oocytes was not practical due to low oocyte numbers for cDNA library preparation [15,16]□.  
76 Moreover, oocytes were heterogeneous populations, and should thus not be represented with average results. Since  
77 cell-to-cell variations are common among isolated oocytes, loss of real features among them in bulk analysis  
78 could bring about considerable bias in the analysis [15].

79         Recently, low-input and single-cell RNA sequencing (RNA-Seq) methods have been implemented with  
80 oocytes acquired from various mammals [3,15,17]□. Since such technologies allow robust transcript profiling  
81 analysis down to a single cell, they enable researchers to explore biological variations among oocytes with  
82 reduced bias [3,16,18]□. Interestingly, an increase of such data among a variety of mammalian species even  
83 allowed cross-species meta-analysis of oocytes. Since cross-species transcription profiling meta-analysis of  
84 oocytes could imply universal orthologs regulated among the species in the environment of interest [19],  
85 implement of such analysis with RNA-Seq data of oocytes obtained from different *in vitro* and *in vivo* melatonin  
86 environments should imply us insight into such orthologs, as well.

87         In several transcription profiling studies, genomic classification or subtyping of unknown samples can  
88 be achieved by various machine learning algorithms. Support vector machine (SVM) is a robust learning model  
89 widely implemented on a genome-wide scale to prioritize orthologs for sample classification [20,21]□. Novel  
90 classifier orthologs were identified by SVM in several topics including aging, cancers, infections, and metabolic  
91 disorders [22–25]. Regardless of its flexibility, SVM application in oocyte study was still limitedly demonstrated.  
92 It was noteworthy that the integration of SVM with cross-species meta-analysis of different melatonin  
93 environments could thus help us determine novel classifier orthologs for evaluating the melatonin sufficiency  
94 of oocytes. In other words, such cognition would preliminarily assess the beneficial effects of melatonin  
95 supplement on target oocytes presented in environments of interest.

96         To gain insight into universal classifier orthologs for melatonin environments, cross-species  
97 transcription profiling meta-analysis of *in vitro* matured bovine oocytes and *in vivo* matured human oocytes in  
98 low and high melatonin environments was performed in the current study. The pooled cDNA library based on  
99 human-bovine orthologs was constructed with bias correction. With restricted verification procedures and SVM  
100 classification, the universal classifier orthologs for melatonin environments of oocytes were successfully implied  
101 for the first time in this study.

102

## 103 **2. Materials and methods**

104

## 105 2.1 Sample datasets

106 RNA-Seq data of bovine and human oocytes were retrieved from the sequence read archive (SRA)  
107 database (<https://www.ncbi.nlm.nih.gov/sra>) (Table 1). Both bovine and human RNA-Seq data were presented in  
108 NCBI with accession numbers GSE122738 [3]□ and GSE125300 [26]□ (<https://www.ncbi.nlm.nih.gov/geo/>).  
109 The sources of bovine [3]□ and human oocytes [26]□ utilized for data preparation were summarized in Table 1.  
110 In brief, bovine cumulus- oocyte complexes (COCs) were collected from slaughterhouses and cultured in the  
111 presence (In\_vitro\_MII+Mela) or absence (In\_vitro\_MII) of  $10^{-9}$  M melatonin for 24 hours, respectively.  
112 Metaphase II (MII) oocytes were isolated from COCs and divided into 6 samples—each containing 7 oocytes of  
113 their groups. On the contrary, each human MII oocyte sample was retrieved directly from single follicles of  
114 young (In\_vivo\_young\_MII), or old (In\_vivo\_old\_MII) patients by ultrasound-guided technique. In this study,  
115 the term—melatonin environment was used to infer the amount of melatonin in the environment surrounding the  
116 target oocyte samples. Through this, the bovine oocytes matured in the absence of melatonin supplement, and  
117 human oocytes acquired from old patients were regarded as oocytes in the low melatonin environment. On the  
118 other hand, bovine oocytes matured in the presence of melatonin supplement, and human oocytes acquired from  
119 young patients were considered as oocytes in the high melatonin environment.

120

## 121 2.2 Data pre-processing

122 Data pre-processing was performed with the process similar to that described in our previous study  
123 [27]□. In brief, all sequences were trimmed to remove contaminated adapter sequences and unqualified  
124 sequences (length < 25 nucleotides and mean Phred score < 25). Duplicated sequences removal, sequence  
125 alignments (>85%), and GC base-balance bias correction were accomplished prior to gene-level counting. In this  
126 study, the read counts were normalized by the trimmed mean of M-values (TMM). bovine-human orthologs were  
127 assigned with the previously reported procedure [19]□. Data quality control was carried out [28]□ along with  
128 cell-cycle effect removal to minimize the bias from within-cell-type heterogeneity [29]□. Count read numbers of  
129 all samples were  $> 14 \times 10^6$ , and none of them was prone to apoptosis (Fraction of mitochondrial reads < 1.5%).  
130 The mutual nearest neighbors (MNN) in ortholog expression space was identified for cross-species batch  
131 correction [30]□ of pooled cDNA library created from bovine and human oocyte count data using ‘scater’  
132 package [31]□.

133

## 134 2.3 Candidate ortholog selection for oocyte sample categorization based on melatonin environments

135 Candidate orthologs for oocyte sample categorization by melatonin environment were selected using  
136 procedures as described in our previous study [19]. In brief, candidate bovine-human orthologs differentially  
137 expressed between oocytes in low and high melatonin environments were selected from 1,000 learning datasets  
138 using the linear model and empirical Bayes methods. In this study, we consider differentially expressed orthologs  
139 with importance value  $\geq 2.5$ . Only orthologs declared significant both by intra-species and cross-species analyses  
140 would be selected as candidate orthologs. With expression values of the candidate orthologs, Partitioning around  
141 medoids (PAM) clustering (2-4 clusters) was performed, and clustering validation was determined by  
142 Connectivity, Dunn, and Silhouette scores.

143

#### 144 *2.4 Verification of the candidate orthologs*

145 Concerning about unrealistic mean-variance relationship produced by the batch-corrected data, the  
146 candidate orthologs were verified by Zero-inflated regression analysis with proper model matrix using raw count  
147 data to avoid the bias from the batch correction process with ‘MAST’ package [32]. The same criteria of  
148 candidate ortholog selection was applied—by which the verified candidate orthologs must be differentially  
149 expressed between low and high melatonin environments both by intra-species and cross-species analyses (FDR  
150  $\leq 0.05$ ).

151

#### 152 *2.5 Oocyte sample classification by Support Vector Machine*

153 To determine the potential of the verified candidate orthologs as classifiers for melatonin environments,  
154 Support Vector Machine (SVM) with linear kernel classifier was performed. Different training data were  
155 generated by 5-fold cross-validation, Bootstrap, and Monte-Carlo-cross-validation strategies—1,000 iterations  
156 for each strategy. All procedures could be archived by class prediction functions using the ‘CMA’ package [20].

157

#### 158 *2.6 Data visualization*

159 Heatmap illustrating candidate orthologs was drawn using the “ComplexHeatmap” package [33]. The  
160 silhouette width values acquired from PAM clustering were illustrated by the bar plot drawn by the ‘ggplot2’  
161 packages. A bubble plot manifesting  $-\log_{10}$ FDR values and  $\log_2$ fold-change values of the verified candidate  
162 orthologs was drawn using the ‘ggplot2’ package. Probability plots and Receiver operating characteristic (ROC)  
163 curves were drawn to illustrate the classification accuracy of the verified candidate orthologs by the ‘ggplot2’  
164 package.

165

### 166 3. Results

167

168 *3.1 Oocyte samples could be clustered by their melatonin environments using the candidate orthologs as*  
169 *features.*

170 Analysis of bias-corrected data by linear model revealed 284 candidate orthologs differentially  
171 expressed between oocytes in low and high melatonin environments (importance value  $\geq 2.5$ ). According to PAM  
172 clustering results based on expressions of these candidate orthologs (Table 2 and Fig 1), oocyte samples were  
173 correctly categorized into 2 clusters corresponding to their melatonin environments regardless of the species (Fig  
174 2).

175

176 *3.2 All verified candidate orthologs were significantly expressed in oocytes in the high melatonin environment.*

177 Zero-inflated regression analysis revealed 158 orthologs differentially expressed between oocytes in  
178 low and high melatonin environments ( $FDR \leq 0.05$ ). However, only 15 orthologs were found intersected with the  
179 candidate orthologs acquired from the previous selections—AAMP, ADM, ATP6V0A2, BCL7B, CYB561D2,  
180 ENDOV, F2RL1, KIAA0586, LPCAT4, NR1H3, PLEKHM1, SLC47A1, SYT11, RFT1, and MRPS18C (Fig 3).  
181 While all oocytes in the high melatonin environment expressed these orthologs higher than those in the low  
182 environment, only ADM, CYB561D2, ENDOV, KIAA0586, NR1H3, PLEKHM1, SLC47A1, SYT11, RFT1, and  
183 MRPS18C rendered substantial fold-change values ( $\geq 1.5$ ). According to our review of their biological functions,  
184 such orthologs were associated with a variety of cell metabolisms (Table 3).

185

186 *3.3 Verified candidate orthologs were potential classifiers for the melatonin environments of oocytes.*

187 The verified candidate orthologs were used as classifier orthologs in SVM performance. Consistent  
188 results were acquired among learning datasets—by which all of them rendered Misclassification rates  $< 0.18$ , and  
189 their ROCs had Area Under Curves (AUCs)  $> 0.9$  (Fig 4).

190

### 191 4. Discussion

192 A cross-species transcription profiling meta-analysis of *in vitro* matured bovine oocyte data and *in vivo*  
193 matured human oocyte data to identify classifier orthologs for different melatonin environments was performed  
194 in this study. Despite the discreet procedures manifested, some limitations in result interpretation should be  
195 greatly concerned to avoid over-assumption. As described in the methodology, the *in vitro* and *in vivo* matured  
196 oocyte samples used in this study were procured from different studies with limited information about them. For



197 instance, melatonin concentrations in follicles containing *in vivo* matured human oocytes were not indicated,  
198 while the concentrations were clearly stated in bovine oocytes' culture systems. Undetermined differences in  
199 melatonin concentrations between such environments should reduce the sensitivity of differential analysis in this  
200 study. With more available *in vivo* matured oocyte profiling data with well-defined melatonin environments in  
201 the future, the list of classifier orthologs could be improved for better generalization.

202 As expected, our previously reported procedures [19]□ were also applicable to the selection of  
203 candidate orthologs differentially expressed between low and high melatonin environments (Fig 2 and Table 2).  
204 While these orthologs could be used to annotate essential melatonin functions on *in vivo* and *in vitro* matured  
205 oocytes, this study aimed to identify the most reliable classifier orthologs among them. Due to such delegation,  
206 we further verified the candidate orthologs with the zero-inflated regression test to reduce the mean-variance bias  
207 (Fig 3).

208 Encoded products of the verified candidate orthologs were associated with in various oocyte  
209 metabolisms (Table 3). In detail, ADM encoded an angiogenic factor, adrenomedullin—an important  
210 vasorelaxant required during folliculogenesis and oocyte maturation processes [34,35]. CYB561D2 encoded a  
211 member of trans-membrane (TM) protein family involving stress defense and iron metabolism [36]. ENDOV  
212 encoded Endoribonuclease V—an enzyme involving with the DNA repair process of oocyte [37]□. KIAA0586  
213 encoded a conserved centrosomal protein required for sonic hedgehog/SHH signaling during the oocyte  
214 developmental process [38]. NR1H3-activating ligands helped with oocyte meiosis resumption [39]□.  
215 PLEKHM1 could regulate the autophagy process which was crucial in the regulation of oocyte metabolism [40–  
216 42]□. SLC47A1 encoded an extrusion protein for drug and toxic metabolites [43]□. SYT11 encoded  
217 synaptotagmin-11 for exocytosis and endocytosis regulation [44]□. RFT1 encoded a major folate transporter  
218 [45], which helped to accumulate folate for normal oogenesis and early embryogenesis [46]□. Lastly, MRPS18C  
219 encoded a mitochondrial ribosomal protein and thus was required for other mitochondrial protein syntheses  
220 [47]□.

221 The high melatonin environment was shown to enhance expressions of all verified candidate orthologs  
222 (Fig 3 and Table 3). Supporting this notice, enhanced expression of KIAA0586 by melatonin was already  
223 notified in the other study [8]. Since functions of these orthologs were strongly associated with qualified oocyte  
224 development as previously mentioned, their diminished expressions in the low melatonin environment should  
225 thus adversely affect the oocyte development (Table 3). As far as we know, shreds of evidence of enhanced  
226 functions of most verified orthologs by melatonin were still not warranted in mammalian oocytes—by which  
227 future studies to confirm such effects were encouraged.

228 In the current study, the classifying potential of the verified orthologs was determined by SVM to assess  
229 their reliability and reproducibility. According to the result, low misclassification scores and high AUCs  
230 manifested among tested learning datasets (Fig 4) hereby indicated such orthologs as potential classifiers for  
231 melatonin environments. Though utilization of cross-species data in this study partially suggested such classifier  
232 orthologs to be universally applicable among mammalian oocytes, further confirmation in other mammals with  
233 known melatonin environments were still required. Since overall expression levels of acquired classifier  
234 orthologs could imply sufficiency of melatonin, we, on a certain level proposed them as novel parameters to  
235 provide some clues about beneficial effects of melatonin supplement on target oocytes in any culture  
236 environments of interest. Such knowledge greatly relieved burdens for researchers to screen for universal  
237 melatonin effects on all ortholog expressions which were not possible by the conventional approach. It should  
238 however be noted that these classifier orthologs might not only be specific to the melatonin environment and thus  
239 required further classification tests with other possible oocyte environments.

240

## 241 **5. Conclusion**

242 Novel classifier orthologs of oocytes matured under low and high melatonin environments were  
243 identified by cross-species transcription profiling meta-analysis in this study. Such an outcome not only offered  
244 us future melatonin research topics to evaluate classifying potentials of the verified orthologs at biological levels  
245 but also the complete process to acquire them from available RNA-Seq data. We would like to note that the  
246 analytical procedures demonstrated in the current study could be applied with other oocyte environments of  
247 interest—by which identification of other novel parameters of other oocyte culture environments was possible.

248

## 249 **Conflict of interest**

250 We certify that there is no conflict of interest with any financial organization regarding the material and  
251 methods discussed in the manuscript.

252

## 253 **Acknowledgments**

254 The current study was fully supported by the Thailand Research Fund (TRF) through New Research  
255 Scholar Program (Grant No. TRG5880003). We would like to show our gratitude to the Office of Academic  
256 Resource and Information Technology (ARiT), Rajamangala University of Technology Tawan-OK (RMUTTO)  
257 for the contribution of the server computer for the burden analysis.

258

259 **References**

- 260 [1] Wang J, Sauer M V. *in vitro* fertilization (IVF): A review of 3 decades of clinical innovation and  
261 technological advancement. *Ther Clin Risk Manag* 2006;2:355–64.  
262 <https://doi.org/10.2147/tcrm.2006.2.4.355>.
- 263 [2] Zhang M, Lu Y, Chen Y, Zhang Y, Xiong B. Insufficiency of melatonin in follicular fluid is a reversible  
264 cause for advanced maternal age-related aneuploidy in oocytes. *Redox Biol* 2020;28:101327.  
265 <https://doi.org/10.1016/j.redox.2019.101327>.
- 266 [3] An Q, Peng W, Cheng Y, Lu Z, Zhou C, Zhang Y, et al. Melatonin supplementation during *in vitro*  
267 maturation of oocyte enhances subsequent development of bovine cloned embryos. *J Cell Physiol*  
268 2019;234:17370–81. <https://doi.org/10.1002/jcp.28357>.
- 269 [4] Do LTK, Shibata Y, Taniguchi M, Nii M, Nguyen T V., Tanihara F, et al. Melatonin Supplementation  
270 During *in vitro* Maturation and Development Supports the Development of Porcine Embryos. *Reprod*  
271 *Domest Anim* 2015;50:1054–8. <https://doi.org/10.1111/rda.12607>.
- 272 [5] Soto-Heras S, Catalá MG, Roura M, Menéndez-Blanco I, Piras AR, Izquierdo D, et al. Effects of  
273 melatonin on oocyte developmental competence and the role of melatonin receptor 1 in juvenile goats.  
274 *Reprod Domest Anim* 2019;54:381–90. <https://doi.org/10.1111/rda.13378>.
- 275 [6] El-Raey M, Geshi M, Somfai T, Kaneda M, Hirako M, Abdel-Ghaffar AE, et al. Evidence of melatonin  
276 synthesis in the cumulus oocyte complexes and its role in enhancing oocyte maturation *in vitro* in cattle.  
277 *Mol Reprod Dev* 2011;78:250–62. <https://doi.org/10.1002/mrd.21295>.
- 278 [7] Jing T, Shile S, Sun Y, Li H, Li WP, Cong Z, et al. Melatonin levels in follicular fluid as markers for IVF  
279 outcomes and predicting ovarian reserve. *Reproduction* 2017;153:443–51. [https://doi.org/10.1530/REP-](https://doi.org/10.1530/REP-16-0641)  
280 16-0641.
- 281 [8] Lee S, Jin JX, Taweetchaipaisankul A, Kim GA, Ahn C, Lee BC. Melatonin influences the sonic  
282 hedgehog signaling pathway in porcine cumulus oocyte complexes. *J Pineal Res* 2017;63.  
283 <https://doi.org/10.1111/jpi.12424>.
- 284 [9] Igarashi H, Takahashi T, Nagase S. Oocyte aging underlies female reproductive aging: biological  
285 mechanisms and therapeutic strategies. *Reprod Med Biol* 2015;14:159–69.  
286 <https://doi.org/10.1007/s12522-015-0209-5>.
- 287 [10] Malhi PS, Adams GP, Singh J. A bovine model to study reproductive aging. *Biol Reprod* 2005;73:45–53.  
288 <https://doi.org/10.1095/biolreprod.104.038745>.
- 289 [11] Malhi PS, Adams GP, Mapletoft RJ, Singh J. Oocyte developmental competence in a bovine model of  
290 reproductive aging. *Reproduction* 2007;134:233–9. <https://doi.org/10.1530/REP-07-0021>.
- 291 [12] Thouas GA, Trounson AO, Jones GM. Effect of Female Age on Mouse Oocyte Developmental  
292 Competence Following Mitochondrial Injury1. *Biol Reprod* 2005;73:366–73.  
293 <https://doi.org/10.1095/biolreprod.105.040956>.
- 294 [13] George K, Kamath M. Fertility and age. *J Hum Reprod Sci* 2010;3:121–3. [https://doi.org/10.4103/0974-](https://doi.org/10.4103/0974-1208.74152)  
295 1208.74152.
- 296 [14] Hansen PJ. Realizing the promise of IVF in cattle - An overview. *Theriogenology* 2006;65:119–25.  
297 <https://doi.org/10.1016/j.theriogenology.2005.09.019>.
- 298 [15] Brayboy LM, Wessel GM. The double-edged sword of the mammalian oocyte - advantages, drawbacks  
299 and approaches for basic and clinical analysis at the single cell level. *Mol Hum Reprod* 2015;22:200–7.  
300 <https://doi.org/10.1093/molehr/gav064>.

- 301 [16] Chambers DC, Carew AM, Lukowski SW, Powell JE. Transcriptomics and single-cell RNA-sequencing.  
302 *Respirology* 2019;24:29–36. <https://doi.org/10.1111/resp.13412>.
- 303 [17] Liu XM, Wang YK, Liu YH, Yu XX, Wang PC, Li X, et al. Single-cell transcriptome sequencing reveals  
304 that cell division cycle 5-like protein is essential for porcine oocyte maturation. *J Biol Chem*  
305 2018;293:1767–80. <https://doi.org/10.1074/jbc.M117.809608>.
- 306 [18] Chitwood JL, Burrueal VR, Meyers SA, Ross PJ. 131 RNA-Seq TRANSCRIPTOME PROFILING OF  
307 INDIVIDUAL RHESUS MACAQUE OOCYTES AND PRE-IMPLANTATION EMBRYOS. *Reprod*  
308 *Fertil Dev* 2014;26:179. <https://doi.org/10.1071/rdv26n1ab131>.
- 309 [19] Chokeshaiusaha K, Puthier D, Nguyen C, Sananmuang T. Construction of pooled oocyte expression  
310 profiles of rhesus monkey and mouse for concurrent meta-analyses. *Thai J Vet Med* 2015;45:523–33.
- 311 [20] Slawski M, Daumer M, Boulesteix AL. CMA - A comprehensive Bioconductor package for supervised  
312 classification with high dimensional data. *BMC Bioinformatics* 2008;9. [https://doi.org/10.1186/1471-](https://doi.org/10.1186/1471-2105-9-439)  
313 [2105-9-439](https://doi.org/10.1186/1471-2105-9-439).
- 314 [21] Fernandes de Mello R, Antonelli Ponti M, Fernandes de Mello R, Antonelli Ponti M. Introduction to  
315 Support Vector Machines. *Mach. Learn.*, 2018, p. 163–226. [https://doi.org/10.1007/978-3-319-94989-](https://doi.org/10.1007/978-3-319-94989-5_4)  
316 [5\\_4](https://doi.org/10.1007/978-3-319-94989-5_4).
- 317 [22] Huang S, Nianguang CAI, Penzuti Pacheco P, Narandes S, Wang Y, Wayne XU. Applications of support  
318 vector machine (SVM) learning in cancer genomics. *Cancer Genomics and Proteomics* 2018;15:41–51.  
319 <https://doi.org/10.21873/cgp.20063>.
- 320 [23] Wang HY, Chang SC, Lin WY, Chen CH, Chiang SH, Huang KY, et al. Machine learning-based method  
321 for obesity risk evaluation using single-nucleotide polymorphisms derived from next-generation  
322 sequencing. *J Comput Biol* 2018;25:1347–60. <https://doi.org/10.1089/cmb.2018.0002>.
- 323 [24] Kumar A, Jeya Sundara Sharmila D, Singh S. SVMRFE based approach for prediction of most  
324 discriminatory gene target for type II diabetes. *Genomics Data* 2017;12:28–37.  
325 <https://doi.org/10.1016/j.gdata.2017.02.008>.
- 326 [25] Arabfard M, Ohadi M, Rezaei Tabar V, Delbari A, Kavousi K. Genome-wide prediction and  
327 prioritization of human aging genes by data fusion: A machine learning approach. *BMC Genomics*  
328 2019;20. <https://doi.org/10.1186/s12864-019-6140-0>.
- 329 [26] Zhang JJ, Liu X, Chen L, Zhang S, Zhang X, Hao C, et al. Advanced maternal age alters expression of  
330 maternal effect genes that are essential for human oocyte quality. *Aging (Albany NY)* 2020.  
331 <https://doi.org/10.18632/aging.102864>.
- 332 [27] Chokeshaiusaha K, Puthier D, Nguyen C, Sudjaidee P, Sananmuang T. Factor Analysis for Biclust  
333 Acquisition (FABIA)revealed vincristine-sensitive transcript pattern of canine transmissible venereal  
334 tumors. *Heliyon* 2019;5:e01558. <https://doi.org/10.1016/j.heliyon.2019.e01558>.
- 335 [28] Luecken MD, Theis FJ. Current best practices in single- cell RNA- seq analysis: a tutorial. *Mol Syst*  
336 *Biol* 2019;15. <https://doi.org/10.15252/msb.20188746>.
- 337 [29] Barron M, Li J. Identifying and removing the cell-cycle effect from single-cell RNA-Sequencing data.  
338 *Sci Rep* 2016;6. <https://doi.org/10.1038/srep33892>.
- 339 [30] Shafer MER. Cross-Species Analysis of Single-Cell Transcriptomic Data. *Front Cell Dev Biol* 2019;7.  
340 <https://doi.org/10.3389/fcell.2019.00175>.
- 341 [31] Haghverdi L, Lun ATL, Morgan MD, Marioni JC. Batch effects in single-cell RNA-sequencing data are  
342 corrected by matching mutual nearest neighbors. *Nat Biotechnol* 2018;36:421–7.  
343 <https://doi.org/10.1038/nbt.4091>.

- 344 [32] Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, et al. MAST: A flexible statistical  
345 framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA  
346 sequencing data. *Genome Biol* 2015;16. <https://doi.org/10.1186/s13059-015-0844-5>.
- 347 [33] Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional  
348 genomic data. *Bioinformatics* 2016;32:2847–9. <https://doi.org/10.1093/bioinformatics/btw313>.
- 349 [34] Malamitsi-Puchner A, Sarandakou A, Baka SG, Tziotis J, Rizos D, Hassiakos D, et al. Concentrations of  
350 angiogenic factors in follicular fluid and oocyte-cumulus complex culture medium from women  
351 undergoing *in vitro* fertilization: Association with oocyte maturity and fertilization. *Fertil Steril*  
352 2001;76:98–101. [https://doi.org/10.1016/S0015-0282\(01\)01854-4](https://doi.org/10.1016/S0015-0282(01)01854-4).
- 353 [35] Manau D, Balasch J, Jiménez W, Fábregues F, Civico S, Casamitjana R, et al. Follicular fluid  
354 concentrations of adrenomedullin, vascular endothelial growth factor and nitric oxide in IVF cycles:  
355 Relationship to ovarian response. *Hum Reprod* 2000;15:1295–9.  
356 <https://doi.org/10.1093/humrep/15.6.1295>.
- 357 [36] Asard H, Barbaro R, Trost P, Bérczi A. Cytochromes b561: Ascorbate-mediated trans-membrane electron  
358 transport. *Antioxidants Redox Signal* 2013;19:1026–35. <https://doi.org/10.1089/ars.2012.5065>.
- 359 [37] Cao W. Endonuclease V: An unusual enzyme for repair of DNA deamination. *Cell Mol Life Sci*  
360 2013;70:3145–56. <https://doi.org/10.1007/s00018-012-1222-z>.
- 361 [38] Lee S, Jin JX, Taweetchaipaisankul A, Kim GA, Ahn C, Lee BC. Sonic hedgehog signaling mediates  
362 resveratrol to improve maturation of pig oocytes *in vitro* and subsequent preimplantation embryo  
363 development. *J Cell Physiol* 2018;233:5023–33. <https://doi.org/10.1002/jcp.26367>.
- 364 [39] Steffensen KR, Robertson K, Gustafsson JÅ, Andersen CY. Reduced fertility and inability of oocytes to  
365 resume meiosis in mice deficient of the Lxr genes. *Mol Cell Endocrinol* 2006;256:9–16.  
366 <https://doi.org/10.1016/j.mce.2006.03.044>.
- 367 [40] Sagrillo-Fagundes L, Bienvenue-Pariseault J, Vaillancourt C. Melatonin: The smart molecule that  
368 differentially modulates autophagy in tumor and normal placental cells. *PLoS One* 2019;14.  
369 <https://doi.org/10.1371/journal.pone.0202458>.
- 370 [41] Lin FH, Zhang WL, Li H, Tian XD, Zhang J, Li X, et al. Role of autophagy in modulating post-  
371 maturation aging of mouse oocytes article. *Cell Death Dis* 2018;9. <https://doi.org/10.1038/s41419-018-0368-5>.
- 373 [42] Shen XH, Jin YX, Liang S, Kwon JW, Zhu JW, Lei L, et al. Autophagy is required for proper meiosis of  
374 porcine oocytes maturing *in vitro*. *Sci Rep* 2018;8. <https://doi.org/10.1038/s41598-018-29872-y>.
- 375 [43] Yonezawa A, Inui KI. Importance of the multidrug and toxin extrusion MATE/SLC47A family to  
376 pharmacokinetics, pharmacodynamics/toxicodynamics and pharmacogenomics. *Br J Pharmacol*  
377 2011;164:1817–25. <https://doi.org/10.1111/j.1476-5381.2011.01394.x>.
- 378 [44] Wang C, Wang Y, Hu M, Chai Z, Wu Q, Huang R, et al. Synaptotagmin- 11 inhibits clathrin- mediated  
379 and bulk endocytosis. *EMBO Rep* 2016;17:47–63. <https://doi.org/10.15252/embr.201540689>.
- 380 [45] Umopathy NS, Gnana-Prakasam JP, Martin PM, Mysona B, Dun Y, Smith SB, et al. Cloning and  
381 functional characterization of the proton-coupled electrogenic folate transporter and analysis of its  
382 expression in retinal cell types. *Investig Ophthalmol Vis Sci* 2007;48:5299–305.  
383 <https://doi.org/10.1167/iovs.07-0288>.
- 384 [46] Meredith M, MacNeil AH, Trasler JM, Baltz JM. Growing Mouse Oocytes Transiently Activate Folate  
385 Transport via Folate Receptors As They Approach Full Size1. *Biol Reprod* 2016;94.  
386 <https://doi.org/10.1095/biolreprod.115.137687>.

- 387 [47] Van Blerkom J. Mitochondrial function in the human oocyte and embryo and their role in developmental  
388 competence. *Mitochondrion* 2011;11:797–813. <https://doi.org/10.1016/j.mito.2010.09.012>.  
389

390 **Figure 1.** Silhouette plot of PAM clustering. The average silhouette width was 0.395 when dividing oocyte  
391 samples into 2 clusters.

392  
393 **Figure 2.** Heatmap of PAM clustering results. Rows of heatmap represented scaled expression values of the  
394 candidate orthologs clustered by k-means clustering, while the columns represented oocyte samples clustered by  
395 PAM.

396  
397 **Figure 3.** Bubble plot of the verified candidate orthologs. The blue and red colors indicated scaled expression  
398 levels of target orthologs expressed by oocytes in the low melatonin environment comparing to those obtained  
399 from the high melatonin environment. The color intensities were corresponded to the  $-\log_{10}\text{FDR}$  values, while  
400 the dot sizes were proportional to the absolute  $\log_2\text{fold-change}$  values of differentially expressed orthologs.

401  
402 **Figure 4.** Predicted probability plots and ROC curves. Predicted probabilities plots were displayed by colored  
403 dots, where red dots corresponded to class 0—low melatonin environment, and green dots for class 1—high  
404 melatonin environment. The misclassification rate acquired from each learning dataset was shown under the plot  
405 (5-fold cross-validation, Bootstrap, and Monte-Carlo-cross-validation). Adjacent to each predicted probability  
406 plot, the ROC curve generated from its same learning dataset was drawn with AUC value demonstrated.