The impact of age on single Oocytes and their cumulus cells Using sc- RNA sequencing

Arwa Abdulaziz Alageel

Linacre college

University of Oxford

A thesis submitted for the degree of Doctor of Philosophy at the University of Oxford.

Dedication

To my deeply missed mom, Latifa, who sadly left us more than a decade ago, I remember you as this perfect human being who was, to me, the greatest mother who ever walked on earth.

Mom, you spirit has accompanied me throughout my entire journey since I left the country until my last day in the UK.

Your beautiful eyes watched over me in my moments of weakness and sadness, and I felt your hand gently wipe my tears and hold me tight.

I felt your happiness and pride whenever I succeeded year after year until I reached the end of my academic journey.

My wound from losing you will never be healed; even the pain will not be bearable over time. What is most painful for me is that you never witnessed how I earned any of my degrees, from bachelor to PhD now. For this reason, I believe that my only consolation is to name my thesis after you, 'Latifa'.

Without a doubt, a PhD degree from the University Oxford in the field of women's infertility is the most valuable thing that can bring honour to your name.

My PhD thesis is dedicated to the memories of you, until we meet again in heaven, Mom.

To my father, thank you for giving me strength to reach for the stars and chase my dreams.

To my family, the symbol of love and giving, particularly my dearest sister, Layla, who is a soul therapist, the one who leads me during my low moments with her light of hope and support during my academic journey

To my great brother, Omar, the best gift I was given in life, my lifeline who stands by me when things look bleak and, with his heavenly presence, has provided me the mental strength to complete this degree

To my beloved sister, Omiama, the angle who lights my path, the best friend God ever made for me and who gives of herself in countless ways.

And to all those people who touched my heart, making London a beautiful and memorable place for me, I humbly dedicate this research to you with all my love.

Acknowledgment

It is a pleasure to thank the many people who have helped made this thesis possible. First, I would like to express my immense gratitude to my supervisor, Dr Suzannah Williams, for the opportunity to undertake this DPhil. Her guidance and encouragement are a source of my inspiration and passion for scientific research. I likewise express my deep appreciation to my co-supervisor, Professor Dagan wells, for his support and guidance in various ways throughout the course of my DPhil.

It was a great pleasure to work with a distinguished number of people (current and former), whose contributions, discussions and pieces of advice have made the process of completing this thesis much easier. My sincerest thanks go to Dr Belinda lo, my beloved colleague who has always been there for me whenever I needed help.

I also gratefully acknowledge the expertise of Dr Thomas Tapmeier. I appreciate his comments, explanations and crucial pieces of advice. My enormous thanks likewise go to Dr Karl Morten for his precious insights that were helpful in solving technical issues.

I thank the Weatherall Institute of Molecular Medicine. In particular: Dr Neil Ashely and his Single Cell Facility where I have done the first part of my research. I particularly would like to thank Dr Emmanouela Repapi and Dr Nicki Gray from Computational Group.

My thanks go to the sequencing group at the Wellcome Trust Centre for Human Genetics at University of Oxford.

My sincerest gratitude goes to my thesis evaluation committee, which consisted of Professor Christian Becker and Professor Tom Fleming who served as the external examiner of my thesis.
Finally, I express my enormous and special gratitude to Professor Krina Zondervan for her unceasing support and selfless contributions that helped me get me where I am today.

I hope that this thesis is useful and serves a good purpose in the field of women's health. I thank everyone who contributed to the success of my DPhil at Nuffield Department of women's and reproductive and Health.

Table of Content

Abstract……..….xiii
Abbreviation……..…… ……..xii
Chapter 1: Literature Review……xvii
1. Introduction 1
1.1. The function of the ovaries 1
1.1.1. Follicle development 1
1.1.2. Oogenesis 4
1.1.3. Oocyte maturation 6
1.1.4. Gonadotropins and steroidogenesis 7
1.1.5. Oocyte metabolism 8
1.1.5.1. Glucose metabolism 8
1.1.5.2. Lipid metabolism 8
1.2. Age and oocyte competence 9
1.2.1. The clinical indication of pre-pubertal oocytes 9
1.2.2. Previous studies on pre-pubertal oocyte quality 10
1.2.3. Transcriptome studies on pre-pubertal oocytes 12
1.2.4. Advanced maternal age 13
1.2.5. Oocyte ageing studies 17
1.3. Cumulus cells 20
1.3.1. The origin of CCs 21
1.3.2. The role of CCs in follicle development 23
1.3.3. The role of CCs in oocyte meiosis 23
1.3.4. The influence of CCs on oocyte competence studies 24
1.4. Overview: Historic development of transcriptomics technologies 27
1.4.1. Hybridisation based method: Microarray methods 27
1.4.2. Sequence-based methods 28
1.4.3. Next-generation sequencing 29
1.5. Single-cell Transcriptomics 30
1.5.1. Single-Cell (sc) RNA Sequencing 30
1.6. Reproductive lifespan in mice 34
1.6.1. Age category selection 34
1.7. Hypothesis 37
1.8. Aims 38
Chapter 2: Material and Method
2.1. Strain selection 41
2.2. Mice maintenance 42
2.3. Ovary stimulation and collection 43
2.4. Cumulus oocytes complex (COC) preparation 44
2.4.1. Oocyte isolation……..45
2.4.2. Cumulus cells preparation……..…….45
2.5. Smart-Seq2 Approach 47
2.5.1. Cell lysis buffer set-up 49
2.5.2. Reverse transcription (RT) 49
2.6. PCR-based Amplification 51
2.6.1. Primer sequences 51

2.6.2. Amplification protocol 51
2.6.3. PCR clean-up 52
2.7. Quality control 52
2.7.1. Qubit® 2.0 flourometric assay (ThermoFisher, UK) 52
2.7.2. 2100 Bioanalyser assessment (Agilent Technology, UK). 53
2.7.3. PicoGreen measurement 54
2.7.4. Assay set-up 55
2.7.5. TapeStation assay (Agilent, UK) 55
2.8. Sequencing Library 56
2.8.1. Nextera XT construction 56
2.8.2. Pooling and multiplexing 57
2.9. Bioinformatics analysis 58
2.9.1. Sample analysis 58
2.9.2. Normalisation method 58
2.9.3. Differential gene expression 59
2.10. Bioinformatics visualisation method 61
2.10.1. Principle component analysis (PCA) 61
2.10.2. Functional analysis: Ingenuity Pathways analysis (IPA) 61
2.10.3. Functional analysis: Metacore pathways .....**Error! Bookmark not defined.**
2.11. OpenArray real-time PCR validation 62
Chapter 3: Optimising a Preparation Protocol for Single -Cell RNA sequencing Library……...….62
3.1. Introduction 64
3.2. Method 65
3.2.1. Oocyte and CCs Preparation 65
3.2.2. PCR-based Amplification 66
3.2.3. cDNA Assessments 66
3.3. PCR Amplification protocol 67
3.3.1 Pilot 1: cDNA yield generated from a single oocyte and its companion CCs 67
3.3.2. Aim 67
3.3.3. Method 67
3.3.4. Pilot 1 Result 68
3.3.4.1. The Optimal amplification for single mouse 0ocytes 68
3.3.4.2. The Optimal Amplification for the CCs 71
3.4.2. Pilot 2: cDNA Yield of Aged Oocytes and Cumulus Cells 75
3.4.2.1. Aim 75
3.4.2.2 Method 75
3.4.3. Pilot 2 result 76
3.4.3.1. Amplification result of old oocytes 76
3.4.3.2. Cumulus cells amplification 78
3.5. Pilot 3: different cumulus Cell preparation methods 80
3.5.1. Aim 80
3.5.2. Method 80
3.5.3. Results of pilot 3 81
3.6. Pilot 4: The effect of the collection time on RNA integrity 83
3.6.1. Aim 83
3.6.2. Method 83
3.6.3. Pilot 4 Results 84

3.7. Discussion 88
3.8. Summary .....**Error! Bookmark not defined.**
Chapter 4: Experimental Design……..97
4.1. Introduction 99
4.2. Method 99
4.2.1. cDNA synthesis 99
4.2.2. Samples number..... 100
4.2.3. Sample selection criteria 100
4.3. Results: cDNA measurement 104
4.3.1. Nine-week-old (young adult) samples 105
4.3.2. Three-week-old (Pre-pubertal)..... 106
4.4. Sample selected for the subsequent step: Nextera XT library 107
4.4.1. cDNA quality assessment 107
4.5. Final quality assessment..... 109
4.5.1. TapeStation profiles for 1-year-old oocytes 109
4.5.2. TapeStation profiles for 9-weeks-old oocytes 111
4.5.3. TapeStation profiles for 3-weeks-old oocytes 113
4.5.4. TapeStation profiles for 1-year--old CCs 115
4.5.5. TapeStation profiles for 9-weeks-old CCs..... 117
4.5.6. TapeStation profiles for 3-weeks-old CCs..... 119
4.6. Discussion 121
4.7. Summary…….
Chapter 5: Tanscriptomic Analysis of Pre-Pubertal Oocytes….....……122
5.1. Introduction 124
5.2. Method 126
5.2.1. Sample preparation 126
5.2.3. Functional analysis..... 127
5.3. Results 127
5.3.1. Sequencing quality..... 128
5.3.1.1. Mapping rate 128
5.3.1.2. Oocyte-specific genes expression 130
5.3.1.3. Number of expressed genes (quality assessment)..... 133
5.3.2. Principle Component Principle (PCA)..... 134
5.3.2.1. Pre-pubertal versus adult oocytes (3 weeks vs 9weeks) 136
5.3.3. Differential genes expression..... 137
5.3.4. Overlapping genes between the three ages 139
5.3.5. Functional analysis..... 140
5.3.5.1 . Canonical Pathways 140
5.3.6. Upstream regulators 152
5.3.7. Networks identification..... 154
5.4. Discussion 161
5.5. Conclusion……170
Chapter 6: Transcriptome Analysis of Aged Oocyte……...……170
6.1. Introduction..... 172
6.2. Method 174
6.3. Results..... 174
6.3.1. Principle Component Analysis 174
6.3.2. Differential Gene Expression..... 177

6.3.3. Functional Analysis 178
6.3.2. Network Identification 194
6.4. Discussion 199
6.5. Conclusion 209
Chapter 7: Transcriptomic Analysis of Cumuluc cells deriverd from single oocytes……...…207
7.1. Introduction 211
7.2. Methods 213
7.3. Results 214
7.3.1. Sequencing Quality 214
7.3.1.1 Mapping Rate 214
7.3.1.2. Cumulus cell-specific gene expression 215
7.3.2. Principal component analysis 221
7.3.2.1. Pre-pubertal versus adult cumulus cells (3 weeks versus 9 weeks) 223
7.3.2.2. Aged versus young-adult cumulus cells (1 year versus 9 weeks) 223
7.3.3. Differential gene expression 226
7.3.3.1 Pre-pubertal versus young-adult cumulus cells 226
7.3.3.2 1-year-old versus 9-week-old cumulus cells 226
7.3.4. Functional network analysis 228
7.4. Discussion 236
Chapter 8: OpenArray Real-Time PCRValidation 246
8.2.1. Pilot 1: Assessment of RNA recovered from a single oocyte and its CCs ... 249
8.2.2. Pilot 2: Real-time PCR of COC 255
8.2.3. Pilot 3: Gapdh and Cyp11a1 expression in a single oocyte and its CCs 257
8.3. OpenArray RT-PCR 262
8.3.2. Experimental design 267
8.3.2.3. Sample preparation 268
8.3.3. Preamplification 268
8.3.4. Real-time PCR preparation 269
8.3.5. Real-time PCR workflow 269
8.3.6. Preliminary experiment 270
8.3.6.1. Aim 270
8.3.6.2. Method 270
8.3.6.3. Results 271
8.4. OpenArray RT-PCR experiment 271
8.4.1. Sample loading 271
8.4.2. Real-time PCR instrument 272
8.5. Data analysis using the comparative Ct method 272
8.6. OpenArray real-time qPCR result 273
8.6.2. Three-week-old oocytes 276
8.6.3. One-year-old oocytes 280
8.6.4. One-year-old CCs versus 9-week-old CCs 283
8.7. Discussion 286
8.8. Conclusion 295
Chapter 9: General Discussion……**……..……..294**
References……304

Index of Figures

Figure 1.1. Follicle development stages from the primordial to preovulatory stage. 3
Figure 1.2 An oocyte at the germinal vesicle (GV) stage. 5
Figure 1.3 Quantitative and qualitative declines of the ovarian follicle pool. 15
Figure 1.4 Birth rate related with maternal age in the United Kingdom,. 16
Figure 1.5 Diagram of an ovarian preovulatory follicle. 22
Figure 1.6 Schematic of sc RNA sequencing experiment showing each step 33
Figure 1.7 The reproductive lifespan of C57BL/6 female mice 36
Figure 2.1 Single oocyte preparation steps 46
Figure 2.2 Smart-Seq2 protocol, Nature protocol 48
Figure 2.3 High-sensitivity DNA nano chip, Agilent Technologies 54
Figure 2.4 High-sensitivity tape screen, Agilent Technologies 56
Figure 2.5 Illustration of bioinformatics steps 60
Figure 3.1 Flowchart of CCs preparation steps……...65
Figure 3.2 Bioanalyser illustrating the amplification profiles of four single oocytess. 69
Figure 3.3 Electropherogram of oocyte 6 amplified with 18 cycles. ……70
Figure 3.4 Bioanalyser electropherograms of the pre-amplified cDNA of the CCs. 73
Figure 3.5 Bioanalyser analysis of the cDNA generated by the cumulus cells 74
Figure 3.6 Amplification profiles of (A) one-year oocyte and (B) nine weeks oocyte.…76
Figure 3.7 Representative examples of samples of cumulus cell amplification……..…78
Figure 3.8 The cDNA yields from the cumulus cells assessed via the 2100 Agilent Bioanalyser. 86
Figure 3.9 The cDNA yields from the CCs……..……..85
Figure 3.10 The cDNA yields from the CCs……..
Figure 3.11 Template switching mechanism (TSO) used in Smart-Seq2 protocol.. 92
Figure 4.1 Experimental design for sc-RNA-seq. 102
Figure 4.2 Sequencing design for sc-RNA experiment. 103
Figure 4.3 (A) TapeStation Profiles for 1-year-old. 109
Figure 4.3 (B) TapeStation Profiles for 1-year-old oocytes:……... 110
Figure 4.4 (A) TapeStation Profiles for 9-weeks-old 111
Figure 4.4 (B) TapeStation Profiles for 9-weeks-old. 112
Figure 4.5 (A) TapeStation Profiles for 3-weeks-old.. 113
Figure 4.6 (A) TapeStation Profiles for 1-year-old Cumulus cells. ……115
Figure 4.6 (B) TapeStation Profiles for 1-year-old Cumulus cells. 116
Figure 4.7 (A) TapeStation Profiles for 9-weeks-old CCs. 117
Figure 4.7 (B) TapeStation Profiles for 9-weeks-old CCs. 118
Figure 4.8 (A) TapeStation Profiles for 3-weeks-old CCs. 119
Figure 4.8 (B) TapeStation Profiles for 9-weeks-old CCs. 120
Figure 5.1 Mapping rate graph. 129
Figure 5.2 depicts a boxplot demonstrating the average expression of the gene-specific oocyte in the sequencing data.. 132
Figure 5.3 depicts a boxplot of the count per million (cpm) 133
Figure 5.4 depicts the PCA of a mouse sc-RNA-seq transcriptome. 135
Figure 5.5 depicts the PCA of oocytes at 3 weeks of age versus 9 weeks 136
Figure 5.6 depicts a scatter plot of average counts (cpm) versus Log2 -fold changes for genes generated by DEseq. 138

Figure 5.7 depicts a Venn diagram showing the number of genes differently regulated between the two comparisons. 139
Figure 5.8 Stacked bar chart of the top 20 canonical pathways in a pre-pubertal oocyte 149
Figure 5.9 The 20 top significant pathways activated in a pre-pubertal oocyte 150
Figure 5.10 Overlapping of the top ten canonical pathways generated by IPA.. 151
Figure 5.11 TGFβ family is predicted as the upregulator effect in pre-pubertal oocyte. . 153
Figure 5.12 Network 1: Development Disorder, organismal injury and abnormalities.... 156
Figure 5.13 Network 2: Cellular assembly and organisation, cardiovascular system 157
Figure 5.14 Network 3: Inflammatory Disease, skeletal and muscular disorders, ... 158
Figure 5.15 Network 4: Cell cycle, dental disease, developmental disorder. 159
Figure 5.16 Network 5: Cell death and survival, cellular assembly and organisation,..... 160
Figure 6.1 Principle component analysis of oocytes at one year of age (aged) versus nine weeks of age (young adult). 176
Figure 6.2 Scatter plot of average counts (cpm) versus Log2 -fold changes for genes generated by DEseq. 177
Figure 6.3 Enriched significant pathways in an aged oocyte……180
Figure 6.4 A stacked bar chart of the top enriched canonical pathways in an aged oocyte ……...181
Figure 6.5 Overlapping of the top 10 canonical pathways generated by IPA 193
Figure 6.6 The top network affected in 1-year oocyte: Embryonic Development, Organismal Development 195
Figure 6.7 Network four: Organismal injury and abnormalities..... 196
Figure 6.8 Network 3: Cell Death and Survival, Cancer. 197
Figure 6.9 Network 4: Lipid metabolism, small molecule Biochemistry. 198
Figure 7.5 Mapping rate graph for CC samples.....**Error! Bookmark not defined.**
Figure 7.2 Boxplot demonstrating the average expression of the CC-specific genes in the sequencing data. 219
Figure 7.3 A boxplot of the number of genes expressed in the CCs from each age group 220
Figure 7.4 Principle component analysis (PCA) plot based on the 500 most variable genes 222
Figure 7.5 Principle component analysis (PCA) of the CCs 224
Figure 7.6 Principle component analysis (PCA) of oocytes at 1 year versus 9 weeks. 225
Figure 7.7 Scatterplot of average counts (cpm) versus log2-fold changes for genes generated by DEseq. Genes were retained if 1 count-per-million (cpm for any given sample. 227
Figure 7.8 The top IPA-generated network enriched in the differentially expressed CC. 234
Figure 7.9: The second network affected in the aged CCs 235
Figure 8.1 Pilot: 1: Gapdh and Cyp11a1 expressions in single oocytes and CCs. 254
Figure 8.2. Pilot 2: Gapdh and Cyp11a1 expression in mouse COC samples. 256
Figure 8.3 Workflow of the RT-qPCR, including the sample preparation, reverse 259
Figure 8.4 Pilot 3: Successful RT-PCR 261
Figure 8.5 A) Photograph of a stainless-steel plate of OpenArray qPCR 265
B) Schematic illustration of the holes showing 224 assays organised in four subarrays 265
Figure 8.6 Illustration of the first column in an OpenArray plate 266
Figure 8.7 Ct value analysis of five candidate housekeeping genes. 274
Figure 8.8 Variation analysis of Ct values of five housekeeping genes in CCs samples . 275
Figure 8.9 Oocyte-specific genes were detected in 3-week-old oocytes 277

Figure 8.10 Relative expression of the upregulated genes in 3-week-old oocytes determined by OpenArray qPCR.. 278
Figure 8.11 Comparison of sc-RNA-seq (green) and OpenArray qPCR (blue) data. 279
Figure 8.12 Oocyte-specific genes were detected in 1-year-old oocytes..... 281
Figure 8.13 Comparison between the expression of upregulated genes generated by sc-RNA-seq and OpenArray qPCR. 282
Figure 8.14 Cumulus cells-specific genes were detected in aged CCs using OpenArray qPCR 284
Figure 9.1: Common pathways between the prepubertal and aged oocytes relative to young adult oocytes. 302

Index of Tables

Table 1.1: Microarrays studies for comparisons between young and aged oocytes. 20
Table 2.1: Mice strains used for ageing studies 42
Table 2.2: Reagents used for lysis buffer preparation 49
Table 2.3: Reverse transcription mix 50
Table 2.4: Reagents used for PCR mix reaction 51
Table 3.1: Summary of the Cumulus Cell Preparation Methods 81
Table 3.2: cDNA Concentration from the cumulus cells samples 85
Table 4.1: Number of oocyte cumulus complexes (COCs) ……97
Table 4.2: cDNA concentration of 1-year-old oocytes and CCs…….. 104
Table 4.3: cDNA concentrations of 9-week oocytes and CCs..... 105
Table 4.5: Final selection of cDNA samples of oocytes..... 108
Table 5.1: The top 20 enriched pathways in pre-pubertal oocytes 142
Table 5.2 demonstrates genes involved in signaling by Rho of GTPases 147
Table 5.3 demonstrates genes involved in ILK signaling pathway. 148
Table 5.4 demonstrates genes involved in PXR/RXR Activation signaling pathway. 148
Table 5.5 Top five networks involved in pre-pubertal oocyte 155
Table 6.1: The top 20 enriched signalling pathways in one-year-old oocytes..... 180
Table 6.2: List of differential expressed genes in a aged oocyte (one-year-old)……...182
Table 6.4: Mitochondrial L-carnitine Shuttle Pathway..... 188
Table 6.5: Type II Diabetes Mellitus ..... .188
Table 6.6: Sertoli Cell-Sertoli Cell Junctions Signalling……..189
Table 6.7; Signalling by Rho Family GTPases……190
Table 6.8: GP6 Signalling Pathway…….....191
Table 7.1: List of CCs specific genes used as positive control genes for assessing the quality of CCs sequencing…….231
Table 7.2: List of the expressed differential genes in 1-year CCs 229
Table 7.3: Top functional networks involved in the aged CCs (1-year-old). 233
Table 8.1 Reverse transcription reagents 252
Table 8.2 Reagents for PCR mix reaction 253
Table 8.3 Reagents used for the preamplification mixture 260
Table 8.4 Reagents for pre-amplification reaction for the OpenArray qPCR 269
Table 8.5 Reagents used for OpenArray qPCR 269
Table 8.6: RNA-seq data showing 20 downregulated expressed genes.. 285

Abstract

Oocyte quality is the major determinant of successful fertilisation and subsequent normal embryo development. Maternal age is the greatest influencing factor limiting the success of infertility treatment. Current assisted reproductive technologies (ARTs) are unable to recover the fertility of females in two categories: pre-pubertal females who have survived cancer and older women who have delayed childbearing several years beyond their peak reproductive capacity. In the first category, women are at high risk of premature ovarian failure because of gonadotoxic chemotherapy. Thus, ovarian tissue preservation becomes an integral component of the treatment plan. However, the progress of *in vitro* maturation (IVM) of the pre-pubertal oocytes remains poor. In terms of the second category, oocytes from older women are correlated with poor fertilisation, low implantation rates and spontaneous abortion. Both pre-pubertal and aged oocytes have a limited developmental competence for sustaining successful fertilisation and normal embryo development. Also, it is unclear whether follicular microenvironment, which support oocyte development, have contributed to diminish the quality of these oocyte. To improve the ART success rate, it is necessary to elucidate the mechanisms by which maternal age negatively affects oocyte quality, representing an extremely demanding task. Little is known about the transcriptional activity that underlies the impaired competence of oocytes at these ages.

Using single-cell RNA sequencing (sc-RNA-seq), this thesis investigated the impact of age on the transcriptome of single oocytes and their companion CCs at prepuberty, maturity and advanced age. This was achieved, first, by optimising a preparation protocol for generating high-quality cDNA libraries for Illumina sequencing. Sc-RNA-seq was then successfully performed in a single oocyte, providing an unprecedented depth of investigation, by generating 30 million reads from each sample. sc-RNA-seq revealed novel pathways,

including retinoic acid receptor (RAR) metabolism and Rho signalling, suggesting that these pathways govern the transcriptional dynamics changes required to obtain competence. For aged oocytes, the dysregulation of RAR and mitochondrial L-carnitine was involved in the age-related deterioration. Interestingly, while CCs seemed to deteriorate as a function of advanced age, CCs associated with prepubertal oocytes did not reveal a distinct transcriptome in relation to those in young adults. Finally, the oocyte and CC sequencing data were validated using OpenArray real-time polymerase chain reaction (RT-PCR). In summary, single-cell transcriptomics provided insight into the intrinsic (oocyte) molecular changes and extrinsic microenvironment (CCs) associated with age. The aged follicular cells could contribute to the diminished developmental potential in oocyte senescence.

Abbreviation

Abcbg1	ATP-binding cassette transporter G1
Actb	ß-actin
ALdh1a1	Aldehyde dehydrogenase
ATRA	all-trans retinoic acid
ATP	adenosine triphosphate
AMH	Anti mullerian hormone
ARTs.	assisted reproductive technologies
Bmp15	Bone morphogenetic factor-15
BSA	Bovine serum albumin
cpm	count-per-million
CCs	Cumulus cells
Cebpb	CCAAT/enhancer binding
CoA	acetyl-coenzyme A
COC	Cumulus–oocyte complex
Colla1	Collagen alpha 1
Colla2	Collagen type alpha 2
Cola3a1	Collagen typeIII alpha1
Cpt1a	Carnitine
Ct	Threshold cycle
Ctgf	Connective tissue growth factor
Cyp11a1	Cytochrome P450 family 11 subfamily A member 1
Cyp7A1	Cholesterol 7 -hydroxylase
Dhrs	retinaldehyde reductase
ecG	exogenous hormone
ECM	extracellular matrix
PBS	Phosphate Buffered Saline free of calcium and magnesium
FDR	False discovery rate
FSH	follicle-stimulating hormone
GV	germinal vesicle
Gapdh	Glyceraldehyde 3-phosphate dehydrogenase
GCs	granulosa cells
Gdf-9	growth differentiation factor-9
GVBD	germinal vesicle breakdown
HFEA	Human Fertilisation and Embryology Authority
Hprt1	hypoxanthine guanine phosphoribosyltransferase 1
IGFS	Insulin growth substrate
ILK	Integrin Linked Kinase
Il15a	Interleukins 5

l17ra,	Interleukins 7
IVF	*in vitro* fertilisation
IVM	*in vitro* maturation
IPA	Ingenuity Principle Analysis
ICSI	intra-cytoplasmic sperm injection
LH	luteinising hormone
MPO	Myeloperoxidase
MPPS	massively parallel signature sequencing
MI	meiosis I
MII	meiosis II
mRNA	messenger ribonucleic acid
NAC	no amplification control
NGS	next-generation sequencing
PCA	polar body
PB	Principle component analysis
Pkc	protein kinase C
Pkcb	Protein Kinase C beta
PMSG	Pregnant Mare Serum
PXP/RXR	Pregnane X receptor /Retinoid X receptor
RT-PCR	real-time polymerase chain reaction
RA	Retinoid Acid
RAL	retinaldehyde
RAR	retinoic acid receptor
RT	Reverse transcription
Rhou	Ras homolog family
Rbp4	Retinol-binding protein 4
ROS	reactive oxygen species
Rpl19	ribosomal protein L19
RPKM	Read Per Kilobase Per Million
RT-qPCR	Quantitative, real-time polymerase chain reaction
SAGE	serial analysis of gene expression
sc-RNA-seq	single-cell RNA sequencing
Scd1	Stearoyl-CoA desaturase
Sep9	Septin
Slc27a2	Solute carrier family 27
Smart-Seq2	switching mechanism at 5’ end of RNA template
TCA	tricarboxylic acid
TGF-β	transforming growth factor-β
Tlr9	toll-like receptor
TSO	template switching oligonucleotide
Tubb	ß-tubulin
Vim	Vimentin

Chapter 1

Literature Review

1. Introduction

1.1. The function of the ovaries

The ovaries are the primary female reproductive organs, and they have a main function of producing oocytes and preparing them for the ovulation process. Ovaries are also endocrine glands that secrete steroid hormone that are vital to normal reproductive development and fertility. The process of developing female germ cells and the generation of primordial follicles is referred to as *oogenesis*; the growth and activation of these follicles is called *follicle development* (1).

1.1.1. Follicle development

The ovarian follicle, a functional unit of the ovary, consists of an immature oocyte called a *primary oocyte* and a shell of somatic cells. Follicle development is initiated when primordial follicles develop into the primary, secondary and antral stages of development; following this, one follicle is selected to attain preovulatory maturity, and it is subsequently ovulated (Figure 1.1) (2,3). Follicle development involves multiple endocrine and intraovarian interactions that modulate the intrafollicular environment for optimal oocyte development (4).

Many important aspects of early ovarian follicle development are regulated by paracrine factors secreted through the signalling between oocytes and the surrounding somatic cells (5). These factors regulate a wide range of follicle development, including with the activation of dormant primordial follicles and selection of secondary follicles; locally acting

factors have been identified that appear to exert important effects on the growth and differentiation of oocytes and granulosa cells (GCs; discussed further in section 1.3). Follicle growth is an extremely long and complex process; it takes 85 days for the primordial follicle to reach the preovulatory stage in humans (6) and 21 days in mice (7).

Oocytes arrested in the diplotene stage of meiosis I (MI) become surrounded by a single, squamous layer of somatic cells to form a finite population of nongrowing primordial follicles in the prenatal ovary (8). At puberty, primary follicles begin to be recruited from the primordial pool as oocytes interact with their surrounding somatic cells, namely GCs (7); this process is coordinated by bidirectional signalling mediated by paracrine factors between oocytes and their cumulus cells (CCs). The transition from the antral to the preovulatory follicle is associated with the growth of a fully competent, fertilisable oocyte (8).

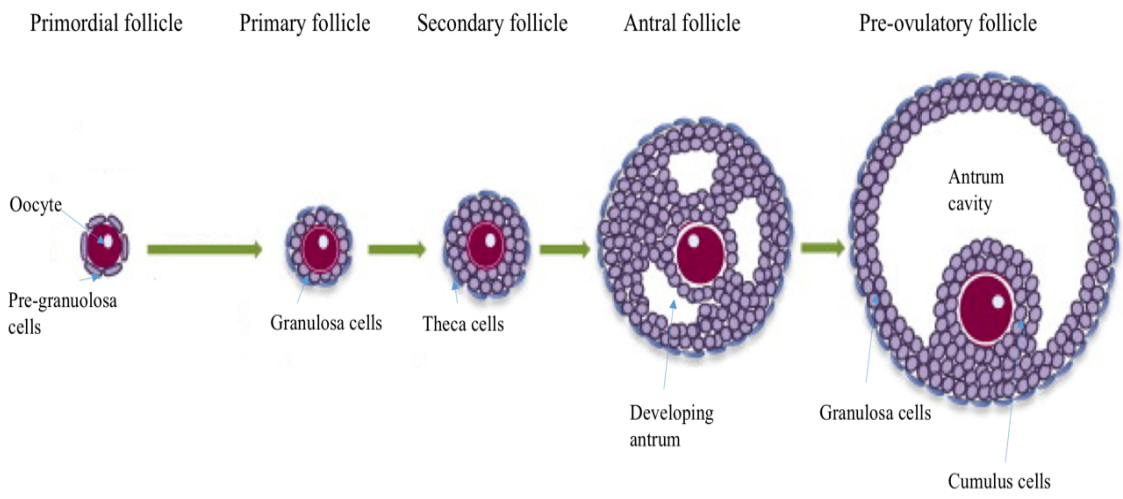


Figure 1.1. Follicle development stages from the primordial to preovulatory stage.

Immature oocytes grow in ovarian follicles, which consist of theca cells situated on the outside of a basement membrane and epithelial cells, GCs, in the basement membrane that have direct contact with the oocyte surface. As the follicles grow, GCs proliferate mitotically until a fluid-filled cavity (antrum) appears at the latest stages of follicle development. The classification of follicles relies on the number of GCs layers surrounding the oocyte: a primary follicle has a single layer; a secondary follicle has two layers; an antral follicle has multiple layers, with an antrum cavity beginning to develop); and finally, in a preovulatory follicle, the GC is differentiated to CCs (cells surrounding the oocyte) and mural GCs (cells lining the follicle), while the antrum is filled with follicular fluid. Figure 1.1 is adopted from Albertini *et al.,* 2009 (9).

1.1.2. Oogenesis

Oogenesis is a complex process that occurs in a coordinated manner with folliculogenesis during foetal development. Oogonia, which originate from primordial germ cells, proliferate by mitosis and enter the initial stage of meiosis (MI) to become *primary oocytes* (10). The diploid primary oocytes do not complete MI; rather, they stop at the first meiotic prophase stage, called *dictyate*. At this stage, the nucleus is called a *germinal vesicle* (GV); thus, this stage is called the GV stage of maturity.

In foetal humans, the peak number of oocytes has formed at approximately 20 gestational weeks; after this point, the number begins to decline, and only 1–2 million oocytes remain in the ovary at birth. By puberty, the number of oocytes has been reduced by 25% (6, 11). During the antral follicular development stage, the oocytes sequentially and gradually reach meiotic and developmental competence; it is during this phase of oogenesis that the oocytes acquire the cytoplasmic and molecular machinery they need to fully support embryo development (8). Especially, CCs support the oocytes' growth through the final phases of their development via gap junctions during antral development (Figure 1.2) (11).

Figure 1.2 An oocyte at the germinal vesicle (GV) stage.

It is surrounded by layers of cumulus cells (CCs); this cumulus–oocyte complex (COC) was retrieved from a preovulatory follicle in a stimulated mouse ovary (Picture was taken during experiments, Williams Laboratory, Oxford).

1.1.3. Oocyte maturation

Oocyte maturation is the final step of oocyte development occurring before ovulation. Oocytes achieve the capacity to undergo two aspects of maturation at the end of their growth phase, namely nuclear and cytoplasmic maturation. *Oocyte maturation* refers to a process in which the oocytes acquire the developmental competence to sustain fertilisation and further embryo development. The oocytes must undergo a successful, coordinated nuclear and cytoplasmic maturation to produce viable oocytes (12).

Nuclear maturation involves the resumption of the meiotic arrest at prophase I and the driving of the progression of MI to metaphase II. The cytoplasmic maturation of oocytes takes place throughout follicular development, from the primordial to preovulatory stages. Cytoplasmic maturation includes all the ultrastructural modifications that occur in the ooplasm during oocyte prematuration. Oocytes synthesise large quantities of messenger ribonucleic acid (mRNA) and ribosomal RNA, which are later used to generate essential proteins required for accomplishing oocyte maturation and are crucial for supporting the oocyte developing into a viable embryo after fertilisation (10,13)

In fact, oocyte maturation goes through a long and complex developmental process while competence is being acquired. Concurrent with chromosome condensation and migration, maturation is accompanied by a remarkable reorganisation of the cytoplasmic organelles, such as mitochondria, lipid droplets, cortical granules (CGs), Golgi apparatus, and endoplasmic reticulum (14), For instance during oogenesis, the number of mitochondria increases from 10 000 in the primordial oocytes to 100 000 in fully grown oocytes (15,16). In addition, CGs are formed from the Golgi and accumulate in the cytoplasm. Once maturation is completed, they localise beneath the oolemma (plasma membrane of the

oocyte) and prepare for releasing their secretions at fertilisation. This extensive remodelling and repositioning of intracellular organelles takes place at germinal vesicle breakdown (GVBD), throughout the transitions to MI, polar body (PB) extrusion and meiosis II (MII).

The oocyte is considered an extremely special cell. Unlike somatic cells, where RNA and proteins usually undergo rapid turnover, oocytes can recruit the stored mRNA and proteins in a highly regulated manner (17). In turn, oocyte genomes drive the oocyte growth until the early stage of embryo preimplantation, when the activation of the embryo genome occurs (18).

1.1.4. Gonadotropins and steroidogenesis

Ovarian follicle development relies on the sequential effects of the two main gonadotropins, namely follicle-stimulating hormone (FSH) and luteinising hormone (LH). Also, gonadotropins stimulate the secretion of steroid hormones, which are components of feedback mechanisms that responsible about gonadotropin secretion via actions upon the hypothalamic-pituitary system. FSH acts on early antral follicles to stimulate growth. LH, in synergy with FSH, acts on the FSH-stimulated follicles to maintain growth, and it is eventually responsible for the processes of luteinisation and ovulation (19).

Steroidogenesis is vital for the synchronisation of follicle growth and oocyte development, and it begins with an uptake of cholesterol from circulating lipoprotein and storage of the cholesterol in lipid droplets. Cholesterol is transported to the mitochondria by converting cholesterol to pregnenolone, which is then converted to progesterone, a key player in ovulation. Thus, cholesterol is the main precursor for the ovarian steroid production, and it plays a central role in follicle development (20,21)

1.1.5. Oocyte metabolism

The energy requirements during folliculogenesis include changing energy utilisation during meiotic resumption. The metabolic requirements of the COC increase throughout maturation, specially, at the final stages. Oocyte maturation require different compounds, such as fatty acids, amino acids, electrolytes, purines and pyrimidines (22).

1.1.5.1. Glucose metabolism

Glucose is the main energy source for the rapid dynamics of oocyte maturation. However, oocytes have a low ability to metabolise glucose, and instead, they rely on CCs for glucose metabolism. CCs can convert glucose to pyruvate, which is then delivered to the oocyte, specifically, the mitochondria, where pyruvate is converted to acetyl-coenzyme A (CoA). The latter then enters the tricarboxylic acid (TCA) and the electron transport chain to produce adenosine triphosphate (ATP). The mature cumulus–oocyte complex (COC) consumes twice as much glucose, oxygen and pyruvate as the immature COC does (22). An early study found that oocytes have a poor capacity to utilise glucose and CCs metabolise the bulk of the glucose consumed by the COC to supply metabolic intermediates to the oocytes (23).

1.1.5.2. Lipid metabolism

Oocytes contain large lipid droplet stores, mainly in the form of triglyceride and cholesterol. The oocytes require more energy at maturation; during the *germinal vesicle breakdown*

GVBD and proceeding to the MI stage, oocytes increase their energy demand and use (24). Moreover, oocyte lipids influence the developmental competence of the oocytes under normal physiological conditions. Evidence from a previous studies found an association between lipid droplets quantities in mouse oocytes and their developmental competence (25).

1.2. Age and oocyte competence

A successful maturation is provided by producing viable oocytes that sustain fertilisation and develop into healthy embryos. However, oocytes with poor quality have less potential. While fertility treatment relies on the oocyte quality, two categories of infertile patients are characterised with poor-quality oocytes – pre-pubertal females and older women. The poor quality is associated with limited competence. The scarcity of oocytes retrieved from these categories has made the animal models, such as bovine, sheep and mouse models, the preferable way to study the oocyte competence. The following review displays the current knowledge about changes associated with oocyte ageing by considering the most relevant findings obtained so far in humans and the mouse model.

1.2.1. The clinical indication of pre-pubertal oocytes

In recent years, the rates of cancer survival among pre-pubertal females have significantly improved, but fertility loss is a common side effect of chemotherapy, which destroys the

primordial follicles in the ovary. Although embryo and oocyte cryopreservation are the current gold standard methods for fertility preservation in cancer patients; this option is largely not available for pre-pubertal females. Hormonal changes have not yet commenced in pre-pubertal girls; as result, the maturation of hypothalamus–pituitary axis has not yet been approached. Ovarian tissue cryopreservation, although considered experimental, is the only viable option for fertility preservation for this category of infertile patients. The use cryopreservation of ovarian cortical biopsies containing the immature primordial follicles, followed by *in vitro maturation* (IVM) – a reproductive technology in which oocytes are retrieved from the antral follicles of unstimulated or minimally stimulated ovaries and matured *in vitro* is a fundamental part of the fertility preservation strategy. After cancer recovery, these immature oocytes from the primordial follicle undergo IVM combined with *in vitro* fertilisation (IVF) (26,27)

1.2.2. Previous studies on pre-pubertal oocyte quality

It is generally known that oocytes that have matured *in vitro* are less developmentally competent than those that have matured *in vivo* (28,29). The outcome of IVM is even poorer when maturing pre-pubertal oocytes compared with adult oocytes. The suboptimal development of pre-pubertal oocytes has resulted in lower cleavage and poorer blastocyst rates in mice (30), sheep (31), cows (26) and humans (32).

The low efficacy of IVM for the pre-pubertal oocytes remains a concern; thus, several ultrastructural studies have investigated the quality of *in vitro* matured pre-pubertal oocytes. Of these, an ultrastructure assessment of *in vitro* matured pre-pubertal lamb oocytes and

adult sheep oocytes, using electron transmission microscopy, was reported during the transition from GV to MII; the meiotic progression of pre-pubertal oocytes was reported at certain time intervals over 24 hours. Although both groups: pre-pubertal and adult oocytes reached maturation after 24 hours, there was a significant delay in the timing of nuclear maturation of the pre-pubertal oocytes compared with the adult GV oocytes. In addition, irregular oolemma (oocyte membrane) and abnormally scattered organelles were the most prevalent characteristics of cytoplasmic alteration in pre-pubertal oocyte (33). Moreover, adding recombinant FSH to the IVM medium improved the maturation rate of adult oocytes, but it did not have the same effect in pre-pubertal oocytes, revealing that oocytes at the prepuberty stage are less responsive to FSH. The study also showed that polyspermy fertilisation rate was higher in pre-pubertal oocytes compared with their adult counterparts (34).

An increased polyspermy rate in pre-pubertal bovine oocytes was reported in prepubertal lamb oocytes compared with adult oocytes following IVM, which was related to the lower volume of CGs in pre-pubertal lamb oocytes (35). The CGs are produced from Golgi complexes and localised below the oocyte membrane during maturation. When the sperm fuses to the zona pellucida of the oocyte, the CG initiates a series of pathways preventing the penetration of the oocyte by more than one sperm at the fertilisation event. It seems that the low volume of CGs present in the immature oocyte contributes to the defective cytoplasm of the pre-pubertal oocyte, leading to poor fertilisation and blastocyst rates. Interestingly, one study attributed the prepubertal oocytes' lower competence to the notion that they that may have been recovered from atretic follicles (36). In fact, it is likely that abnormal or incomplete growth of the oocytes will occur when the oocytes are recovered from variously sized follicles *in vivo* (after ovarian stimulation). Thus, oocytes containing

follicles are likely to exhibit various levels developmental capacity. In most studies, compromised cytoplasmic maturation seems the basis for the low maturation rates of IVM oocytes (37,38).

1.2.3. Transcriptome studies on pre-pubertal oocytes

Despite the enormous number of ultrastructural studies conducted to analyse the quality of pre-pubertal oocytes and improve IVM in the context of oocytes, the molecular mechanisms contributing to the oocyte's competence at prepuberty have not been addressed. Understanding the oocyte's transcriptome at the prepuberty stage is the key limiting factor for improving the success rate of IVM of the pre-pubertal oocyte.

In one study, using real-time polymerase chain reaction (RT-PCR), the transcript levels of six selected genes were compared between oocytes from pre-pubertal and adult bovines and oocytes retrieved from small and large follicles (from both pre-pubertal and adult cows). The findings revealed that follicle size did not have an effect, whereas a reduction in the transcript abundance of the six genes was found in *in vitro* matured pre-pubertal oocytes compared with the adult ones (39).

In a broader study, Paczkowski *et al.* (2011) carried out microarrays to analyse the transcriptome of *in vitro* matured pre-pubertal and adult porcine oocytes, using a pool of 150 denuded oocytes from each group (40). About 450 genes were differentially expressed in pre-pubertal compared with adult swine. However, these genes which have increased abundance in pre-pubertal females were classified as transcribed loci, or genes that have no

gene identification whereas genes increased in oocytes derived from cyclic genes were involved in processing genetic information and metabolism. Although differences in transcriptomes were detected between the two ages, the study did not clarify which aspects of the metabolism were altered in pre-pubertal oocytes (40). It seems that microarray studies have a major technical limitation, in that low-density arrays were used. One of the most important issues in microarray study is the necessity of having previous knowledge of the probes Therefore, microarray studies have limited ability to identify the genes involved in specific biological functions unless the probes have been determined previously (microarray technology is discussed in section 1.5).

As clarified above, most of the previous literature supports the claim that pre-pubertal oocytes have limited potential to sustain normal embryonic development and produce viable offspring. However, the mechanisms and components that govern oocytes' developmental competence remain unknown. Elucidating the activity of the pre-pubertal transcriptome oocyte at a molecular level is the first step in improving their quality. Considering the research from this angle is required to achieve a proper *in vitro* environment for enhancing the competence of pre-pubertal oocytes.

1.2.4. Advanced maternal age

It is well documented that fertility declines as women age. While the primary reason for this drop is the dramatic depletion of the follicle count, it further coincides with diminished oocyte quality toward the end of women's reproductive life. This decline starts in a woman's early 30s and speeds up after age 35; upon approaching 40, there is only a 5% chance of becoming pregnant in any month (Figure 1.3) (23). Age-related infertility is potentially due to diminished oocyte quality rather than endometrium function. The chance of successful

implantation increases when oocytes are taken from donors younger than 35 and then transferred to older recipients after maturing them in vitro (42).

Despite advances in the IVF success rate, a woman's age remains a limiting factor influencing IVF outcomes (Figure 1.4, adopted from Human Fertilisation and Embryology Authority [HFEA], UK) (43). The age-related decline in oocyte quality has been correlated with the increased number of spontaneous abortions, meiotic division errors and poor implantation rates in older women (44,45). Since primordial follicles are formed during foetal development, some authors think that normal oocytes may be prioritised to be recruited earlier in life, and then, the more abnormal oocytes remain for later recruitment (46). Others contend that oocytes are arrested and remain in the middle of prophase I until 40 years of age for women, which may result in the accumulation of abnormalities (47).

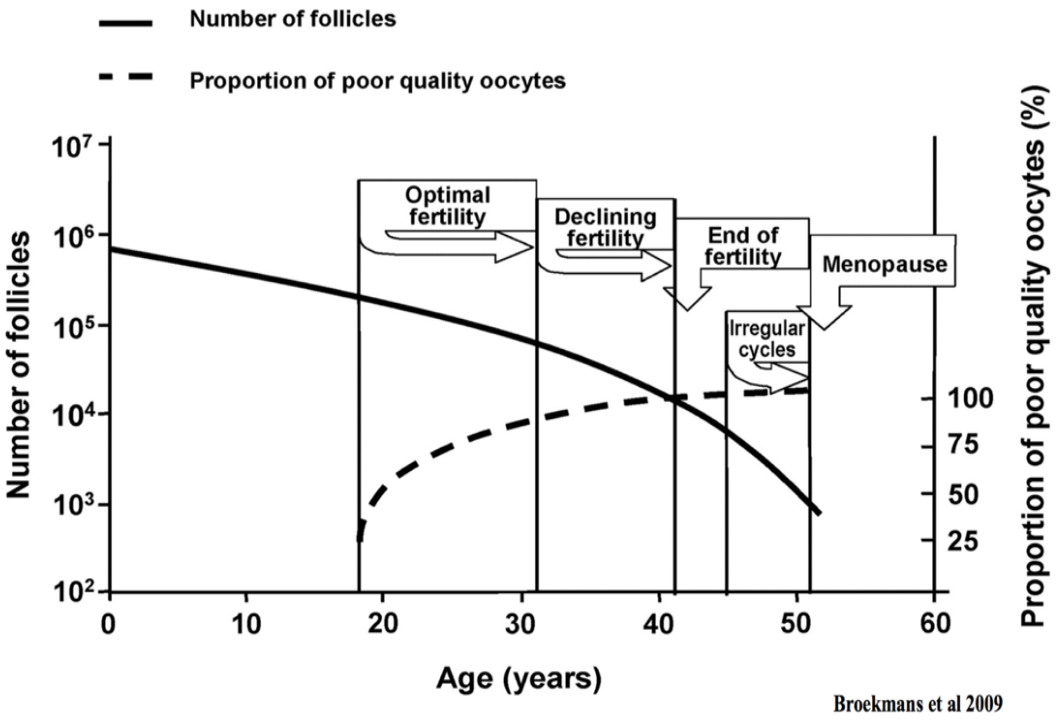


Figure 1.3 Quantitative and qualitative declines of the ovarian follicle pool.

Age-related female infertility is mainly based on changes in the ovarian reserve. The ovarian reserve can be defined as the number and quality of follicles and oocytes remaining in both ovaries at a given age (48).

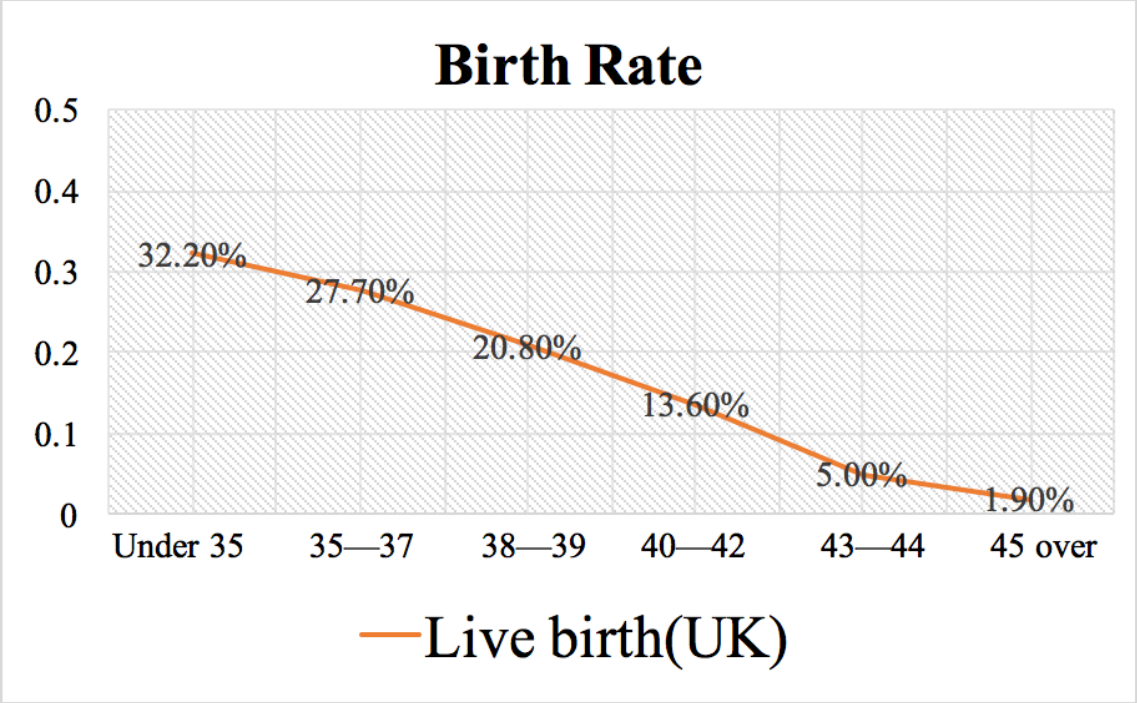


Figure 1.4 Birth rate related with maternal age in the United Kingdom, annual report by Human Fertilisation and Embryology Authority (HFEA, 2014-2016). The live birthrate per oocyte steadily decreases from 32.2% in younger women (<35) to just 5% for women at 43 (43).

1.2.5. Oocyte ageing studies

Most research on age-related fertility has highlighted the effect of an increased incidence of oocyte aneuploidy on older women's reproductive potential. Abnormal nuclear division, including failure in meiotic pairing, synapsis, recombination and segregation and nondisjunctional mechanisms have all been defined as causes of oocyte aneuploidy (39).

The limited potential of aged oocytes has led several researchers to carry out transcriptomic analysis of aged oocytes, as shown in Table 1.1. One of the earlier human studies used oocytes that were unsuitable for IVF, taken from two groups of women, divided based on whether they were younger or older than 36 years. The study examined the abundance of transcripts for only two genes related to the spindle function at the GV stage and MI and MII; the researchers found a significant decreased of mRNA abundance of these genes, indicating that mRNA degradation increases as women age (49).

The development of a mouse microarray platform by Carter *et al.* in 2003 for analysing stem cells and early embryo development (50) encouraged Hamatani *et al.* (2004) to take advantage of this microarray platform for investigating the differences in mRNA abundance in MII oocytes from mice aged 5–6 and 42–45 weeks. The microarray revealed relatively few differences between oocytes from young and old mice, showing only 99 genes with 2-fold differences. Most of the altered genes were related to mitochondrial activity, oxidative stress, DNA methylation and genome stability (51). In a similar concept, but with a different microarray chip, Pan *et al*. (2008) found 5% differences in the transcriptomes of aged GVs compared with young GVs; this number of altered genes increased to 33% in aged MII oocytes. The increased differences in transcript abundance after oocyte maturation raises an

interesting point concerning the possible failure of maternal RNA degradation in aged oocytes compared with what occurs during young oocyte maturation (52).

Using genome-wide analysis, Steurwald *et al*. (2007) compared the gene expression in oocytes of young women aged <32 years and women aged >40 years after failed IVF procedures. The study identified 136 genes with more than 2-fold difference between young and old eggs involved in the central biological functions of oocytes, including cell cycle regulation, cytoskeletal structure, energy pathways and stress responses (53); this is partially consistent with Hamatani *et al*.'s findings (51).

It is noteworthy that most published human studies have been conducted on oocytes that failed to fertilise after IVF or intra-cytoplasmic sperm injection (ICSI), where the oocytes did not possess the capacity to sustain the processes necessary for fertilisation and further development. Hence, the age factor cannot be identified as a major reason for the fertilisation failure of these oocytes, as unfertilised eggs represent suboptimal pools of oocytes, regardless of the woman's age.

In contrast to the previous human studies, Grondhal *et al.* (2010) based their study on mature oocytes donated by women younger than 36 years compared with those from women aged 37 to 39. Using an Affymetrix microarray platform, about 342 genes were differently expressed in aged oocytes (54). The findings of this study did not consider mitochondrial function or oxidative stress as significantly affecting the genes, as revealed in Hamatani *et al*.'s study (51); rather, the researchers discussed genes related to cell cycle regulation, chromosome alignment and sister chromatids (49). The variation of the affected genes in aged oocytes between studies mainly depends on the age category chosen for the comparison, as more abnormalities accumulate in more advanced age. The analysis is

normally dominated by the most apparent altered genes in the dataset. Labrecque and Sirard observed that the number of biological processes affected in aneuploid oocytes is similarly altered in oocytes from older women, which suggests a relationship between the two conditions (55).

In a study published in late 2017, the researchers conducted RNA-seq on individual mouse primordial follicles, demonstrating that protein metabolism was disrupted with advanced age. The dysregulation of protein metabolism was due to alteration in gene expression related with endoplasmic reticulum, nucleolus and protein processing (56).

Finally, studying aged human oocytes is challenging due to the limited number of oocytes in older patients. For this reason, it has been necessary to use animal models for investigating the developmental competence of aged oocytes in large-scale analyses.

Table 1.1: Microarrays studies for comparisons between the transcriptomes of young and aged oocytes.

Study	Single/pooled	Type/ strain	Age/stage	Findings
Hamatani *et al*., 2004	Pooled oocytes, 3 subsets of 500 oocytes	Mouse C57BL/6	5–6 wks old 42–45 wks old Meiosis II (MII) stage	99 genes are significantly differentially expressed, related to mitochondrial function and oxidative stress
Steurwald *et al.*, 2007	Pooled eggs 10X pooled oocytes per	Human	<32 years old >40 MII stage	Global decrease in transcripts related to cell cycle regulation, energy pathways
Pan *et al*., 2008	Pooled oocytes 4 sets (25 oocytes/sample) (35 eggs/sample)	Mouse B6D2F1	6–12 weeks 66 weeks GV & MII	5.2% of genes were differentially expressed in old oocytes; the difference increases to 33% between young and old eggs
Grondahal *et al*., 2010	14 Individual oocytes	Human	MII stage <36, 37–39 years old	342 of 2747 genes, involved in cell cycle, chromosome alignment, oxidative stress

1.3. Cumulus cells

Oocytes' growth and development are coordinated by the bidirectional communication between oocytes and their surrounding CCs. CCs communicate with the oocyte via

specialised transzonal cytoplasmic projections that penetrate through the zona pellucida and form gap junctions with the oocyte, thereby forming an elaborate structure called the COC (31). The mature COC consists of the secondary oocytes, which are arrested at metaphase II following the extrusion of the first PB and surrounding CCs.

1.3.1. The origin of CCs

Cumulus cells are derived from undifferentiated GCs, which are the main cell type in the ovarian follicle (32). The differentiation and function of cumulus cells depend on oocyte-derived paracrine factors. The development of the primordial to primary follicle corresponds with alterations in GCs' morphology from squamous to cuboidal shape. GCs proliferate and differentiate in synchrony with the growing oocytes. Oocytes act on the neighbouring follicular cells, which secrete several paracrine factors that promote follicular growth (59). With the formation of the follicular antrum, the GCs differentiate into two distinct cell lines of mural GCs, which line the follicular wall, and CCs, which form an intimate association with the oocytes.

GCs' differentiation is mediated by both hormonal and oocyte paracrine signalling. With the formation of preantral follicles, which almost correspond to the end of the oocyte growth phase, oocyte-specific members of the transforming growth factor-β (TGF-β) superfamily, especially growth differentiation factor-9 (Gdf-9) and bone morphogenetic factor-15 (Bmp15), participate in regulating the oocyte–granulosa loop during preantral follicle development (22). As a response to this interaction, the GCs produce paracrine secretions, including activin, inhibin, anti-mullerian hormones (AMH) and kit ligand (Figure 1.5) (60).

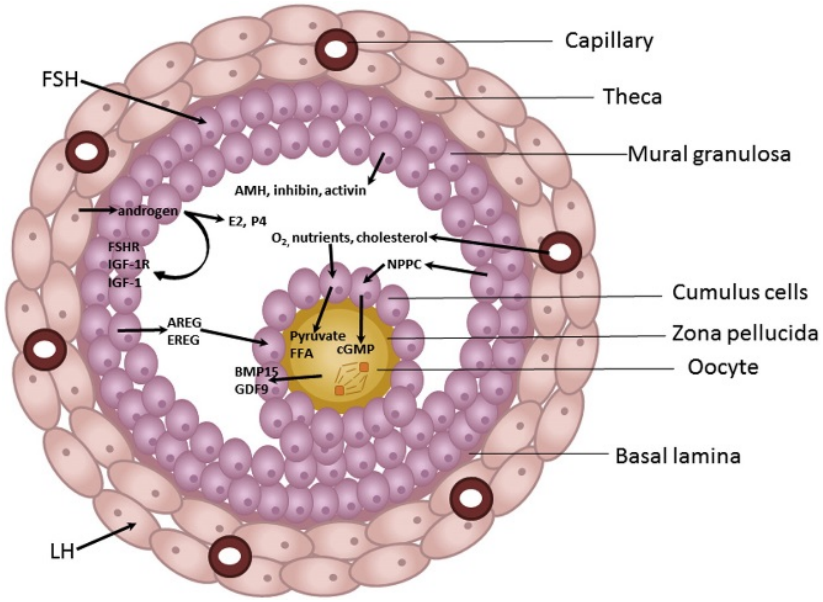


Figure 1.5 Diagram of an ovarian preovulatory follicle.

The figure is demonstrating how somatic cells provide oocytes via steroidogenesis, cellular signalling and metabolism. Abbreviations: AREG: amphiregulin, Bmp15: bone morphogenic protein 15, EREG: epiregulin. Figure 1.5 adopted from (295).

1.3.2. The role of CCs in follicle development

Studies have demonstrated that the knockout of the homozygous Gdf9 allele in mouse ovaries has blocked the follicle development by preventing the formation of the secondary follicles. In addition, the knockout of the Gja4 gene encodes an oocyte gap junction protein, resulting in the perturbation of oocyte growth and blocking follicle development. Thus, there is clear evidence that complex junctional communication mediates the information flow during oocyte maturation in the ovarian follicle. Taken together, the findings show that this regulatory loop between the oocyte and surrounding follicular microenvironment is fundamental for oocyte competence acquisition.

1.3.3. The role of CCs in oocyte meiosis

Early studies have documented that meiosis is controlled by CCs. Findings have shown that the CCs are responsible in different ways for keeping the oocyte under meiotic arrest, inducing meiotic resumption and supporting cytoplasmic maturation (57,61,62). An early observation by Pinctus and Enzmann (1935) have documented that the oocytes spontaneously resumed meiosis when they were isolated from their follicular environment (63). Eppig *et al*. (1991) demonstrated that substances are transferred between CCs and oocytes that are crucial in keeping the oocytes in meiotic arrest (62). Eppig and Down revealed that the CCs secrete a meiotic inhibitor, cAMP, that participates in the mechanism of maintaining oocytes in meiotic arrest (64). This has increased the knowledge concerning the CCs' necessity for providing a normal microenvironment for oocyte health (62).

Cumulus cells also facilitate the meiosis progression in oocytes by regulating the GVBD. Basically, prior to the ovulation process, CCs allow transmission of the LH surge, which triggers the resumption of MI in oocytes. In concordance with this, the CCs produce hyaluronic acid, which accumulates and expands in the spaces in the CC layers, thereby leading to what is called *cumulus expansion* (57). The CCs' expansion is detrimental for oocytes, and it has been documented that the inhibition of hyaluronic acid synthesis significantly decreases ovulation rates in mice (65). In addition, the role of CCs remains important even after fertilisation has occurred, as the CCs continue to attach to oocytes after fertilisation, thereby supporting the transport of oocytes through the fallopian tubes and facilitating the capture of the oocytes by the ciliated epithelial cells (66).

1.3.4. The influence of CCs on oocyte competence studies

Recent studies aimed to understand the influence of the CCs on oocyte competence. Due to the perception of the CCs' role in transmitting the signals in the growing follicle and the support of oocyte maturation, several groups have focussed on CC gene expression analysis to obtain insights into oocyte health and viability (67–70)

Yang *et al.*(2009) has reported an increase in apoptotic CC rates in abnormal morphological oocytes compared with morphologically normal oocytes Degenerate structures were observed in abnormal oocytes (71).

Two studies compared the transcriptomes of the CCs associated with oocytes at different maturation stages (GV, MI, MII). Ouandaogo *et al.*'s study (2011) revealed that the transcriptomes of CCs surrounding GV, MI and MII are similar, indicating that companion

CCs are not necessarily associated with the maturation situation of the corresponding oocyte (72). This finding suggests that mature oocytes in MII could be surrounded by CCs with gene expression representing that of earlier stages (GV or MI). However, this could reveal that a lack of synchrony between the oocyte and the CCs could negatively affect oocyte competence, although evidence is needed to confirm this assertion. In contrast, using a different microarray chip, Feuerstein *et al* (2012). identified 734 differential genes in the CCs of MII compared with the GV, involved in regulation of a total of 16 processes, including the mitogen-activated protein kinase pathway, lipid biosynsthesis and apoptosis (73). The disagreement in the results between the two studies demonstrates that the transcriptomic method used adds a significant difference in revealing delicate findings that have important implications for ovary biology. (The difference between the transcriptomic methods is elucidated and further explained in section 1.5). Although many studies have addressed the CCs' transcriptome, none have considered the transcriptome associated with prepubertal oocytes.

An increased focus on the ovarian environment, especially CCs, associated with aged oocytes has emerged in recent years. A study conducted by Pacella *et al.* (2012) aimed to determine whether an altered follicular environment is affected by maternal age; they focussed on the GCs and follicular fluid from two groups of patients, namely those of young maternal age (<35 years) and advanced maternal age (>40 years). The study found that the expression levels of genes involved in the regulation of progesterone, prostaglandins and apoptosis were higher in the CCs of older women and reduced the ovarian reserve (74). In addition, alteration of the metabolite levels was found in the follicular fluid of the patients with advanced age. Given that follicular fluid comprises hormones and carbohydrates that are basically synthesised by follicular cells (75), the

altered levels of the components of follicular fluid reflect the perturbed environment in both GCs and CCs in aged follicles.

A microarray study revealed that angiogenic genes, including Angiopoietin Like 4 (ANGPTL4), Leptin receptor (LEPR), Transforming Growth Factor Beta Receptor 3 (TGFBR3) and **fibroblast growth factor 2** (FGF2) were significantly overexpressed in older patients (>40 years) compared with patients in the younger groups. In contrast, genes implicated in TGF-β signalling pathways, such as AMH, TGFB, inhibin and the activin receptor, were underexpressed. It seems that, in these pathways affected by aging in CCs, the oocyte quality declines with age (76). Recently, Molinari *et al.* (2016) used RNA-seq for comparing the transcriptome profiles of CCs collected from young patients (<35 years) and older patients (>40 years). The study identified that 45 genes were differentially expressed, and most of the altered genes were involved in hypoxia, glycolysis and vascular development (77).

The CCs form an intimate association with the oocyte in the follicle, and they can have a significant influence on oocyte meiotic and developmental competence. This enables CCs' use as a non-invasive tool for measuring oocyte competence. Human studies have shown a perturbed environment, which may contribute to impaired oocyte developmental competence. This review has discussed studies that recently published (during this study) as the CC studies are extremely limited. More importantly, CCs associated with pre-pubertal oocytes have never been investigated due the infeasibility of such research. The next section of this chapter provides an overview of the most common transcriptomics technologies.

1.4. Overview: Historic development of transcriptomics technologies

1.4.1. Hybridisation based method: Microarray methods

Early transcriptome studies developed various strategies to study transcription activity. Of these methods, hybridisation-based microarray technology has always been the most widely used. In principle, it involves using predefined prop chips with fluorescently labelled cDNA (extracted RNA converted to single-strand cDNA mixed with fluorescent nucleotide) that is hybridised to these predesigned probes (78). The microarray approach has successfully enriched transcriptome studies, and its strength comes from its ability to generate a large amount of data at a relatively moderate cost (53). This method has led to advances in the understanding of biological problems such as the identification of differential gene expression between healthy and diseased cases (79).

However, the microarray approach suffers from certain limitations that prevent an accurate characterisation of the transcriptome. First, it suffers from background noise caused by cross-hybridisation, non-specific hybridisation, and single saturation (80,81). This factor limits the detection of genes, especially those expressed at low abundances, which might be physiologically relevant as a result of low signal saturation. Another major issue is that microarray chip design requires a previous knowledge of the genes contributed to the condition tested. Predefined probes are therefore designed and developed to meet the research purpose. As a consequence, discovering unknown transcript splice variants is not possible. Moreover, this approach has a limited dynamic range of gene quantification due to the inaccurate quantification of the high folds of gene expression. The collective variations

and biases in the generated data become a matter of concern when comparing expression levels across different studies (82). These drawbacks limited the progress of deep transcriptome profiling, thereby stimulating the development of newer methods capable of solving these issues.

1.4.2. Sequence-based methods

The development of sequencing-based methods eliminated the challenges posed by microarray technologies. The sequencing approach, in principle, has the advantage of being able to directly access and sequence RNA or cDNA and produce a digital count of the transcript abundance. The resulting reads are then mapped to the source genome or reference transcript and counted to obtain the number of reads for the gene of interest, which represents its level of abundance (83) In this way, the method avoids the biases and variations associated with hybridisation and labelling efficiencies in the microarray approach. Subsequently, the digital quantification allows for the capture of a broader dynamic range (approximately five logs) and offers a reliable, comprehensive measurement of the transcriptome (84).

Historically, Sanger sequencing is considered the first generation of sequencing technology (85). However, the low throughput and high cost of the Sanger technique made it less than ideal. As a result, demand grew for a modified Sanger-based method available at an affordable cost. Therefore, tag-based methods such as serial analysis of gene expression (SAGE) and massively parallel signature sequencing (MPPS) were developed to produce accurate digital gene expression data in a high-throughput manner, which is advantageous in comparison to the analogue-style signal obtained from the microarray method (79,80). Nevertheless, a significant portion of short tags are technically difficult to be uniquely

mapped to the reference genome. Moreover, isoforms are not distinguishable. Due to these disadvantages, the tag-based method is limited as it is not suitable for large-scale studies because of issues related to sequencing capacities and costs (86).

1.4.3. Next-generation sequencing

Despite many technological improvements in microarray and traditional sequencing technologies, demand has risen for revolutionary technology delivering fast, inexpensive, and accurate transcriptome information. Therefore, recent years have seen a substantial shift away from applying these methods for analysing genomes and transcriptomes. RNA sequencing is a newly developed approach that permits the direct sequencing of transcripts using next-generation sequencing technologies. It has strong potential for replacing microarray technologies for whole-genome transcriptome profiling (87)

The introduction of next-generation sequencing (NGS) technologies (also called *high-throughput sequencing*) has revolutionised transcriptomic studies (88). This category includes SOLiD, Hiseq, Illumina, and Roche454 Life technology. These NGS methods facilitate RNA analysis through cDNA sequencing on a massive scale. In fact, NGS platforms can sequence as many as one billion DNA strands in parallel per instrument run, resulting in straightforward, higher sequencing performance at a lower cost. Any of these high-throughput technologies can be used for RNA sequencing. All NGS platforms have similar workflow steps, and each has some advantages of specific application over others (89–92).

1.5. Single-cell Transcriptomics

Until recently, multiplexed quantitative real-time reverse transcription -polymerase chain reaction (q-RT-PCR) has been the method of choice for single-cell gene expression profiling in individual cells (93). In fact, multiplexed q-RT-PCR is known as the gold standard for accurately quantifying mRNA, however, the data generated is considerably biased toward the genes selected for examination. Besides this, it is not feasible to survey the entire transcriptome because q RT-PCR remains limited for hundreds of genes of interest (94)

The microarray approach has enabled single-cell transcriptome analyses within large-scale studies (95). The initial global mRNA expression profile of single cells was carried out on hippocampal neurons using a microarray approach (96). Since then, several studies have applied similar methods to other cell tissues, such as pancreatic, retinal, stem and cancer cells (97–100). These studies have used similar single-cell amplification protocols followed by array-based approaches. Different versions have become available, although all of these follow a similar principle. The lack of sensitivity and artefacts associated with hyperdisation were the most consistent drawbacks restraining further progress in terms of developing this technology (101). This limits the ability of both q-RT-PCR approaches and microarrays to quantify diverse RNA molecules over a wide range of levels (102).

1.5.1. Single-Cell (sc) RNA Sequencing

Due to the technical advances in RNA sequencing in recent years, this approach offers clear advantages over existing sequencing approaches for bulk-cell analysis. Therefore, substantial effort has been made to adapt RNA sequencing for single cell applications, and

not only for populations of cells (103,104). Single-cell approaches circumvent the averaging artefact associated with the traditional population cells and reveal genetic expression dynamics that could be masked in bulk analysis.

Deep sequencing technologies have permitted the characterisation of transcriptome variations across tissues at unprecedented resolutions. Recent advances in whole transcriptome amplification have enabled and facilitated the quantitative sequencing of the minutes' amounts of RNA in single cells (105). Thus, such approaches do not require the thousands of input cells normally involved in bulk analyses. More crucially, studying cells at the single-cell level provides a unique opportunity to investigate the transcriptional heterogeneity within tissues or cell types and to dissect the interplay between intrinsic cellular processes and extrinsic environment. This approach also has the potential to trace cell lineages, thereby providing insights into diverse processes, such as development, gene expression dynamics, tissue heterogeneity and disease pathogenesis (106)

Despite the fact that RNA sequencing dominates most transcriptomic studies, this approach also faces critical technical challenges involving sample preparation, library construction, data generation, analysis and bioinformatics. These technical issues must be overcome to gain the anticipated biological insights using the single-cell approach.

The RNA sequencing experiment is a population of RNA molecules converted to a library of cDNA fragments ligated with adapters. These short fragments are then amplified and parallel sequenced in a high-throughput manner to obtain short sequences. Successful sc RNA sequencing experiments mainly rely on the successful combination of two independent steps: (A) the isolation of the individual cell of interest from a tissue or dissociated cell

suspension, during which minute quantities of cellular RNA is recovered and converted into cDNA, and (B) the massive parallel sequencing of the cDNA libraries (Figure 1.6).

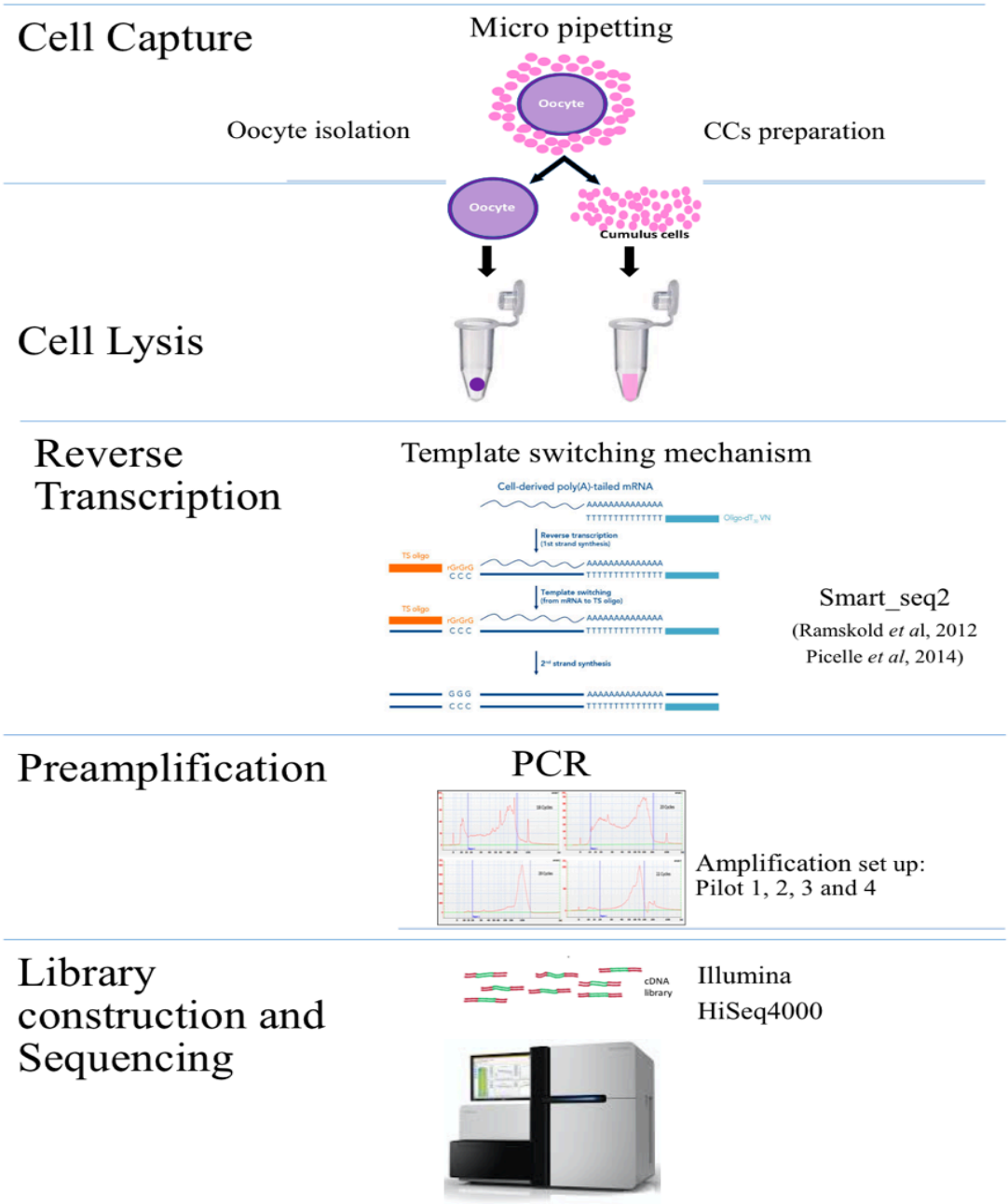


Figure 1.6 Schematic of RNA sequencing experiment for single oocyte and their companion cumulus cells (CCs). Each individual cumulus oocyte complex (COC) was mechanically separated and immersed in lysis buffer. The lysates were then primed and reverse transcribed to complementary DNA (cDNA) using Smart-seq2 (developed by Ramskold et al, 2012). cDNA samples were amplified by PCR. The cDNA libraries were fragmented and tagmented using Nextera XT protocol, sequenced using High seq4000.

1.6. Reproductive lifespan in mice

Female mice become sexually mature when they are six weeks old. The reproductive cycle in mice, called the estrous cycle, lasts 4-5 days, and they have a gestation period of about 20 days. Follicles in mice require 18-21 days to develop (7). Although mice are unlike humans in that they continue to produce litters even at an advanced age, their reproductive potential dramatically declines after eight months old, and this decline becomes obvious at 12 months for the selected strain C57BL/6. The reproductive lifespan for this strain ends earlier than other strains, such as B65JlF1 and MF1, which is about two years. Therefore, there is a lower level of oocytes in the ovaries of 12-month-old C57BL/6 mice (107) (Figure 1.6)

1.6.1. Age category selection

The age categories selected were based on critical time points in the reproductive lifespan of the mice, assuming that each egg from each stage should have distinct genes expressions based on the clarifications discussed below.

A. 3 weeks old (sexually immature): A pioneering study by Hirshfield (108) suggested two populations of primordial follicles in post-natal rat ovaries, with each dominant in a particular period of the reproductive lifespan. In a recent study by Zheng *et al.* (2015), the primordial follicles were traced using a tamoxifen-inducible reporter mouse model (Fox2-CreERT2; mT/mG), and it was demonstrated that the first population of growing follicles are not activated until days 45-50 and had become exhausted by three months. This wave is functioning for puberty, and then it is replaced by the second wave, which dominates the ovary and by that, provides fertility throughout the reproductive lifespan (109). Thus, it seems that the two classes of primordial follicles are age-dependent, and

they are crucial to understanding the mammalian reproductive lifespan. Therefore, the ovaries of 3-week-old mice were expected to exhibit a different transcript from those in sexually mature mice.

B. 9 weeks old: Mice are at the prime reproductive age. They are sexually active and produce an average of 6-10 pups after the first follicle wave expires. Therefore, samples at this age were used as a reference for comparison.

C. 1-year-old: This ageing model is similar to the human premenopausal stage.

A quantitative cytological analysis demonstrated that primordial and growing follicles were exhausted by 13-14 months (110). In addition, the number of oocytes collected from the ovaries of 9-month-old and 1-year-old mice was also assessed in our lab. The average number of eggs was about 16 and 6, respectively. There was a dramatic drop in eggs number toward approaching one year old. Therefore, the one year was appeared to be the most appropriate age reflecting the human pre-menopausal age in female patients.

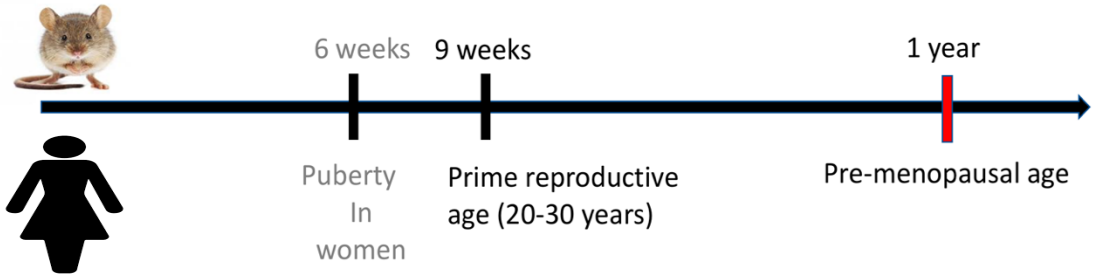


Figure 1.7 The reproductive lifespan of C57BL/6 female mice compared to the female human reproductive lifespan.

At three weeks old, ovaries in mice resemble the ovaries of female humans prior to puberty. Female mice become mature at 6 weeks-old which resemble the puberty age. Nine weeks is the prime reproductive age for mice, which parallels female humans where the best fertility rate occurs between the ages of 20 and 30 years old. Finally, at one year of age for C57BL/6 mice, the oocytes become old, which is thought to be a good model for the oocytes of human females at the period just before menopause.

1.7. Hypothesis

Poor oocyte quality is a major reason for poor IVF treatment outcomes, and this is attributed to the oocytes' limited developmental competence. Pre-pubertal oocytes are found to be less competent for sustaining fertilisation and normal embryos than adult oocytes are (31,34). Despite the poor IVM outcomes for pre-pubertal oocytes, the reasons for their limited competence have not been sufficiently investigated. Furthermore, aged oocytes have a diminished quality, representing a major reason of age-related infertility in older women.

Since successful fertilisation and healthy embryo development are heavily dependent on oocyte competence, we hypothesise that a limited oocyte competence can be attributed to alterations in transcriptional activity associated with age. We suggest that certain pathways have modulated the transcriptome, which has seemingly had an adverse influence on the normal biological process of oocytes, thereby causing a decline in oocyte quality. Yet, a lack of knowledge on the influence of age on the transcriptomes of oocytes at different ages and whether these changes, if they exist, arise before or after maturation, is still evident. Both pre-pubertal and aged GVs present a good model for studying the limited developmental competence of oocytes, as oocytes at pre-pubertal or advanced ages have normal morphological features making them undistinguished from those young adult oocytes with full competence. The mechanism involved in the transcriptome of the pre-pubertal oocyte has never been addressed, especially at the molecular level. In addition, a detailed transcriptome assessment of the oocytes at advanced age has been poorly defined using high-throughput techniques, in which the entire transcriptional activity is investigated.

As oocyte growth and maturation represent a complex process coordinated by bidirectional communication between oocytes and their companion CCs (11), it is possible that deterioration in the function of these critical cells (CCs) may play a role in declining oocyte competence. As such, aged CCs could to be involved in the age-associated decline in oocyte quality. Despite the importance of CCs, characterisation of changes to the transcriptome related to age has seldom been undertaken using RNA-seq, as most studies have involved pooled oocytes or CCs, which are likely to be heterogeneous in terms of competence. In addition, it is still unclear whether a poor oocyte quality is an outcome of the impaired CCs or the CC development has influenced that oocytes. Therefore, we have employed RNA-seq to study single oocytes and CCs retrieved from the same COCs; this will allow us to gain insights into the intrinsic (oocyte) and extrinsic factor (microenvironment: CCs) underlying the molecular changes associated with oocyte quality.

1.8. Aims

The overall objective of this thesis is to investigate the impact of age on the transcriptome of immature, mature and aged mouse oocytes and their companion CCs using RNA-seq. This thesis adopts the following specific aims for accomplishing this objective:

- **Aim 1:** To develop a preparation protocol for RNA recovery from individual oocytes and their companion CCs;
- **Aim 2:** To optimise a cDNA library for the sc-RNA-seq;
- **Aim 3:** To investigate the differences in the transcriptome profiles of a single pre-pubertal oocyte (3 weeks old) compared with a young adult oocyte (9 weeks old) using RNA-seq;

- **Aim 4:** To investigate the differences in the transcriptomic profiles of a single aged oocyte (1-year-old) compared with a young adult oocyte (9 weeks old);
- **Aim 5:** To determine whether age has an effect on the CCs' transcriptome at both ages – pre-puberty and aging – in comparison with young adulthood; and
- **Aim 6**: To validate the sc-RNA-seq data of individual oocytes and their accompanying CCs using OpenArray RT-PCR in all age categories – pre-pubertal versus young adult, aged versus young adult.

Chapter 2

Material and Method

2.1. Strain selection

Mice are the ideal mammalian model for studying the genetics of ageing. The advantages of using mice are their short reproductive lifespan, the genetic manageability of its genome, the ease of breeding and the large amount of available baseline phenotype, which collectively allow for obtaining insights into the pathogenic mechanisms of the ageing process (111,112). Several age-related infertility studies have used different strains of the mouse model (52,107,113–117) (Table 2.1). Female C57BL/6 mice, the most commonly used inbred strain, have been used to investigate the impact of age on the transcriptomic profile of oocytes and cumulus cells (CCs) during certain stages of the reproductive lifespan.

C57BL/6 is an inbred strain produced by 20 generations of brother-sister mating, which ensures that this strain has an identical genetic background, making it an attractive choice for studying age-related molecular changes (111). The advantages of this strain are defined genomes and stable phenotypes, and because this strain is the only mammalian species that has its entire genome profile published, extensive data are available from published studies and databases, such as mouse genome informatics.

Table 2.1: Mice strains used for ageing studies

Demonstrated	Mice strain	Age	Study
Age-associated molecular changes in aged oocytes	C57BL/6	4-6 weeks 42-45 weeks	Hamatani *et al*, 2004
Age-associated increase in aneuploidy	B65JLF1 B6D2F1/J CF1	6-12 weeks 60-70 weeks	Pan *et al*, 2008
Age-related meiotic segregation errors in aged oocytes	C57BL/ICF9	5-6 weeks 42-45 weeks	Lister *et al*, 2010
Kinetochore microtubule attachment	MF1	5-6 months 15-16 months	Shomper *et al*, 2010
Spindle Assembly Checkpoint	B6D2F1/J	6-14 weeks 16- 19 weeks	Duncan *et al*, 2009
Weakened centromere cohesion in aged oocytes	B6D2F1/J	6-8 weeks 15-18 weeks	Chiang *et al*, 2010
Age-associated telomeres in aged oocytes	C57BL/6J	6-8 weeks 42-48 weeks	Yamada-Fukug *et al*, 2013

2.2. Mice maintenance

C57BL/6 mice were purchased from Charles River, Oxford. The number of mice was as follows:

1. 5 mice at age 3-weeks-old
2. 5 mice at age 9-weeks-old
3. 8 mice at age 1-year-old

Mice that were intended for use at age 3 weeks and 9 weeks were purchased at 2 weeks and 8 weeks, respectively, to enable them to acclimatise them to the pathogen-free environment of the Biomedical Science unit (BMS), Oxford (3-week-old mice were delivered with the mother, as they had not been weaned at the time of delivery). For 1-year-old mice, eight mice were purchased at age 9 months and were kept at the BMS facility until they were 1-year-old.

Lengthy samples processing has adverse effects on the RNA integrity of the collected samples because RNA is susceptible to degradation. Therefore, to obtain samples of good quality, mice were processed as follows: two mice per day were processed for mice aged 3 and 9 weeks old, whereas 1 mouse per day was processed for 1-year-old mice to ensure that the delayed procedure was avoided, as aged ovaries have a very low number of follicles, which require careful attention to avoid losing any follicles.

2.3. Ovary stimulation and collection

To obtain a uniform development stage of the follicle population for conducting a single cell-RNA-sequencing (sc-RNA-seq) experiment, mice were stimulated with 5 IU Pregnant Mare Serum (PMSG) (Invitrogen, UK) 48 hours prior to ovary collection. The given injection is to synchronise follicle development. At the time of collection, mice were humanely sacrificed, and each mouse was cleaned with RNaseZap Wipes (Ambion, UK). Dissection tools and the work area were also sprayed with RNaseZap (Ambion, UK) to guarantee an RNase-free environment. Mice were then dissected, and the ovaries were collected and transferred into Eppendorf tubes containing Dulbecco's Phosphate Buffered

Saline free of calcium and magnesium (BSA/PBS) (Gibco, UK). The tubes were kept on ice while transferring the samples to the lab.

2.4. Cumulus oocytes complex (COC) preparation

The ovaries were cleaned of fat and washed in PBS. Pre ovulatory follicles were individually punctured using 30-guage syringes (Sigma-aldrich, UK) under a stereomicroscope. The COCs were then released from the follicle and serially transferred into small drops of Phosphate Buffered Saline containing Bovine Serum Albumin (/PBS/ BSA, 1 mg/ml) (Gibco, UK) with the aid of mouth pipetting, and only large follicles were chosen. This to ensure that the collected follicles are at the final stage of maturation. The COC was washed in a serial transfer through 10 µl drops of BSA/PBS, which removes any contaminating follicular cells. Only a good morphology of COCs, which were completely closed by tight layers of CCs, were selected for the analysis, and oocytes with clear zona surrounded by layers of CCs were processed.

2.4.1. Oocyte isolation

Each cell type was processed separately. As follow: the oocytes were denuded from the cumulus cells by drawing them into fine glass capillary tubes (Hershmann® microcapillary pipette, Sigma-Aldrich, UK) using mouth pipetting. The 20 µl capillaries were pulled to achieve a smaller diameter fitting the COC and to avoid overloading the media. Each oocyte was washed sequentially in PBS to ensure that it was free of CCs. Since any CC that remained attached could contaminate the oocyte transcriptome. If this had occurred, an

overlap between the two cells' transcriptomes would have led to inaccurate sequencing results.

Each oocyte was treated individually in a 5 µl drop of acid Tyrode's solution (Cat. no. T1788 Sigma-Aldrich, UK) for less than 30 seconds until the zona pellucida disappeared. The oocyte was then sequentially washed in three tiny drops of PBS/BSA. It was transferred in a negligible amount of PBS/BSA media (the lowest possible amount [0.3 µl of PBS/BSA] is preferable, which a microcapillary pipette makes possible) and immediately placed in 4.45 µl of 0.4% lysis buffer (contained RNase inhibitor, lysis buffer preparation is described in section 2.4.2). It was then vortexed and centrifuged for 30 seconds at 2000g to ensure that the oocyte was immersed and lysed by the lysis buffer.

2.4.2. CCs Preparation

For the CCs, the establishment of optimisation protocol is detailed in section 3.3. Briefly, each matched CC were placed in 5 µl PBS/BSA, centrifuged at 120g. Supernatant was discarded and 0.5 µl of the cells pellet was transferred and immersed in a lysis buffer, as indicated by the Smart-Seq2 protocol. Cumulus cell samples were collected in the same way throughout this project. Normlising CCs samples to cell number per sample was unfeasible because CCs preparation necessitated lengthier procedures than used for the oocytes especially in single oocyte experiment. As retrieving a single follicle followed by separating the oocyte and CCs, processing them individually at the same time was the most important criteria for conducting this study.

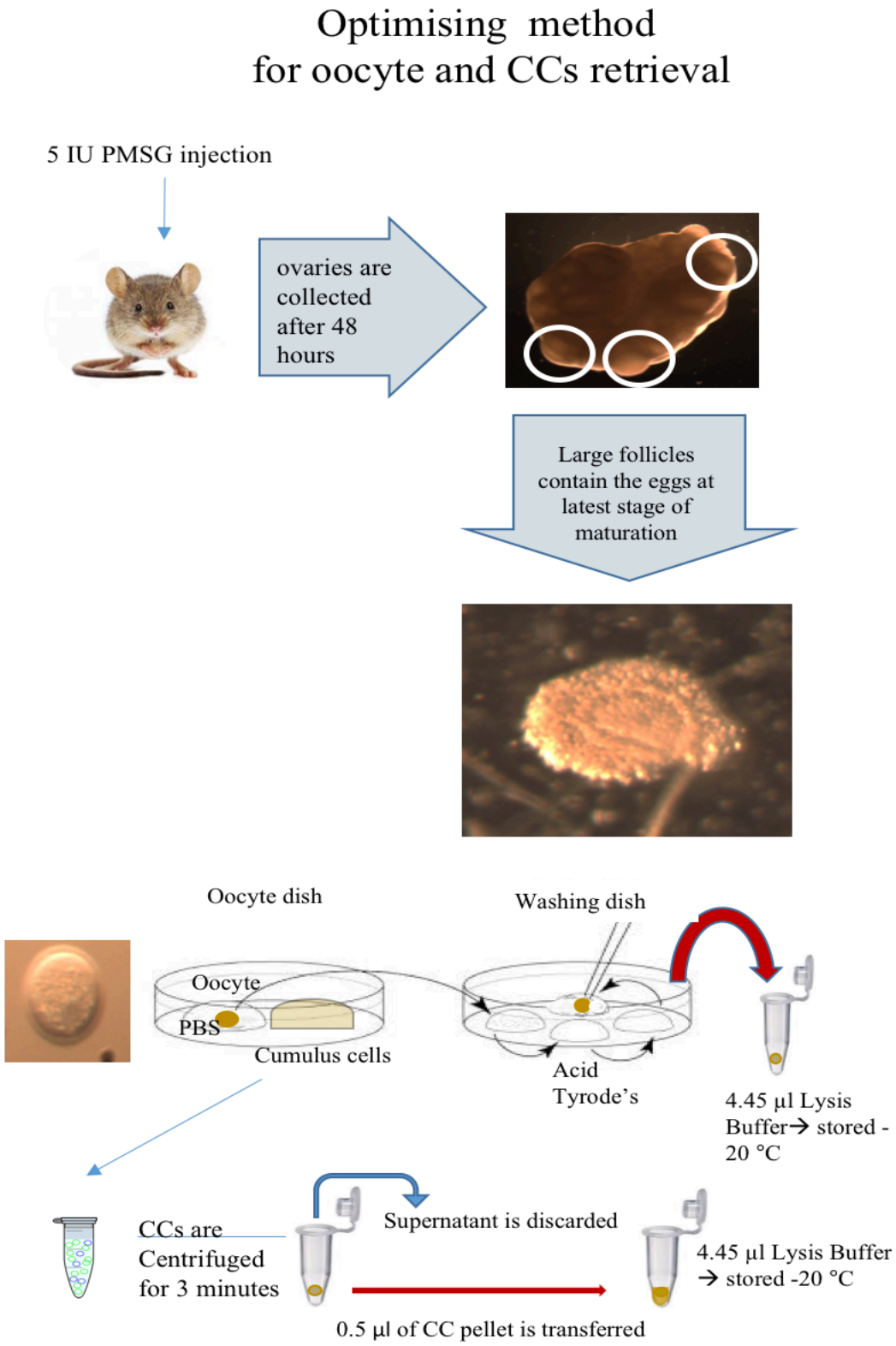


Figure 2.1 Single oocyte preparation steps: (A) an ovary containing large follicles (white circles), (B) a COC, and (C) a stripped oocyte at the GV stage.

2.5. Smart-Seq2 Approach

For single-cell work, a single tube approach was used for rapid cell lysis (including RNA priming), followed RT and PCR. Using the single tube approach is highly critical in terms of eliminating the possible loss of RNA during repeated transfers between tubes, as occurs in other RNA preparation methods.

The switching mechanism at 5' end of RNA template (Smart-Seq2; developed by Picelle *et al* (2014) approach was adopted for use in preparing cDNA libraries from single oocytes and their CCs. In brief, the RNA was primed using Oligo(dT) primers added to the lysis buffer (Thermo Fisher, UK; Figure 2.2). This permitted the selection of RNA with poly adenylated tails during reverse transcription. This strategy enabled the efficient capturing of the informative mRNA, while the rRNA and tRNA were removed (118). The distinct element of this approach is the template switching mechanism (template switching oligonucleotide [TSO]) for the second strand synthesis, which insures improved transcript coverage. Using TSO primer (added in the PCR reaction) enabled coverage of more transcripts at the 5' end and consequently maximised the cDNA yield (118) comparison to the poly tail priming, such as that used in the Tang protocol and QuartzSeq where a bias in the 3' end is apparent in the sequencing data (119)'(120), as the chapter 3 later addresses.

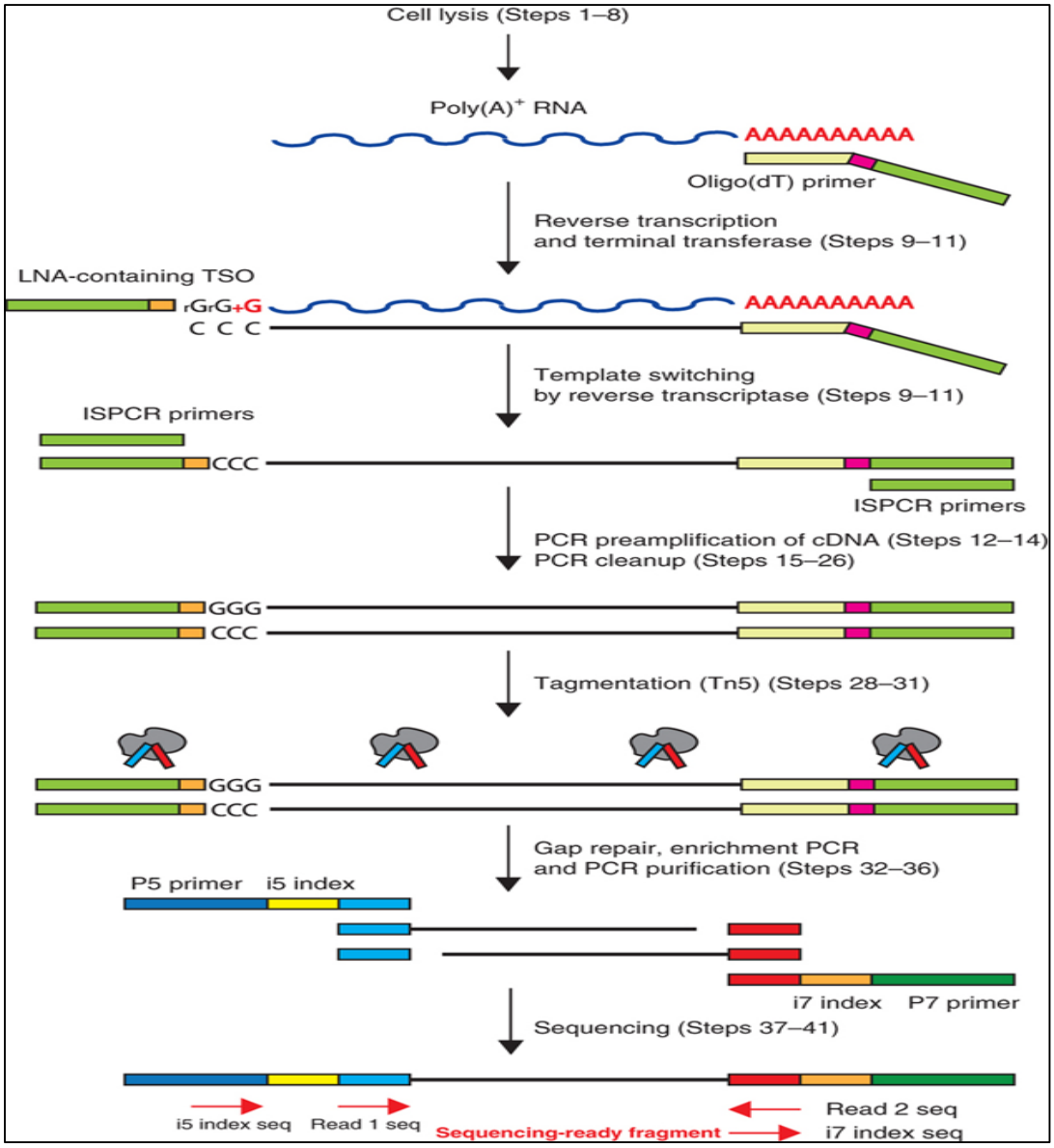


Figure 2.2 Smart-Seq2 protocol, developed by Ramskold *et al* 2012, Nature protocol.

2.5.1. Cell lysis buffer set-up

Lysis buffer was prepared using the quantities and reagents indicated in Table 2.2. The lysis buffer was prepared around an hour prior to the collection time and kept frozen until use. During samples processing, the lysis buffer was kept in fresh ice until the sample was loaded. The lysate buffer tube was then immediately placed in dry ice (keeping the samples on dry ice should not exceed two hours) or kept at -80°C for long-term storage.

Table 2.2: Reagents used for lysis buffer preparation

Reagent	**Volume**
Triton (0.4%)	19 µl
RNase inhibitor	1 µl
Total	20 µl
0.4% Triton + 1ul X RNase inhibitor	2.4 µl
dNTPS(10 mM)	1 µl
Oligo dt (10uM),	1 µl
ERCC Spike in (-80C)	0.06 µl
Total	**4.45 µl**

2.5.2. Reverse transcription (RT)

All the lysates of the oocytes and cumulus cells were converted into single-strand cDNA by reverse transcription. As 0.5 µl of sample, either oocyte or CCs sample was loaded onto 4.45

µl of lysis buffer. About 5 µl of this lysate (template) was added to 10 µl of reaction mixture, which included 2 µl SuperScript II first strand buffer (5X, Invitrogen, cat.no 18064-014), 0.05 µl DTT (100mM Invitrogen, cat.no 18064), 2 µl Betaine, 0.01 ul MgCl2 (1M), 0.25 µl l RNaese inhibiter (40U/ µl), 0.01 µl TSO (100uM) and 0.025 µl Superscript II RT (200U/ µl) (Table 2.3).

2.5.3. Reverse Transcription reaction

The protocol consisted of incubating the samples at 42°C for 90 minutes, followed by 10 repeated cycles of 2 minutes at 50°C, followed by 2 minutes at 42°C. The reaction temperature was then raised to 70°C for 15 minutes.

Table 2.3: Reverse transcription mix

Reagent	**Volume (µl per reaction)**	**Final Concentration**
SuperScript II first strand buffer (5x)	2 µl	1x
DTT (100mM) (freezer C.R -20)	0.5 µl	5mM
Betaine (5M) (C.R.4)	2 µl	1M
MgCl2 (1M) (C.R.4)	0.1 µl	0.01M
RNase inhibitor (40U/ul) (C.R.-20)	0.25 µl	1U/ µl
TSO (100 µlM) (-80)	0.1 µl	1 µlM
Water (RT-PCR grade)	0.79 µl	-
Superscript II RT (200U/µl)	0.25	50U
Total	**5.9965**	

2.6. PCR-based Amplification

Following cDNA synthesis, the samples were amplified in reaction volumes of 15 µl of PCR mix, which is added to each sample as indicated in Table 2.4:

Table 2.4: Reagents used for PCR mix reaction

Reagent	**Volume**
KAPA HIFI HS Ready Mix (2X)	12.5 µl
ISPCR primers (10uM)	0.0125 µl
Water	2.375 µl
Total	**15 µl**

2.6.1. Primer sequences

1. TSO (5′-AAGCAGTGGTATCAACGCAGAGTACATrGrG+G-3′) At the 5′ end
2. Oligo-dT30VN (5′–AAGCAGTGGTATCAACGCAGAGTACT30VN-3′).
3. ISPCR oligo (5′-AAGCAGTGGTATCAACGCAGAGT-3′). This oligo acts as PCR primer in the amplification step after RT

2.6.2. Amplification protocol

Samples were run into a protocol that consisted of 3 minutes at 98°C for denaturation followed by steps aimed at annealing at 67˚C for 15 seconds and extension at 72˚C for 15 minutes.

2.6.3. PCR clean-up

The PCR products were cleaned using Ampure SPRI beads (Beackman Coulter, cat.no. A 63881). Briefly, 22 µl of PCR products were incubated with 22 µl of Ampure beads for 8 minutes, and then supernatants were removed without touching the beads. 200 µl of 80% ethanol was added to each PCR tube and removed immediately; this step was repeated twice. Samples were left to air dry for 8 minutes. The beads were then resuspended in 17.5 µl of an elution buffer and left for 5 minutes. Finally, 17.5 µl of the supernatant was collected and transferred into fresh 0.2 ml thin-walled PCR tubes (Thermo-Fisher, UK).

2.7. Quality control

The quantity and integrity of the amplified cDNA were determined by a combination of the following assays: a quantity assessment assay and a quality assessment assay.

2.7.1. Qubit® 2.0 flourometric assay (ThermoFisher, UK)

The Qubit fluorometer rely on using fluorescent dyes to determine the concentration of nucleic acids in a sample. The more DNA, RNA in the sample, the more light is absorbed at specific wavelength. The Qubit assay, fluorometric quantification was used to measure the concentration of cDNA samples using a Qubit dsDNA HS Assay kit (Cat. no. Q32851, ThermoFisher, UK). To prepare the samples, a working solution (200X concentrate in DMSO) was diluted 1:200 in a Qubit® dsDNA HS buffer 2 µl of each cDNA sample was added to the tube containing 198 µl of a diluted working solution. Similarly, 10 µl of Qubit® dsDNA HS Standard 1 (0ng/ µl in TE buffer) and Standard 2 (known DNA, 10 ng/ µl in TE

buffer) were added to tubes containing 190 µl of diluted working solutions. Samples and standards tubes were then vortexed for 2-3 seconds and incubated for 2 minutes.

cDNA concentrations were read using a Qubit® flouromertic instrument. Standard 1 and Standard 2 values were measured prior to the sample measurement to calibrate the instrument. The concentration values were calculated and calibrated automatically by the instrument according to the two standard values. Finally, to calculate the actual cDNA concentrations in the samples (before dilution), the values were calculated manually using this equation: cDNA sample= QF value × 200/X, where the QV value is the value given by the Qubit® 2.0 Fluorometer, and x is the number of microlitres of the sample added to the assay tube (2 µl).

2.7.2. 2100 Bioanalyser assessment (Agilent Technology, UK).

The Agilent Bioanalyzer is based on microfluidics technology that is capable of performing electrophoretic separations of DNA fragments. The electrophoretic assays on the Agilent Bioanalyzer are similar to those in the traditional gel electrophoresis principles but in a chip format. The chip format has significant advantages that allow automated sizing and quantitation information in a digital format. By that, it saves the samples consumption as well as the time needed for the separation of DNA fragments.

Using the Agilent High Sensitivity DNA kit (Agilent, cat. no. 352095), cDNA samples were loaded into a high sensitivity chip (Figure 2.3) as follows.

First, the chip was vortexed with an IKA vortex mixer for 1 minute at 2400rpm. 9 µl of the marker was dispensed into the well marked G, 5 µl of dye-gel mix was loaded slowly (due to its high viscosity) into each sample well and 1 µl of yellow ladder was dispensed into the ladder well. Finally, 1 µl of each cDNA sample was loaded into each well (1-11 well), and the loaded chip was vortexed using an IKA Vortex mixer for 1 minute at 2400rpm. The chip was then run in the Bioanalyser machine.

Figure 2.3 High-sensitivity DNA nano chip, Agilent Technologies

2.7.3. PicoGreen measurement

PicoGreen is an analytical assay for the DNA quantitation that based on the use of a PicoGreen probe which interact with nuclci acid. When the PicoGreen fluorescent probe binds dsDNA, it forms a luminescent complex that assessed in comparison to the free dye in the sample solution. The PicoGreen method (Quant-it PicoGreen Kit, ThermoFisher, cat. no. P7589) was used to quantify cDNA concentrations of the CC samples. First, an aqueous working solution of Quant-iT was prepared by making 200-fold dilutions of the concentrated DMSO (Quant-iT PicoGreen dye) in 1X TE (the 20x TE buffer was diluted to a 1x TE buffer with nuclease free water). Preparation was performed using a plastic container, and it should be stored in the dark at room temperature.

2.7.3.1. PicoGreen assay set-up

About 29 µl of 1X TE was dispensed into each well in a multiwell plate, and 1 µl of each cDNA sample was added to each well, mixed properly and incubated for 3-5 minutes. Then, 30 µl of the aqueous solution was added to each well and was mixed.

The standard curve was made by serially diluting the Quant-iT reagent. About 8 concentrations were considered a good range for a Smart-Seq2 cDNA yield. Therefore, it was prepared in the last row of the multiwall plate as follows: 30 µl of 1X TE was dispensed in each well from well 2 until well 9. Then, 6 µl l of Lambda DNA standard (100µg/mL in TE) was dispensed into well 1 and serially diluted from well 2 until well 8, for which 30 µl was discarded. For instance, 30 µl was transferred from well 1 into the second well and mixed up and down, and then 30 µl was transferred into well 3. Thus, well 9 was DNA-free.

After that, 30 µl of the PicoGreen diluted solution was added to each well (from 1 to 9), and the plate was covered from light and centrifuged in a plate shaker and then incubated for 2-5 minutes. The plate was then read using a fluorescent plate reader. The cDNA value was calculated using the standard curve equation, where X is the sample, and Y is the fluorescent value (intensity read): Y=111091x+1358.5.

2.7.4. TapeStation assay (Agilent, UK)

The Agilent TapeStation system is an automated electrophoresis tool for DNA and RNA sample assessment. The Agilent TapeStation was used to assess and analyse the quality of cDNA samples using High-Sensitivity D5000 reagents (Cat. no.5067-5593) and High-

Sensitivity D5000 screen tape (Cat. no. 5067-5592). The method was applied as follows. The ladder was prepared by mixing a 2 µl High-Sensitivity D5000 sample buffer with a 2 µl High-Sensitivity D5000 ladder in an optical tube strip (8X). Then, 2 µl from each cDNA sample was mixed with a 2 µl High-Sensitivity D5000 Sample Buffer in the same tube strip. Tube strip was then spun down and vortexed using an IKA mixer for 1 minute at 2000rpm. Samples were run in a 2200 TapeStation instrument. Using the 2200 TapeStation Controller Software, cDNA profiles were generated and analysed. (Figure 2.4).

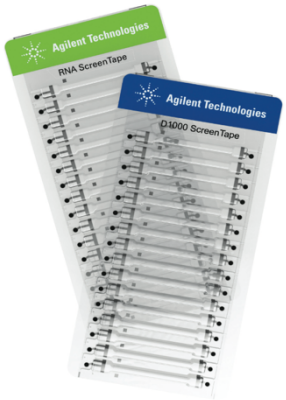


Figure 2.4 High-sensitivity tape screen, Agilent Technologies

2.8. Sequencing Library

2.8.1. Nextera XT construction

The samples were sent to Welcome Trust for Human Genetics for the Nextera XT library preparation (Illumina, UK). This project was the first to analyse a single oocyte, and it was thought that the cDNA concentration of a single oocyte was too high to process using standard input for the single cell Nextera reaction protocol, making the process of preparing

and running the oocytes a bit challenging as may lose the library complexity leading to weak coverage of the transcriptome. Therefore, it has been suggested that each oocyte should be run in triplicate, which allows for the preparation of three Nextera libraries for each single oocyte. Although, technically, the complexity of oocyte library was more likely to be determined during Smart-Seq2 reaction at the early stages of preparation; however, caution was taken to avoid loosing deep coverage.

All samples were tagmented using indices, and the indices were added using the Nextera XT DNA sample preparation kit from Illumina (Cat. no. FC-131-1096). The protocol developed by Illumina was used as follows. For the tagmentation reaction, 1.25 µl of a diluted cDNA sample was added to a 2.5 µl tagmented DNA buffer (1x) and 1.25 µl of an amplicon tagmented mix (1X), making up 5 µl of a volume reaction for each sample. The reaction mixture was incubated at 55°C for 5 minutes. Then, samples were indices by adding 3.75 µl of a Nextera PCR master mix and 1.25 µl each from an index 1 primer (N7xx) and an index 2 primer. The PCR was then run for these samples.

2.8.2. Pooling and multiplexing

72 libraries of the 24 oocytes were pooled into a 72-plex and run in a single lane using Illumina HighSeq 4000 instrument (the standard minimum reads output generated is 270-300 reads for each lane). Similarly, 24 CCs samples were pooled into a 24-plex and run in a separate lane. After the reads of the first two lanes (oocytes and CCs) were generated, the quality of data was assessed. The parentage of duplication and the GC content were scored using FastQC to monitor the sequencing errors. Multiplexing samples in one lane reduced the batch effect that creates technical variations between samples.

Because the data from the first two lanes passed the quality score, both oocytes and CCs were run in the second two lanes to obtain deeper sequencing. A small amount of starting material of RNA and the PCR amplification may limit the depth to which single-cell libraries can be sequenced productively (121).

2.9. Bioinformatics analysis

2.9.1. Sample analysis

The Bioinormatics analysis was performed by the computational group at Weatherall Institute of Molecular Medicine, University of Oxford. Each sample successfully generated an average of 30 million 75-bp paired end reads. Following the QC analysis with the FastQC package (http://www.bioinformatics.babraham.ac.uk/projects/fastqc), the reads were aligned against the mouse genome assembly using STAR (122) (GRCm38 [mm10] UCSC transcripts). Non-uniquely mapped reads and reads that were identified as PCR duplicates using Samtools (123) were discarded. The gene expression levels were quantified as read counts using the featureCounts function (124) from the Subread package (125) with default parameters. The read counts were used for the identification of a global differential gene expression between specified populations using the edgeR package (126).

2.9.2. Normalisation method

To normalise for sequence depth (library size) and transcript length (gene size), the reads counted per kilobase of exon per million (RPKM) method was applied (86). RPKM values

were also generated using the edgeR package. Genes were retained if they expressed a threshold of 1 count-per-million (cpm) for any given sample.

2.9.3. Differential gene expression

The EdgeR package was used for differential gene expression. Genes were considered differentially expressed between populations if they had an adjusted p-value, named false discovery rate (FDR) of less than 0.05.

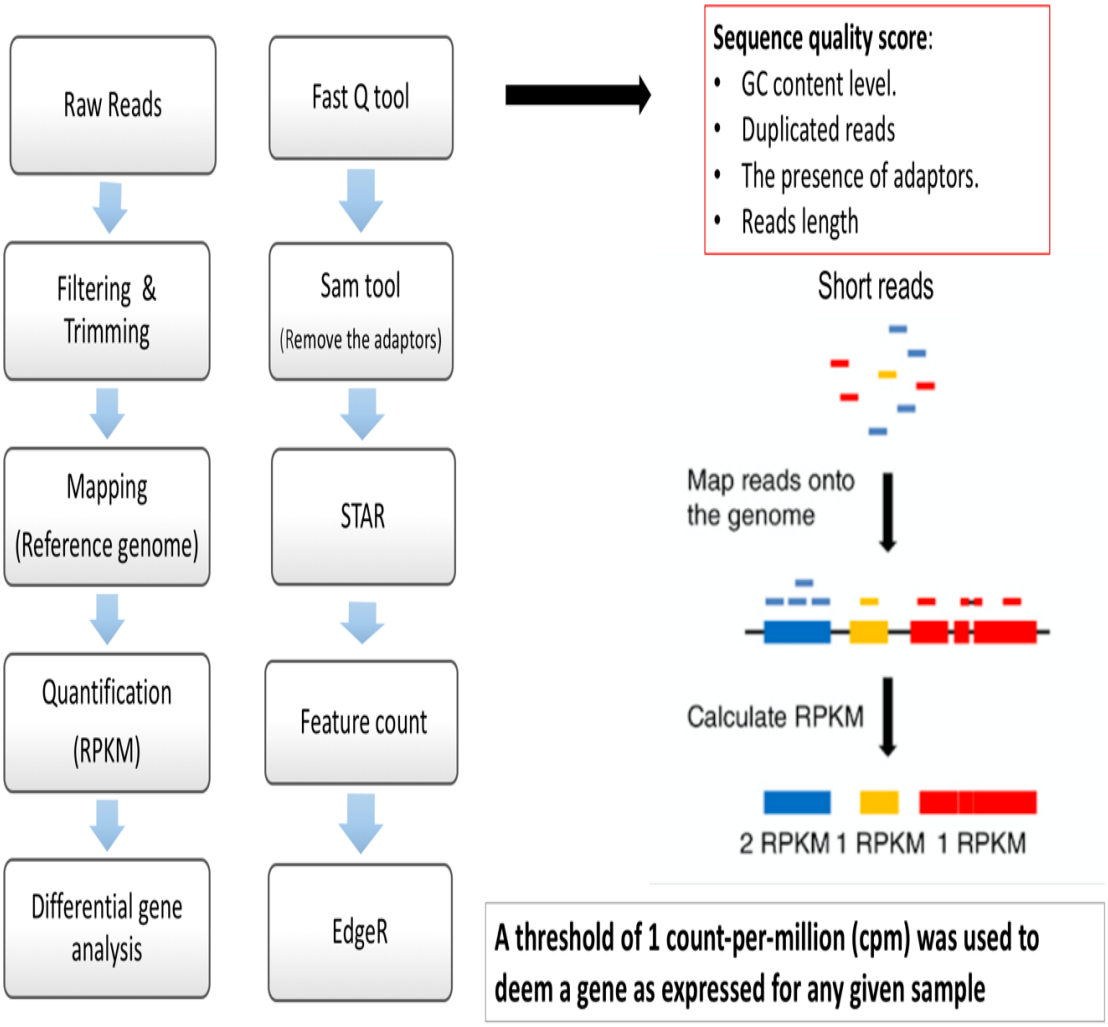


Figure 2.5 Illustration of bioinformatics steps. The schematic diagram illustrates the step of transcriptomic analysis carried in this study. Once the cDNA sequenced, short reads are generated. Then the quality of raw reads is checked for assessing the sequence quality score using FastQ tool. This quality score involves: G-C content, number of duplicate reads, the percentages of adaptors and reads length. The reads were then gone through Sam tool where the raw reads were filtered according to these criteria. Then, short reads were mapped to reference mouse genome using STAR. The aggregated reads were quantified according to RPKM scale using Feature count. Finally, the edge R was used for analysing differential gene expressions.

2.10.Bioinformatics visualisation method

2.10.1. Principle component analysis (PCA)

To investigate the gene expression pattern variations between oocytes at three age points, PCA was used to identify single cell global gene expression patterns with a similar global gene expression in different oocytes. Therefore, sample clustering based on the global gene expressions was performed as a quality assessment for the sequencing data, which is described in detail in chapters 4 and 5.

2.10.2. Functional analysis: Ingenuity Pathways Analysis (IPA)

To identify the pathways and functions associated with significantly changed transcripts, Ingenuity Pathway Analysis (IPA) was used. Transcript gene identifiers were mapped to their corresponding gene objects in the IPA knowledge database (http://www.ingenuity.com). The significant values for the canonical pathways were calculated based on a right-tailed Fisher's exact test which calculates the probability of association between transcripts.

Differentially expressed genes were enriched for a pathway analysis using Metacore software (Thompson router). An FDR <0.05 was considered the threshold for a significant enrichment of the genes, and the pathways were deemed to be enriched if the enrichment over background was at least 2-fold.

2.11.OpenArray real-time PCR validation

The differentially expressed genes were validated using the OpenArray Taqman Real-Time PCR. This new method used for analysing high throughput data because it provides multiplexing, which means it is capable of accommodating more than 3000 reactions per array. Assays panels were designed by ranking differential genes (DG), considering FDR and fold changes as the main metrics for ranking differential genes based on the sequencing results. All details of validation are explained separately in a chapter 8.

Chapter 3

Optimising a Preparation Protocol for Single-Cell RNA Sequencing Libraries.

3.1. Introduction

The advent of high-throughput approaches has revolutionised transcriptome analysis studies. These methods have the potential to examine and uncover the entire transcriptome such that that every single molecule is surveyed and quantified (127). Indeed, high-throughput techniques reveal the functional elements of the genome and provide essential insights into the molecular components constituting a specific development stage or other physiological condition. In turn, knowledge of changes in transcript expression levels facilitates the process of interpreting the factors that lead to a defect or disease.

Single cell-RNA-sequencing (sc-RNA-seq) has become a powerful method for high-throughput transcriptomic analyses of cell dynamics, providing rapid, straightforward, and more importantly, cost-effective technology (128,129). sc-RNA-seq has made substantial contributions to the understanding of transcriptome complexity and transcription regulation. However, this method must become even more sensitive, if the goal is to detect rare transcripts in single cells, particularly in technical and bioinformatic aspects which are still undergoing on-going improvements. Sc-RNA-seq is expected to replace most transcriptomic technologies whilst the cost declines.

The main objective of this chapter is to present the process of optimising a preparation protocol of single-cell cDNA libraries that allows conducting sc-RNA-seq. These cDNA libraries were optimised to enable the investigation of the impact of age on the transcriptomic profiles of the cells of interest, single mouse oocytes and their companion cumulus cells. The main purpose of the amplification strategy was to identify the minimum number of PCR cycles required to generate a sufficient quantity of cDNA. The goal was to avoid biases associated with high numbers of PCR cycles.

3.2. Method

The initial step in obtaining the transcriptome of a single oocyte is to successfully isolate the individual oocyte from surrounding CCs and process them separately. Then, recovered RNA from each oocyte and CCs has to be sufficient in order to generate an adequate amount of cDNA library. This is known to be the greatest technical challenge, particularly at the single-cell level, as most transcriptomes studies are still conducted on a population of oocytes.

3.2.1. Oocyte and CCs Preparation

Ovaries were collected as described in section 2.2 Briefly, cumulus-oocyte complexes (COCs) were collected from pre-ovulatory follicles about to ovulate, since these follicles are assumed to contain an egg at the last stage of maturation (GV).

Each COC was washed in 10 µl of PBS droplets, and the oocyte was then stripped from the CCs by pipetting up and down until complete denudation was achieved. To obtain a complete, clean CCs transcriptome using a new capillary; a 10ul drop of PBS containing cumulus cells was then transferred into a 0.2 ml into a thin-walled PCR tube (RNAase free, Thermo Fisher, UK). The CCs were centrifuged at 120g for 3 minutes and the supernatant was discarded. Then, 0.5 µl of CCs pellet was transferred into 4.45 µl of lysis buffer. The samples were vortexed and centrifuged at 2000g (Figure 3.1). The lysates were then frozen and kept at -80°C. it was necessary to achieve rapid processing of the samples to maintain good physiological conditions within the ovaries. A lengthy procedure could have resulted

in the loss of a portion of the cell population (such as in CCs), thereby altering the gene expression or causing cell death.

Although, isolating single oocyte and CCS methods were optimised at this stage, the efficiency of the two methods could not be determined until samples were converted to cDNA and amplified. Below, these steps are described.

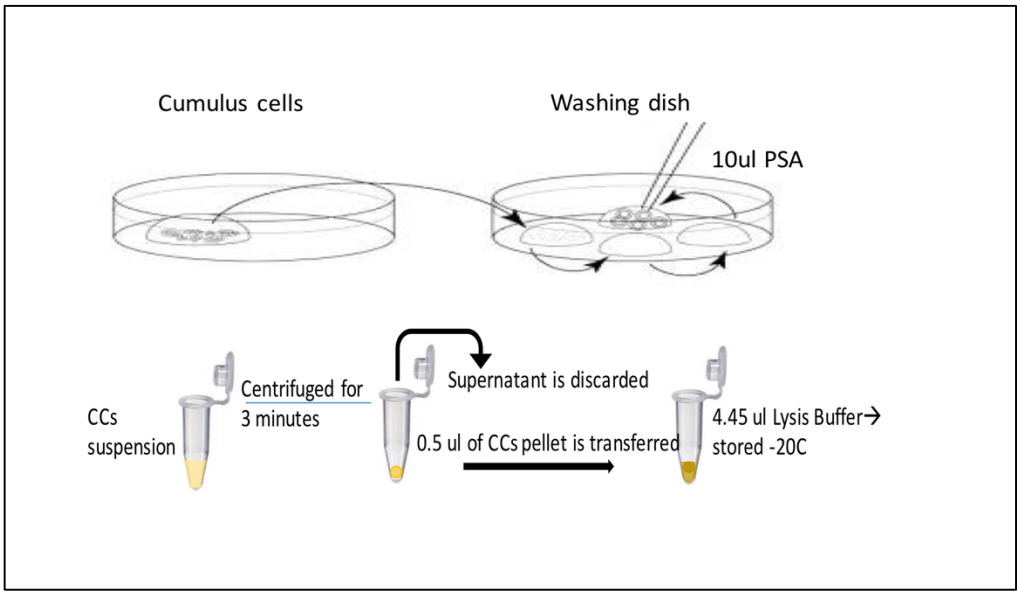


Figure 3.1 Flowchart of CCs preparation steps.

3.2.2. PCR-based Amplification

The oocyte and cumulus cell samples were therefore segregated based on the number of cycles that had been applied. The number of cycles is given for each pilot study.

3.2.3. cDNA Assessments

First, the quality of the cDNA was assessed using the Agilent Bioanalyser 2100 DNA high-sensitivity chip (Agilent Technologies, UK). An ideal amplification profile has a sharp

amplification peak with a distribution size of 1.5 to 2kb lying between lower marker and upper markers, with the ideal size between 50 to 200pb. Second, the quantity (concentration) of the cDNA was measured using the Qubit® flourmetric assay 2.0 (Thermo Fisher, UK).

3.3. PCR Amplification protocol

3.3.1 Pilot 1: cDNA yield generated from a single oocyte and its companion CCs

The success of an RNA experiment depends on the quality and integrity of the recovered RNA. In particular, when starting with low quantities of RNA, obtaining sufficient RNA to analyse entire gene expressions for single cells is challenging. The sensitivity, accuracy and precision of transcript quantification strongly depend on how the mRNA is converted into the cDNA that is eventually sequenced. Therefore, it is extremely crucial to use an efficient RNA extraction protocol with high sensitivity in order to capture the minutes' amount of RNA. The aim is to optimise an efficient non-biased amplification strategy that would allow generating adequate quantities of cDNA yield from single oocyte and their companion CCs.

3.3.2. Aim

To identify whether individual mouse oocytes and CCs can generate sufficient material for a cDNA library for RNA sequencing.

3.3.3. Method

Seven oocytes and 11 CCs samples were collected from 4-week-old female C57B/6 mice and prepared separately, as described earlier. A range of cycles were applied to determine

the best amplification condition for each cell. The oocytes were amplified with 18, 20, 22 and 24 cycles, whereas 16,18, 20, and 22 cycles were selected for cumulus cell amplification. The samples were grouped according to the number of PCR cycles chosen.

3.3.4. Pilot 1 result

3.3.4.1. The Optimal amplification for single mouse oocytes

The quality of the amplified cDNA was assessed with the Agilent 2100 Bioanalyser DNA high-sensitivity chip (Thermo Fisher, UK). Different cycle numbers were applied to the oocytes to amplify the cDNA. As expected, the single oocytes generated varying quantities of cDNA, and the volume increased with the number of cycles (Figure 3.2). Of the cDNA profiles of the 7 oocytes, oocyte 6 (Figure 3.3) appeared to have the ideal amplification peak. This oocyte generated a sharp, broad peak with a distribution size of 1,555bp lying between the upper and lower markers, both of which were clearly defined (approximately 100bp). Although some background noise (short fragments) existed which potentially are a result of an excess of primer dimer, when 18 cycles were used the single oocyte generated high-quality cDNA. This finding thus supported a recent study conducted by Reyes, in which the oocyte amplified at 18 cycles (130).

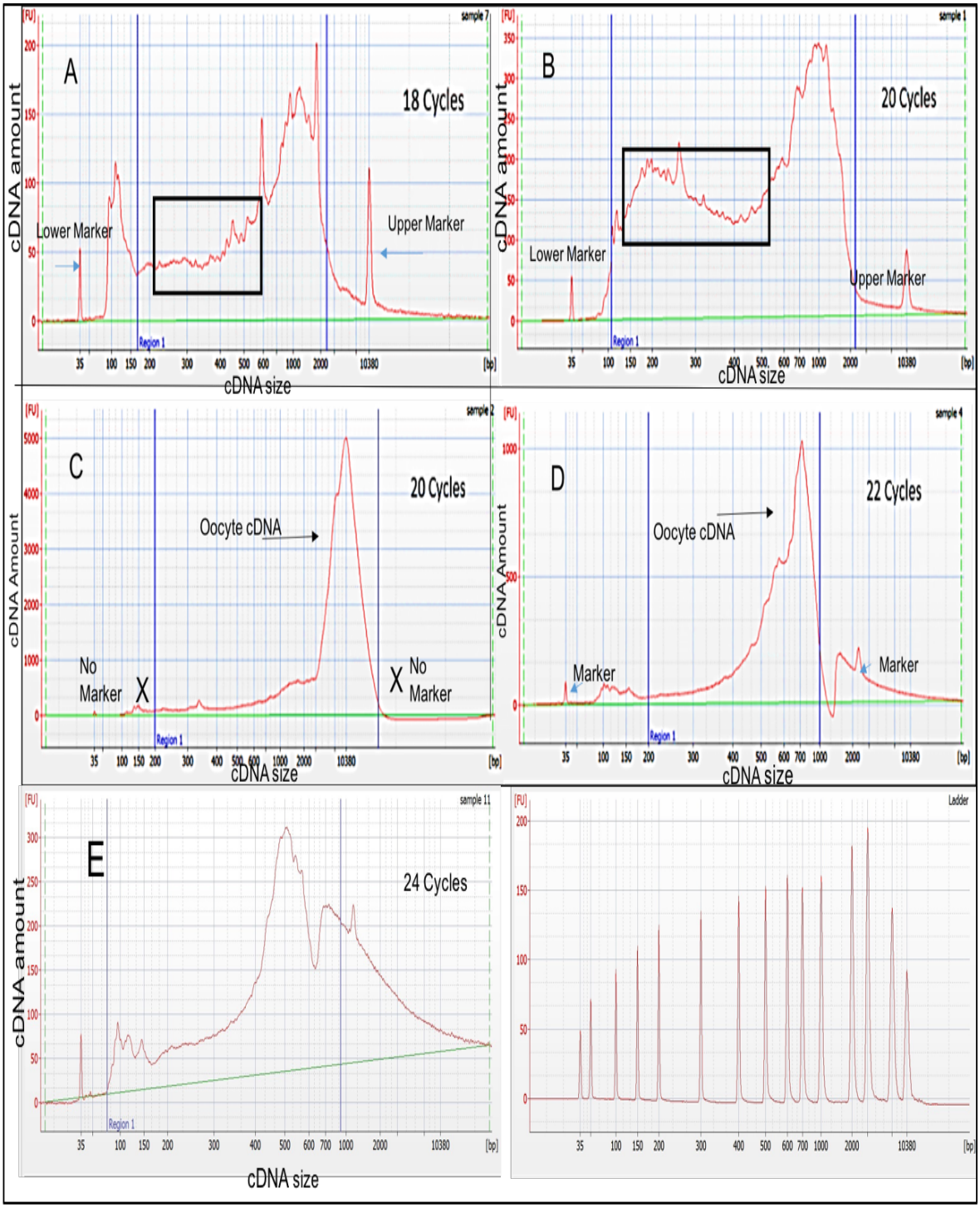


Figure 3.2 Bioanalyser illustrating the amplification profiles of four single oocytes: (A) a good amplification profile, except that the low molecular weight cDNA is partially degraded (black square), (B) cDNA peak is over-amplified with degradation of the low molecular weight cDNA (black square), (C) cDNA profile displays over-amplification with the two markers not detectable, and (D & E) examples of over-amplification in single oocytes amplified by 22 and 24 cycles.

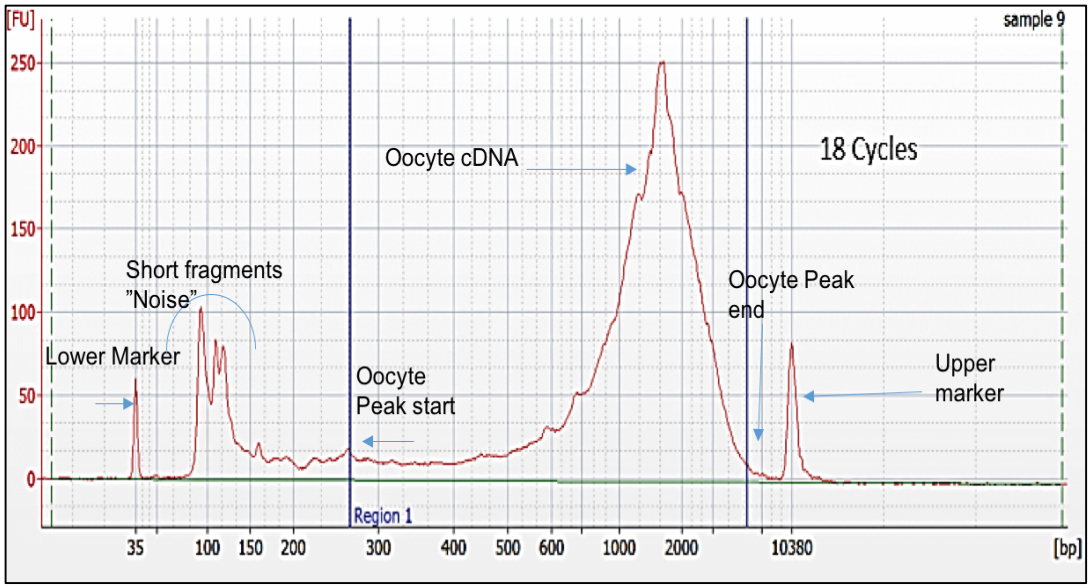


Figure 3.3 Electropherogram of oocyte 6 amplified with 18 cycles. cDNA. The generated peak was sharp and free of fragments.

3.3.4.2. The Optimal amplification for the CCs

Figure 3.4 displays an electropherogram of the cumulus cells amplified with different numbers of PCR cycles. Of these CCs amplification profiles, the CC samples indicated in Figure 3.5 displayed an optimal amplification profile. The broad cDNA peak with a distribution size of 1,280bp and well-defined markers indicated that 16 cycles of PCR seemed to yield the best result over the rest.

Cumulus cells samples amplified with 20 PCR cycles generated an over-amplified cDNA, the peak of which was highly digested and for which only the upper marker was defined (Figure 3.4A). Also, CCs samples amplified with 20 and 22 cycles (Figure 3.4B and D) revealed over-amplification and high levels of technical noise. Since all CCs samples were processed similarly, it was concluded that 20 and 22 cycles of PCR resulted in extensive amplification.

Short fragments were widely distributed in samples amplified with 18 cycles, which indicated completely degraded RNA of the CCs samples (Figure 3.4C). The two CCs samples (Figure 3.4A and C) are examples of heavily digested RNA regardless of the amplification, which can potentially be attributed to lengthy sample processing.

The critical goal of this project was to isolate the oocytes from their cumulus cells and process each of these separately, and this task was highly technically challenging because certain points had to be considered. First of all, it needed to be ensured that the CCs (only the layers that surround the oocyte) were free of other ovarian cells. This required careful and gentle washing by pipetting in and out using a micro capillary with a wide diameter to

keep COC intact as aggressive wash dissociates CCs and results in unwanted mixing of the layers of CCs with other follicle cells. The second point that had to be considered was separating the oocyte from CCs, quickly, and keeping each in separate PBS drops. It was suggested that this long preparation procedure might have resulted in RNA degradation, which existed in some CCs samples but was not seen in oocytes samples. Therefore, some modification for processing time was required to maintain RNA integrity in CCs samples.

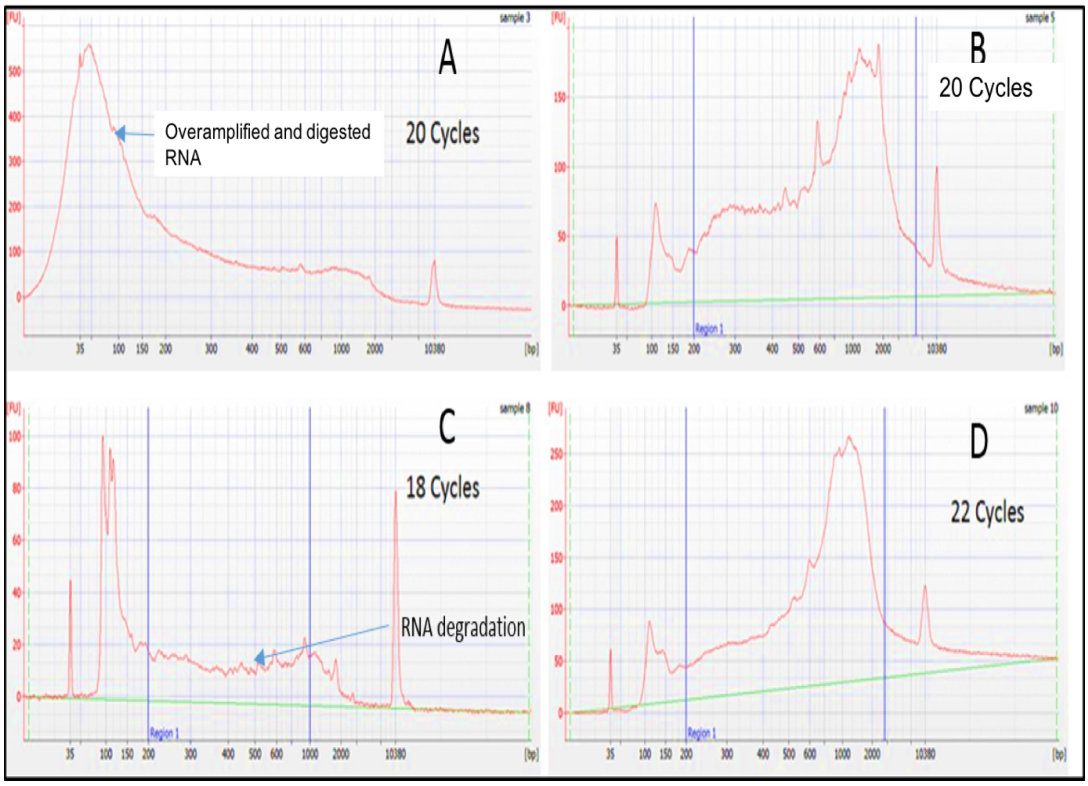


Figure 3.4 Bioanalyser electropherograms of the pre-amplified cDNA of the CCs.

(A) the profile reveals over-amplified and digested cDNA, with the upper marker undefined; (B) a highly concentrated cDNA profile with partial degradation of low molecular weight cDNA; (C) a wide distribution of fragments dominated by primer dimer, indicating RNA degradation; and (D) a large quantity of cDNA generated by 22 cycles.

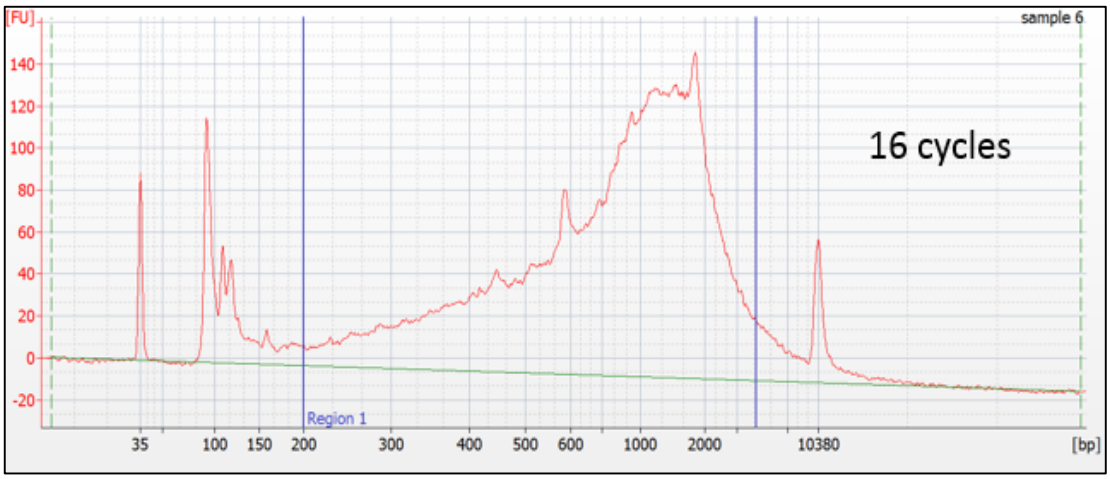


Figure 3.5 Bioanalyser analysis of the cDNA generated by the cumulus cells amplified with 16 PCR cycles

3.4.2. Pilot 2: cDNA Yield of Aged Oocytes and Cumulus Cells (1-year-old Mice)

3.4.2.1. Aim

The oocytes from the aged mouse had fewer cumulus cells that apparently surrounding the aged oocyte. The aim of this study is to determine whether these aged cumulus cells had similar cDNA yields as the young mice.

3.4.2.2 Method

The following samples were collected from female C57BL/6 mice: 14 oocytes and their CCs (14 samples) from a 1-year-old mouse (aged mice);8 oocytes and their CCs (8 samples) from 9-week-old mouse (mature mice).

Oocytes and CCs were prepared in a similar method as applied in Pilot 1. Samples were then reverse transcribed followed by cDNA amplification. However, some additional amendments were considered for the preparation method of CCs to improve the quality of the cDNA yield:

1. The task of handling the samples was challenging in the previous experiment because cells stuck to the glass capillaries' walls, leading to the loss of some oocytes during the collection. Therefore, Bovine Serum Albumin (BSA; Sigma-Aldrich, UK) was added to the PBS to avoid cell stickiness. BSA has been used in other studies and does not appear to affect the amplification efficacy (131).

2. An additional washing step was added to the cumulus cell preparation process to improve the quality of the CCs cDNA yield. This addition was necessary because the cDNA profiles had demonstrated a moderate level of noise (Figure 3.6). Therefore, an additional 10 µl of PBS/BSA was added to the cumulus cell suspension (10 µl) comprising a total of 20 µl of CCs suspension, which was then transferred to a 0.2 ml thin-walled PCR tube and centrifuged for 3 minutes at 120g. Finally, the supernatant was removed, and only 0.5 µl of the residual cells was transferred and placed in the lysis buffer.

3.4.3. Pilot 2 result

3.4.3.1. Amplification result of old oocytes

Two oocytes, one oocyte representing each age group, were selected for analysis with the Agilent 2100 Bioanalyser. The oocyte from the 1-year-old mouse yielded a similar amount of cDNA as the oocyte from the 9-week-old mouse (Figure 3.6). Both samples resulted in comparable amplification peaks, indicating that oocytes of both ages yield analogous quantities of RNA.

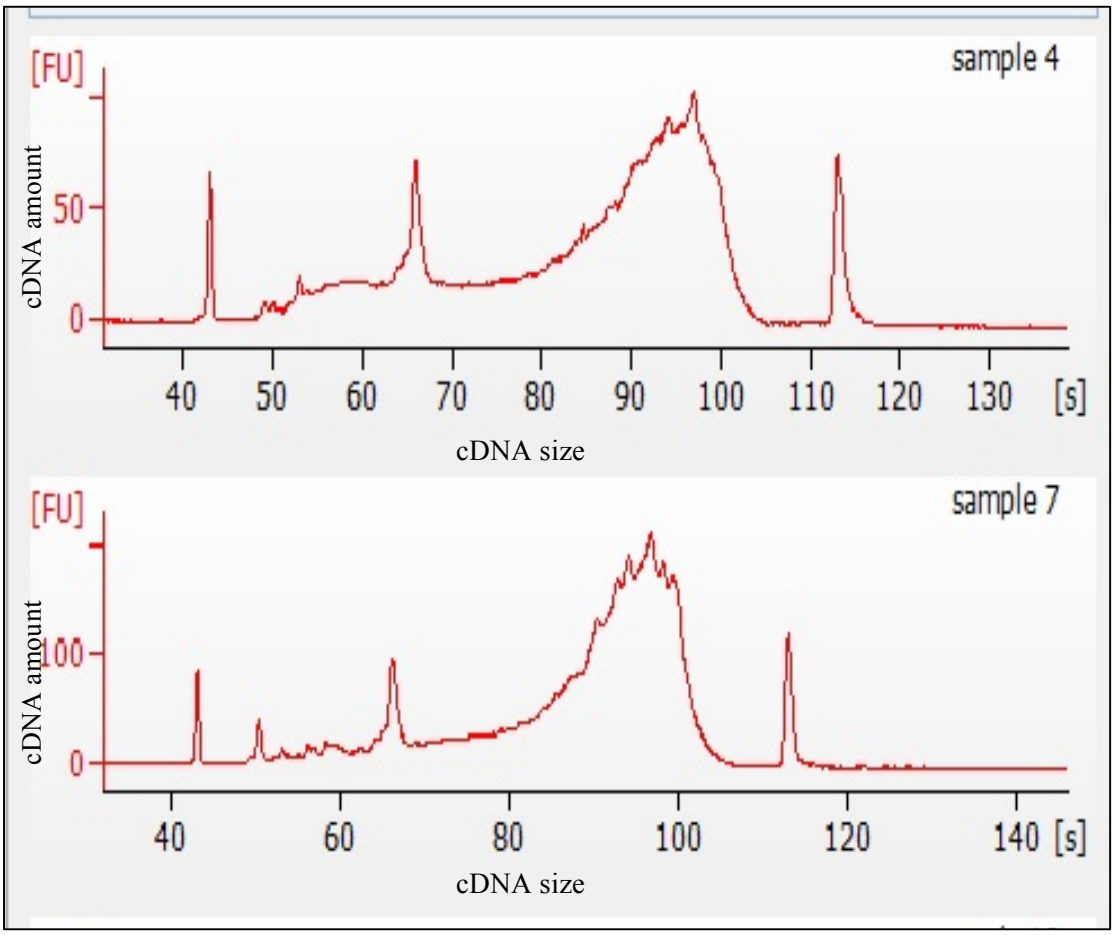


Figure 3.6 Amplification profiles of (A) one-year oocyte and (B) nine weeks oocyte. Both oocytes were amplified with 18 PCR cycles. Oocyte from the 1-year-old mouse yielded a similar amount of cDNA as the oocyte from the 9-week-old mouse.

3.4.3.2. Cumulus cells amplification

All of the CCs cDNA yielded outcomes that were dramatically lower than those from the previous pilot study, implying that the insufficient amount of RNA led to amplification failure (Figure 3.7).

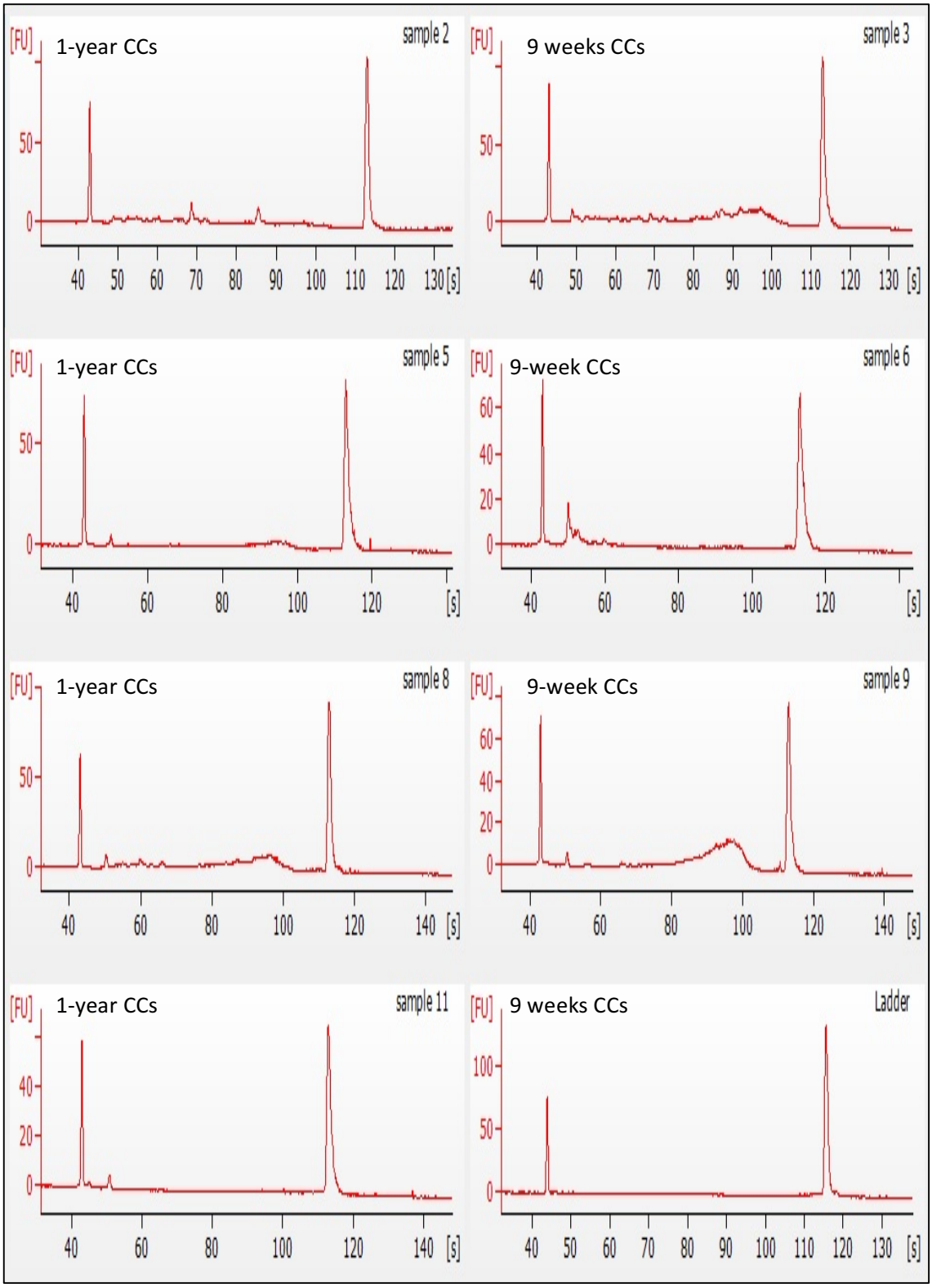


Figure 3.7 Representative examples of samples in which cumulus cell amplification was inefficient. Most of the samples failed to generate cDNA. Results were assessed by the 2100 Agilent Bioanalyser.

3.5. Pilot 3: different cumulus cell preparation methods

3.5.1. Aim

Due to the inconsistent results of the second pilot, Further optimisation was necessary before collecting samples for the actual RNA experiment. The goal was to determine which preparation method produced both higher quality, and sufficient cDNA yield.

3.5.2. Method

The previously mentioned preparation methods (pilot 1 and pilot 2) were applied using 10 cumulus cell samples (Table 3.1). The lysate samples were then reverse transcribed followed with PCR amplification. cDNA samples were then assessed by the Agilent bioanalyser DNA sensitivity chip (Agilent technologies, UK)

Table 3.1: Summary of the Cumulus Cells Preparation Methods

	Pilot 1 method				Similar to Pilot 1 with additional wash step		
PBS only	CCs 1	CCs2	CCs3	CCs4	CCs5	CCs6	CCs7
Wash step	No wash	No wash	No wash	No wash	√	√	√
	Similar to Pilot 2 without additional wash step				Pilot 2 preparation method		
BSA/PBS	CCs8	CCs29	CCs10	CCs11	CCs12	CCs13	CCs14
Wash step	No wash	No wash	No wash	No wash	√	√	√
When oocyte is stripped from CCs, CCs are kept in 10ul media drop (PBS or BSA/PBS). This 10 µl CCs suspension were transferred into 0.2 ml PCR tube, centrifuged at 120g for 3 minutes. If washing step is added: 10 µl of PBS is added to the 10 µl of CCs suspension (total is 20 µl), centrifuged at 120g for 3 minutes, supernatant is then discarded, then 0.5 µl of pellet is transferred into lysis buffer.							

3.5.3. Results of pilot 3

The bioanalyser traces did not reveal any substantial differences between the cumulus cell samples prepared with and without BSA. The samples prepared with BSA actually had better amplification profiles, although with slightly degraded RNA. This demonstrates that using BSA improved the cDNA yield (Figure 3.8). Figure 3.8B illustrates suboptimal amplification in which the upper and lower markers are difficult to detect. Although this amplification is not ideal, it confirmed that the addition of the BSA was not the reason for the suboptimal cDNA profiles of the cumulus cells in Pilot 2 (Figure 3.7). Therefore, it can be surmised that the washing step affected the cell concentrations, potentially explaining the

failure of the CCs in Pilot 2. It is possible that the cell suspension was highly diluted. In other words, in the second pilot, the CCs were suspended in 20 µl of PBS/ BSA and only 0.5 µl of the volume was used for the RT-PCR, whereas 0.5 µl of 10ul was used in Pilot 3.

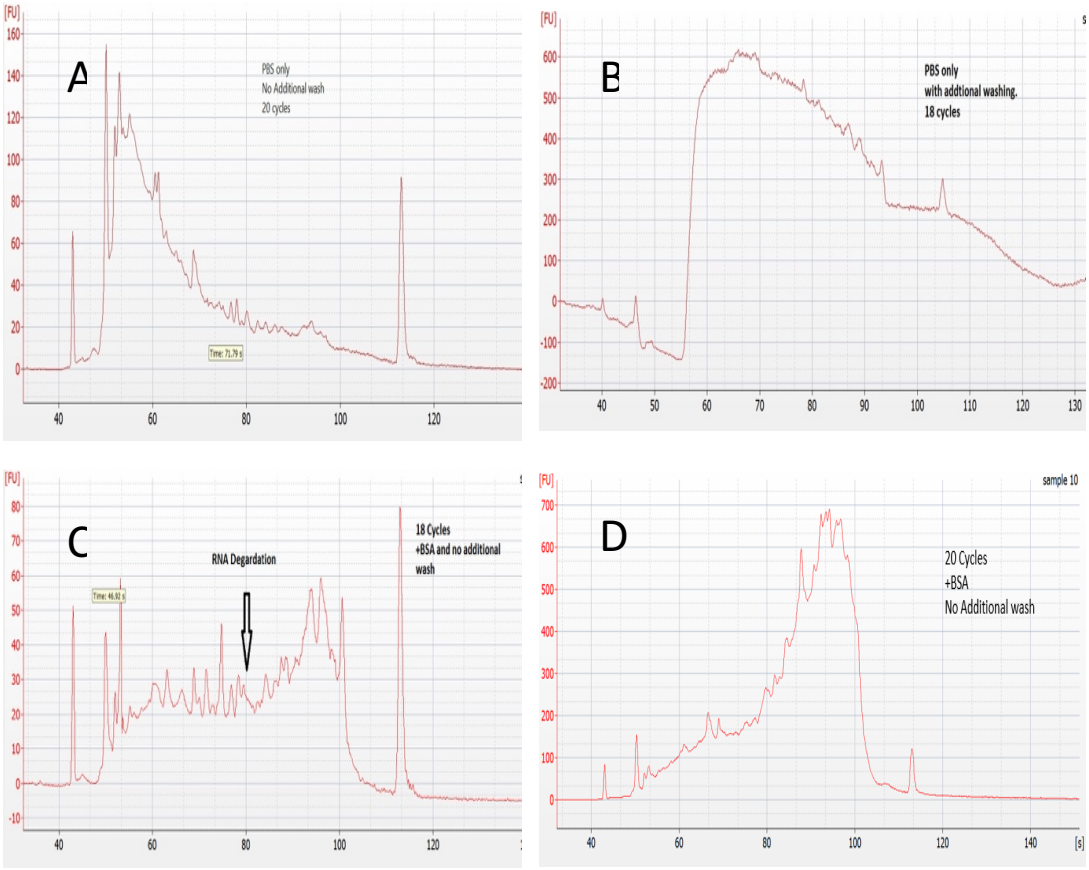


Figure 3.8 Sample electropherograms representing the cDNA generated from CCs prepared in PBS without BSA (A and B) and with BSA/PBS (C and D): (A) sample profile reveals a digested cDNA peak shifted to the right; (B) bioanalyser result showed over-amplified cDNA, and markers could not be detected; (C) the cDNA profile shows a suitable amount of cDNA but suffers from RNA degradation; and (D) representative successful cDNA amplification, with few fragment distributions.

3.6. Pilot 4: The effect of the collection time on RNA integrity

3.6.1. Aim

The aim of this pilot was to investigate whether the time of collection and processing influenced RNA integrity.

3.6.2. Method

Ten cumulus cell samples were used: four CC samples from a 1-year-old mouse, and six CC samples from a 3-week-old mouse. The selected samples were the last samples processed on each collection date. The purpose of this pilot was to ascertain if these samples had any RNA degradation.

As the cumulus cell preparation method was adjusted multiple times, the final iteration was as follows: Each COC was placed in 5 µl of PBS/BSA and then the oocyte was processed separately, as previously described. A 5 µl drop of the PBS/ BSA containing the cumulus cells was transferred into 0.2ul tube (RNAase free, Life Technologies). It was centrifuged at 120g, and the supernatant was then discarded. Finally, 0.5 µl of the cell residue was transferred into 4.45 µl of lysate buffer. These samples were then reverse transcribed, followed by PCR amplification, cycles are indicated in Table 3.2. The cDNA concentrations were measured with the Qubit®2.0 fluorometer and the quality was assessed using the Agilent 2100 Bioanalyser.

3.6.3. Pilot 4 Results

cDNA Concentrations from the CCs Samples were measured using the Qubit 2.0 fluorometer method (Table 3.2). Most of samples have generated enough amount of cDNA; All samples generated a concentration range of cDNA between 0.23 and 2.2 ng/µl which is within the required concentration for generating Illumina sequencing library (discussed in Chapter 4)

The bioanalyser results demonstrated that most of the CC samples produced successful cDNA profiles showing good quality peaks. The exception was sample 1, which failed to generate cDNA (Figure 3.9). Although the cDNA concentration measurements in Table 3.2 agreed with the cDNA profiles, a very low concentration was detected in sample 8. Although the cDNA peak was broad and reflected a high concentration, this might have been due to a pipetting error when loading the samples into the bioanalyser chip.

Table 3.2: cDNA Concentration from the cumulus cells samples measured using Qubit 2.0 Flurometer.

Sample number	Age	PCR tube number	Cycle number	cDNA concentration
CC1	1 year	1	15	0.23ng/ µl
CC2	1 year	2	15	0.848ng/ µl
CC3	1 year	3	16	0.572ng/ µl
CC4	1 year	4	15	0.641ng/ µl
CC5	3 weeks	5	14	0.634ng/ µl
CC6	3 weeks	6	14	0.691ng/ µl
CC7	3 weeks	7	15	0.582ng/ µl
CC8	3 weeks	8	15	0.083ng/ µl
CC9	3 weeks	9	16	0.541ng/ µl
CC10	3 weeks	10	16	2.2ng/ µl
Negative control	3 weeks	11	16	0ng/ µl

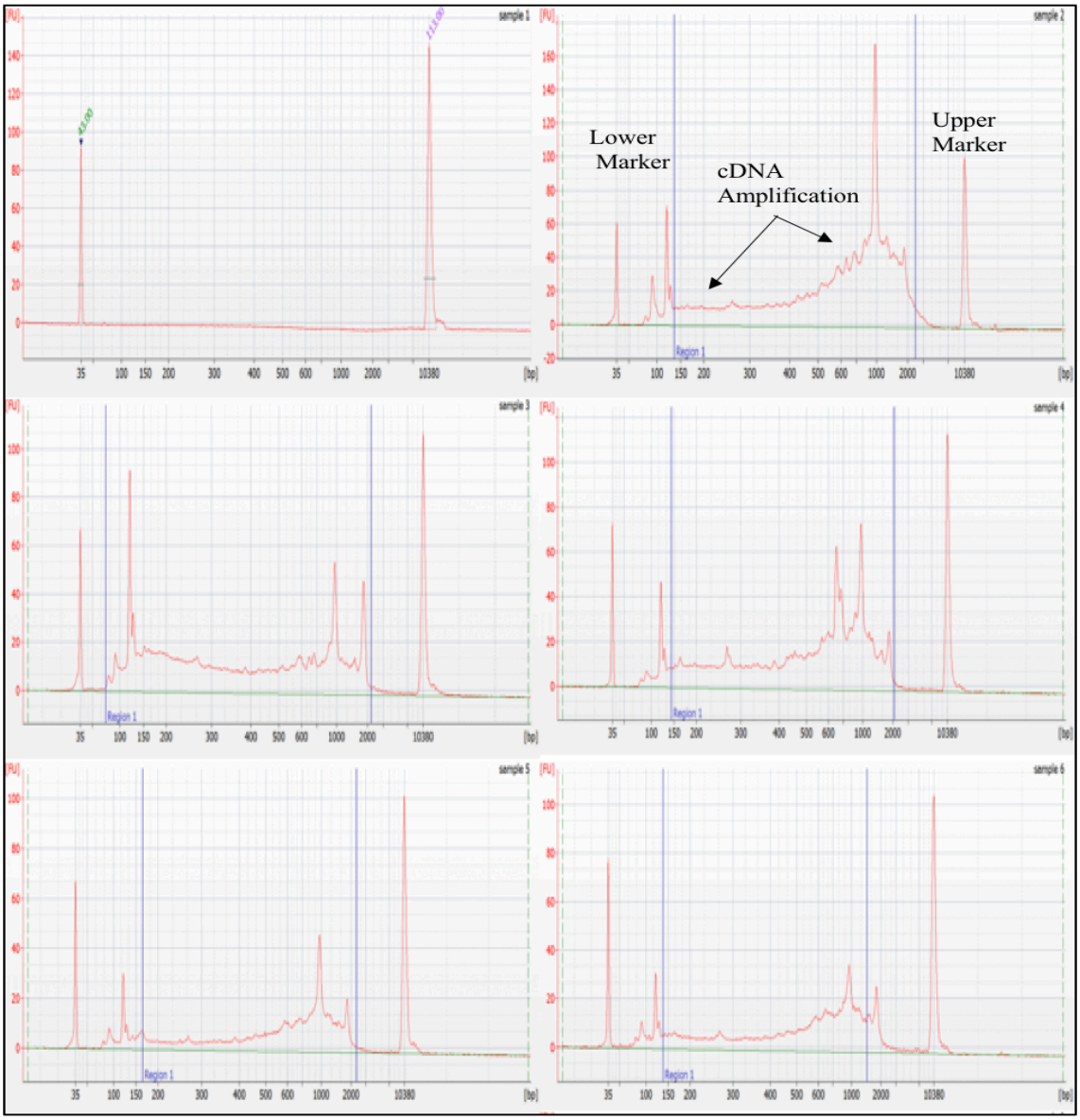


Figure 3.9 The cDNA yields from the cumulus cells assessed via the 2100 Agilent Bioanalyser.

In Sample 1, the bioanalyser failed to show amplification. Sample 2 displays a good cDNA profile, with the exception of the mechanical spike attributable to the PCR artefact. In Samples 3 and 4, heat denaturation was needed to improve the yield. Samples 5, 6 produced sufficient cDNA.

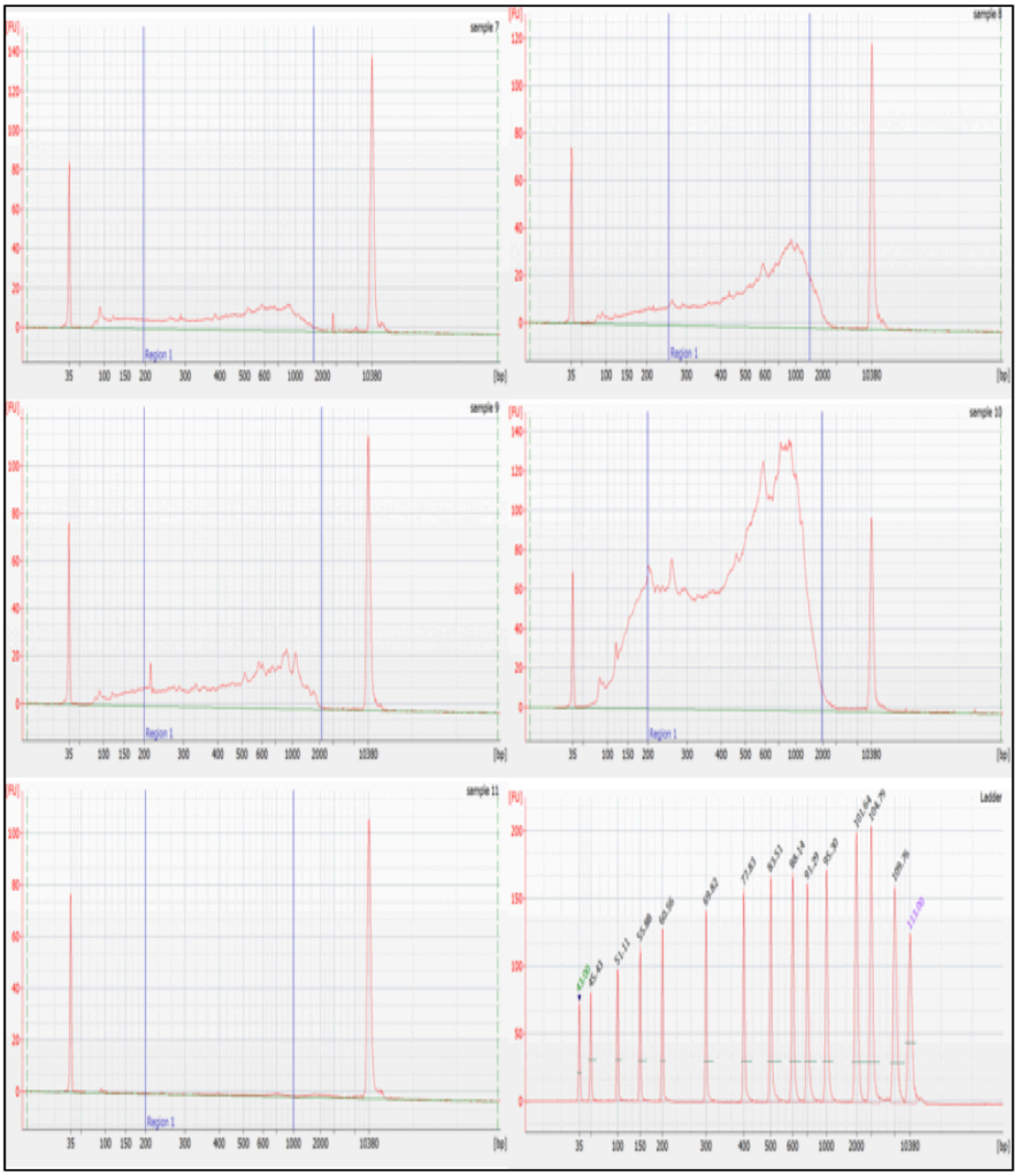


Figure 3.10 The cDNA yields from the CCs.

Sufficient cDNA from sample 7,8 and 9, although there were mechanical spikes due to the PCR artefacts. Sample 10 demonstrates an overabundance of amplified cDNA displaying degradation.

3.7. Discussion

The advent of RNA sequencing has allowed transcriptomic studies to thrive with unprecedented sensitivity and resolution, leading to a deepening understanding of the transcriptome's complexity and the discovery of new, small RNA species, such as microRNA and short-interfering RNA (132). RNA sequencing has played a role in revealing these hidden factors in the cellular RNA pool because these small RNAs are very short to be captured by the microarray approach. Rather, they were discovered using NGS platforms with higher throughputs as part of a genome-wide survey (133,134).

The first potential reason for the failed amplification in pilot1 was thought to be the addition of the BSA to the PBS. The PBS/ BSA was used to prepare both the oocytes and the cumulus cells, and so the high-quality cDNA yield that the oocytes efficiently generated was not affected by the addition of BSA. In contrast, it improved the oocyte processing (avoiding mechanical cell damage) and prevented potential RNA degradation that results from long procedures. The CCs samples were more likely to be highly diluted when prepared in a total of 20 µl compared to the 10 µl CCS suspension used in Pilot 1. The result was the loss of RNA and amplification failure.

Another potential possibility was the centrifugation speed was not optimal for the volume of the CCs suspension used in pilot 2. When centrifugation was optimised for the protocol, different speeds were tested to choose the optimal speed for pelleting the cells down without harming the integrity of the cells' RNA. Therefore, it might that the cells did not pellet down. Minor technical issues seemed significant, and have serious impact on cell preparation.

Pilot 3, which compared modification steps of the preparation method, demonstrated that the cumulus cells could generate cDNA if they remained in adequate concentrations. This finding confirmed that the BSA had not interfered with the reverse transcription or PCR reaction. Therefore, the more likely explanation is that the cells were not concentrated sufficiently in order to obtain adequate quantities of RNA, as it is well known that a low quantity of starting materiel (estimated at 0.1pg of mRNA in each cell) leads to insufficient cDNA volumes. Based on that, adding BSA to the PBS (1mg/ml) does not appear to inhibit the RT-PCR. Rather, this method maintains the samples' integrity, as it prevents undesired delays. In addition, it protects the cells from mechanical damage. This is particularly true for older mice; samples from such mice are valuable and difficult to source. Therefore, scaling the volume of the cumulus suspension down and assessing if this improves yield is suggested for the next pilot experiment.

The consistent results achieved in pilot 4 indicated that the final preparation was efficient and resulted in desired quality and quantity. The small differences in number of amplification cycles (15 or 16 cycles) did not influence the amplification outcome as much as the concentration of starting material. This was evident when CCs were concentrated in 5 µl of BSA/PBS in Pilot 4 and 0.5 µl of CC pellet (added to 4.5 µl lysis buffer) was used. The amplification of these samples generated the desired quality and quantity needed for the subsequent step, Nextera XT reaction, which is a standard method that requires only 0.2 ng/µl. The cumulus cell results from all four experiments illustrate a variety of amplification profiles. These variances stemmed from the various preparation protocols

Library preparation, which begins with template preparation, is a key step for successful RNA sequencing. Its importance lies in the fact that it reflects how closely the cDNA sequencing outcome reflects the original RNA population. The most critical issue is to obtain unbiased cDNA. Although existing RNA protocols only comprise a few steps, several manipulations are necessary when producing cDNA libraries. Each step of library construction (RNA extraction, cDNA synthesis, pre-amplification, library fragmentation) can create a bias in the final sequencing library. Such biases can affect the quality of both the data and the final analysis, leading to decreased sensitivity of transcript quantification.

Significant effort has been put into improving library preparation strategies to avoid potential biases in transcriptome coverage. Therefore, as a first step, recovering high-quality RNA is an essential element for the success of library construction. Recently developed approaches allow for the selection of the RNA type before an RNA sequencing experiment, and these techniques have permitted transcriptomic studies with more focused aims. One of the main aspects of experimental design is the RNA extraction protocol used for selecting the RNA type. Most standard protocols rely on selecting polyadenylated mRNA and depleting ribosomal RNA, which constitute approximately 90% of all cellular RNA (135). It is estimated that each mammalian cell contains 10-20 pg, of which approximately 0.1 pg is mRNA (136). For this reason, dealing with a negligible quantity of RNA has been always a source of concern in single-cell gene expression profiling studies, because material loss and low biochemical reaction efficiency (mRNA capture, RT reaction, PCR) are major challenges for single-cell analysis (137). At the time of optimising this protocol (the beginning of this experiment), most transcriptomics studies were carried out on pooled oocytes in which the RNA extraction kits used were designed for microgram quantities that are not sensitive enough for recovering pictogram quantities in single cell

The objective of this chapter was to establish a method for preparing cDNA libraries for high-throughput sequencing that require very minute amounts of cellular material, and in such a way that the library represents the full length of all cDNA molecules—in other words, the full-length sequence of the expressed mRNAs in the cell. This was achieved in two subsequent stages: first, a preparation protocol for isolating and preparing single mouse oocytes and their companion cumulus cells was optimised; second, an amplification strategy was then optimised for generating sufficient cDNA using the Smart-Seq2 approach with some modification. Yet, existing protocols have either 3′ biased or uneven transcript representations. Back dated to the early era of RNA seq when the first sc-RNA-seq was developed by Tang in 2009. He modified a single cell microarrays protocol to be applied in individual mouse oocytes. This method overcame the bias created by microarrays as thousands of genes were expressed (131). Nevertheless, this protocol suffers from pronounced 3' end bias as well as failing to detect transcripts longer than 1 kb, which consequently limits the transcript coverage (118).

Smart-Seq2 strategy, developed by Picelle *et al,* 2014, has successfully expanded the transcript coverage by achieving nearly 40% coverage through the 5' end of transcripts. It relies on a template switching mechanism using the Moloney murine leukaemia virus (MMLV) reverse transcriptase enzyme and the use of a unique TSO.

This strategy possesses two features: template switching and terminal transferee activity. During first-strand synthesis, when reaching the 5' end of the RNA template, the terminal transferase activity of the MMLV reverse transcriptase adds a few additional (non-templated) cytosine residue to the 3' end of the newly synthesised cDNA strand. These bases function as a TS oligo-anchoring site. Upon base pairing between the TSO and the added

deoxycytidine, the reverse transcriptase switches templates, from cellular RNA to the TSO, and is transcribed to the 5' end of the TSO. This mechanism enriched transcripts with an intact 5’ end, resulting in cDNA containing the complete 5' end of the transcript. Along with tagging the cDNA 3' end with oligo (dT) primers, this protocol enables efficient amplification of the entire full-length transcript pool (138).

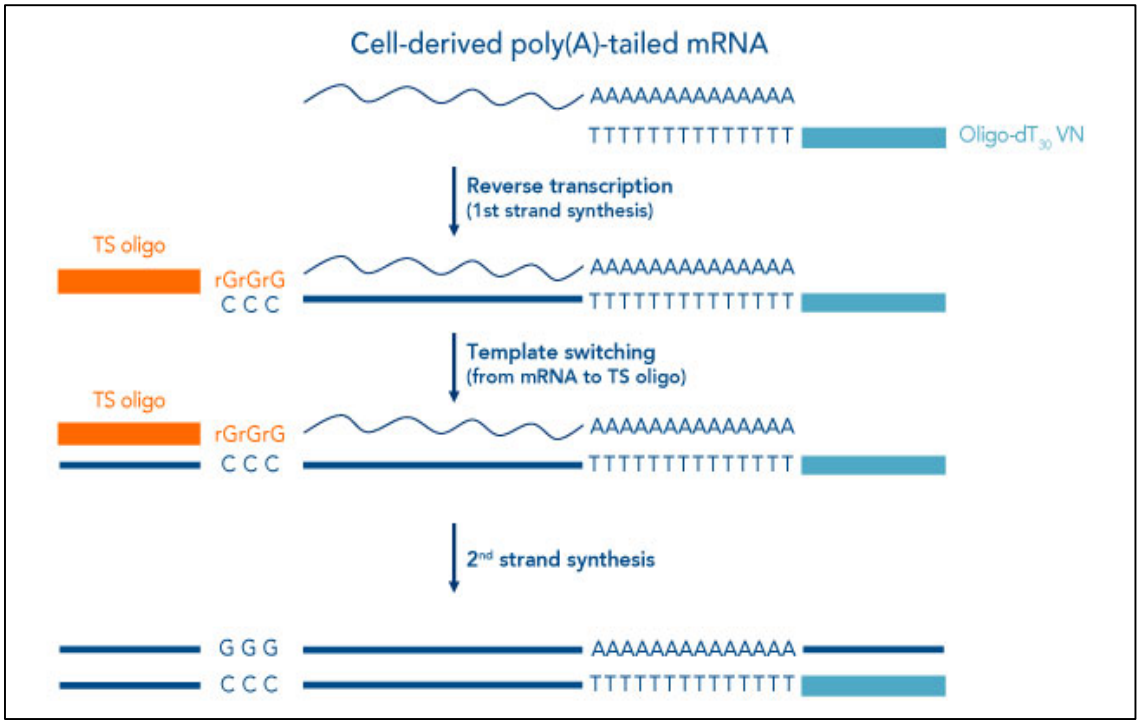


Figure 3.11 Template switching mechanism (TSO) used in Smart-Seq2 protocol, Picture adopted from Picelle *et al*, 2014 (149) .

Blakely *et al.* (2015) conducted an sc-RNA-seq study to investigate the three cell lineages of human blastocysts and used the SMARTer Ultra Low RNA kit for Illumina Sequencing (Clontech Laboratories) for preparing the sequencing libraries. Similarly, Reyes used the

same kit in 2015 to prepare bovine oocyte libraries for sc-RNA sequencing (139). Although both studies successfully generated approximately 30 million reads and achieved their objectives, the SMARTer kit could potentially be incompatible with the single mouse oocytes since a mouse oocyte is about half the size of those in bovines and humans (140). Smart-Seq2 seemed more efficient for recovering the minute amount of RNA in single cells, as it has been utilized successfully in cancer cell analysis showing a significant sensitivity and having half the cost of commercial kits.

Using the Smart-Seq2 strategy for the cells of interest (oocytes and CCs) enabled the recovery of sufficient RNA from a single oocyte. Moreover, the quantity of cDNA achieved was adequate for the standard Nextera Library construction (details in the Chapter 2).

Although it may seem trivial, harvesting cells rapidly and accurately with high efficiency is one of the main challenges in single cell work. Proper sample preparation is therefore crucial when preparing viable sc-RNA-seq libraries. Considering the importance of these technical issues when it comes to sample quality and integrity, micromanipulation techniques were used in isolating single oocyte and CCs from single follicles. Aside from the fact that the use of sterile micro-capillaries (made each day of collection) for cell collection was time consuming, it enabled microscopic supervision of cell isolation, thereby ensuring that the COC was captured during each isolation attempt. Furthermore, it allowed the isolation of each oocyte from its cumulus cells and the processing of each before moving to the next follicle. This method ensured that the oocytes were matched with their cumulus cells and kept in order.

The most common sources of error in downstream steps are the degradation and contamination of input material. Therefore, minimising the time spent in processing the cells during collection was highly prioritised. The use of a single tube approach that kept the RNA material in the same tube until the cDNA synthesised helped to prevent the loss of minute quantities of RNA during pipetting as had occurred in other studies (141). In addition, it saved time spent on sample processing. Also, to avoid technical bias, master mix of reagent was used whenever applicable.

Standard RNA extraction approaches involve using a lysis buffer followed with a stop solution to deactivate its lytic activity before subsequent steps. Employing the SMART-Seq2 protocol allowed for the preparation and use of a hypotonic lysis buffer with small amounts of RNase inhibitor and surfactant. This lysis buffer circumvented the need for post-lysis clean up or the addition of a stop solution prior to reverse transcription, which is often used for single cell RT-PCR. Indeed, keeping the recovered RNA in a minimum volume had significant effect in maintaining RNA.

Cross contamination was also a significant concern throughout this protocol. It is vitally important to keep all workspaces, equipment and reagents RNase free by using products such as RNase ZAP and RNase inhibitors. Avoiding sharing tip and capillaries between the oocytes and the CC samples was taken into account, since clear isolation is necessary to ensure a clean transcriptome profile for each cell type. Additionally, samples were covered to prevent airborne contamination. Collecting and preparing CCs necessitated lengthier procedures than used for the oocytes; CCs tubes were processed in fresh ice, and the lysate tubes were frozen immediately using dry ice.

Preparing the oocytes and CCs required multiple modifications to attain high quality and adequate cDNA yield. A low amount of input material for conducting sc RNA sequencing leads to a high level of technical noise, which complicates data analysis and could mask underlying biological variations (142) The more input starting material, the higher coverage achieved as a result of the increased library complexity. Ideally, larger cells, such as oocytes or embryos, should allow adequate coverage to be obtained for high deep sequencing where a similar level of coverage cannot be achieved in single somatic cell (143).

The main technical challenge encountered when applying this protocol was cDNA pre-amplification (144). As is well known, pre-amplification is a mandatory step when preparing a library for low-input samples, and particularly those similar to this study's single oocytes and cumulus cells. Based on the fact that minor differences in replication efficiency build on each other with every cycle added, it was vital to prepare a library in such away that there was a balance between generating cDNA yield to maximise complexity with the minimum number of enrichment amplification cycles for both the oocytes and cumulus cells. The PCR is reportedly the primary source of systemic bias that amplifies a sequence unequally due to the stability of guanine-cytosine content (G-C) (145). These G-C-rich reigns are actually recalcitrant to this technology, thus, over-amplification of the template will lead to under-presented or reduced quality reads of the transcript rich in G-C region. Thus, a percentage of these sequences will be duplicates, resulting in an uneven distribution of the read coverage across the targeted sequencing regions.

Attempts have been made to avoid this technical issue, with researchers proposing alternative, amplification-free approaches for Illumina sequencing libraries. These new approaches have indeed improved the read distribution. However, these approaches are not

applicable to very minute samples (e.g., single cells) since a large amount of starting material is required (146).

The complexity of a library can reflect the amount of bias generated by experimental design and such bias is a severe issue in data analysis. Libraries and can actually be assessed by the number of genes detected, which demonstrates the depth of coverage achieved (142). However, before this analysis taken, a library can be evaluated by measuring the rate of duplication while sequencing is in process. The higher the duplicate rate, the less complex the library, and the more uneven the coverage of the transcriptome (147). Therefore, when increasing cDNA yield for the purpose of improving library complexity and coverage depth, PCR duplicates should be accounted for, as these are associated with extensive amplification (148).

The present study's amplification protocol involved using the PCR enzyme Kapa HiFi (Kapa Biosystems, UK), which proved to have an advantage in terms of reducing the amplification bias caused by the GC content, as indicated previously (149). This resulted in improved coverage and sensitivity, particularly for GC-rich transcripts. Furthermore, adding betaine and additional magnesium chloride to the mixture improved the performance of reverse transcription. Together, these improvements maximised the Smart-Seq2 cDNA yield as compared to that of the SMARTer technique. This finding supported this study's choice of the Smart-Seq2 method for use in the cells.

3.8. Summary

In summary, the Smart-Seq2 protocol is an efficient approach for preparing single mouse oocytes and their accompanying cumulus cells. This chapter has described how the study

obtained a sufficient amount of high-quality cDNA with the fewest number of manipulation steps, demonstrating why Smart-Seq2 is a highly attractive method for preparing single-cell libraries for Illumina sequencing. In addition to the increased sensitivity of the Smart-Seq2 approach relative to that of other kits, Smart-Seq2 uses standard reagents in minimal volumes, and its cost is half that of its competitors (143). As Smart-Seq2 has shown to be better suited for low volumes of starting material, this protocol was employed to prepare and collect the samples for the whole project. The next step in library construction is the Nextera XT reaction (Illumina, UK), which is detailed in the materials and methods chapter.

Chapter 4

Experimental Design

4.1. Introduction

A fundamental prerequisite for a successful RNA study is that the data have the potential to answer the biological question of interest. An efficient experimental design incorporates biological replicates, as they improve the measurement of variations and detect outliers and increase the precision of the measurement. Selecting the appropriate number of biological replicates is a trade-off between cost and precision. It has been shown that low replicate experiments often have an insufficient statistical power to detect differential gene expression (DGE) correctly, and consequently, obtaining an accurate measurement of the biological variability is not feasible (150). Therefore. This chapter demonstrates the steps for designing single cell-RNA sequencing (sc-RNA-seq) experiment, taken into account these specific criteria set for selecting samples, which is crucial for an accurate data analysis.

4.2. Method

4.2.1. cDNA synthesis

The lysate samples of oocytes and their corresponding CCs were reverse transcribed to cDNA (described in Chapter 2, section 2.5.2). The cDNA was amplified using 18 and 16 PCR cycles for oocyte and CCs amplification, respectively. Although not all samples were used for RNA sequencing, it is preferred to covert all lysate samples to cDNA as they are more stable for long-term use than cell lysates. As many criteria were set to select sample (discussed below in figure 4.1), it was necessary to collect more COCs than required for use. The samples were collected and stored. Two additional CCs suspensions were aliquted in

two thin wall PCR tubes and kept frozen at -80°C, i.e. for each oocyte lysate tube, three aliquots of the matched CCs (0.5 µl :4.5 µl, CCs pellet: Lysis buffer) were stored at -80°C.

Table 4.1: Number of oocyte cumulus complexes (COCs) collected and converted to cDNA.

Age category	Oocyte	CCs	Number of Mice
3-weeks-old	36	36	5 mice
9-weeks-old	36	36	5 mice
1-year-old	46	46	8 mice

4.2.2. Samples number

From the total collected samples shown in Table 4.1, the sample number was reduced to the number indicated below based on selecting an equal number of samples from each mouse.

- 20 oocytes and their 20 matched CCs from 1-year-old mice (mice number =8)
- 16 oocytes and their matched 16 CCs from from 9-week-old mice (mice number =5).
- 16 oocytes and their matched 16 CCs from 3-week-old mice (mice number =5).

4.2.3. Sample selection criteria

A total of 108 samples of oocytes and their matched CCs (Table 4.2, 4.3 and 4.4) were measured using PicoGreen and Qubit.2, respectively. The quality of cDNA for both oocyte

and CCs samples was assessed using TapeStation (methods are described in chapter 2, section 2.7.3 and 2.7.4.). Then, by looking to both criteria: concentration measurement and amplification profile (Figure 4.1), cDNA concentration must be above 0.2 ng/ µl, which is the minimum concentration required for the subsequent reaction: Nextera reaction, as well as each sample should have a high quality amplification profile. In addition, oocytes and CCs must be selected from the same COC.

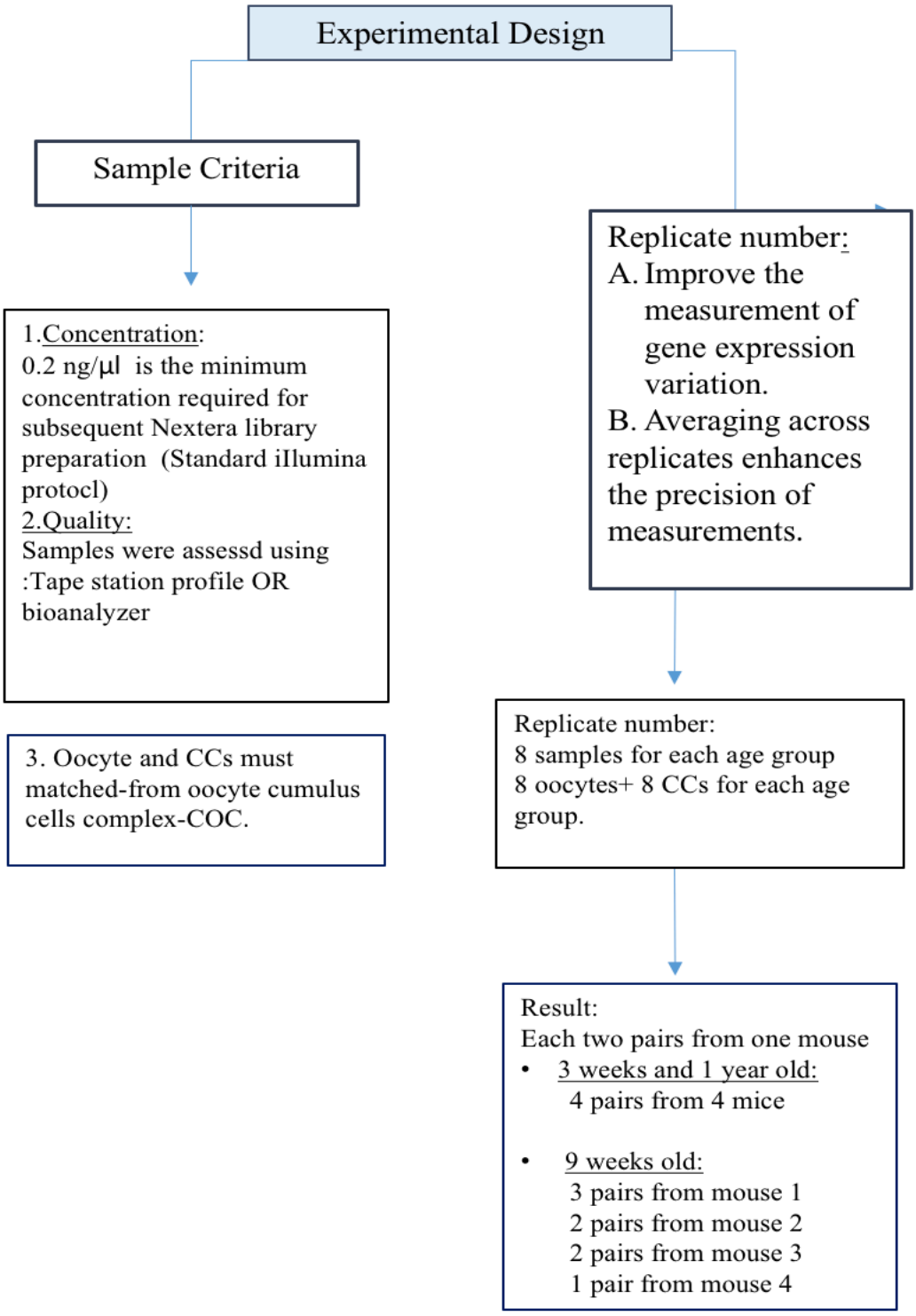


Figure 4.1 Experimental design for RNA-seq.

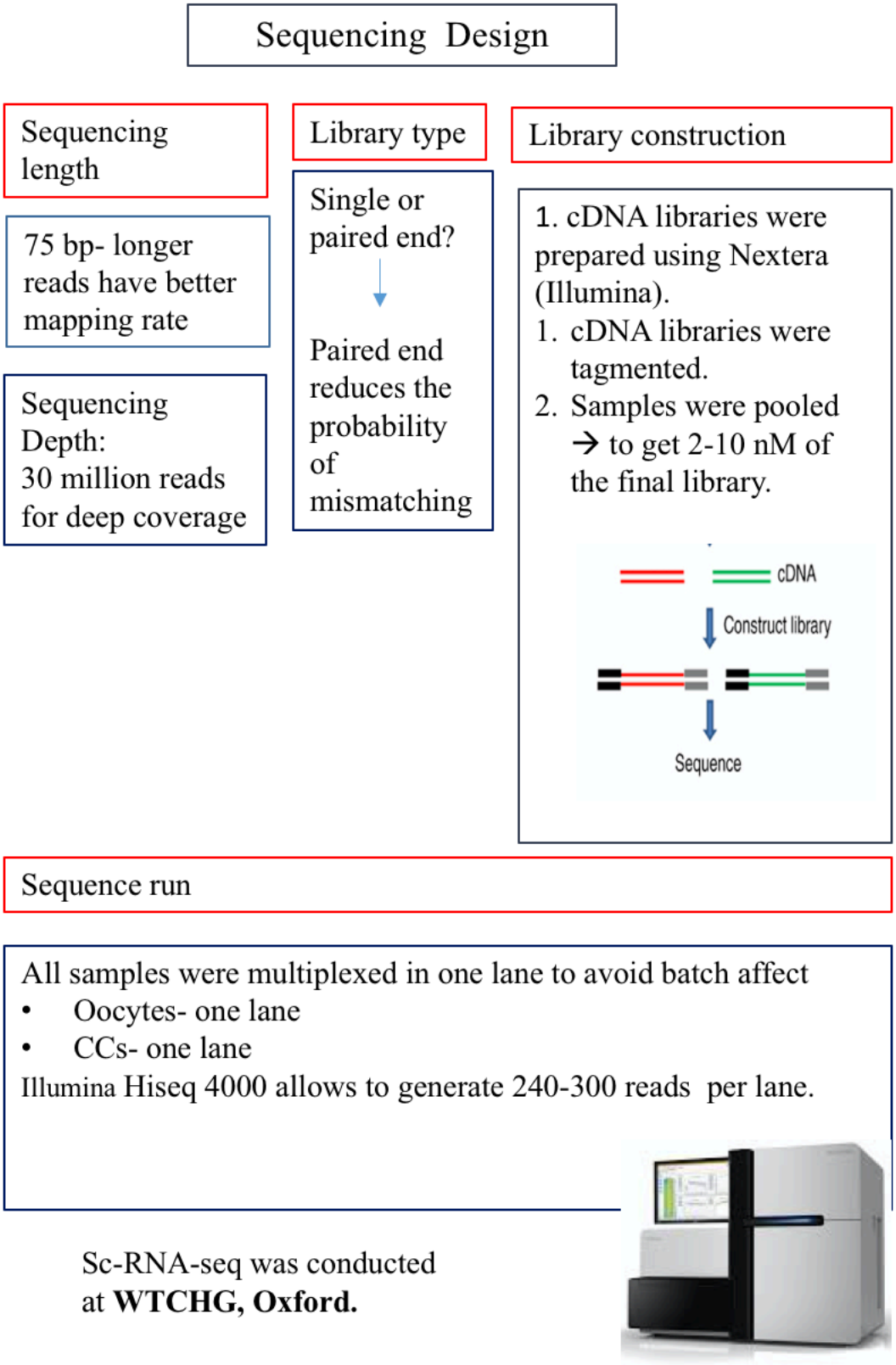


Figure 4.2 Sequencing design for RNA experiment.

4.3. Results: cDNA measurement

The concentrations of cDNA oocyte sample were measured by PicoGreen (Agilent Technologies) wherase CCs samples were measured by Qubit.2 (ThermoFisher). Table 4.2, Table 4.3 and Table 4.4. Samples chosen from the same mouse were highlighted by the same colour.

4.3.1. One-year-old oocytes and their CCs

Table 4.2: cDNA concentration of 1-year-old oocytes and CCs.

Oocyte 1	Oocyte 2	Oocyte 3	Oocyte 4	Oocyte 5	Oocyte 6	Oocyte 7	Oocyte 8	Oocyte 9	Oocyte 10
12.59 ng/ µl	5.4 ng/ µl	9.1ng/ µl	2.7 ng/ µl	0.23 ng/ µl	4.4 ng/ µl	9.9 ng/ µl	0.03 ng/ µl	10.7 ng/ µl	12.5 ng/ µl
Oocyte 11	Oocyte 12	Oocyte 13	Oocyte 14	Oocyte 15	Oocyte 16	Oocyte 17	Oocyte 18	Oocyte 19	Oocyte 20
20.8 ng/ µl	5.6 ng/ µl	6.69 ng/ µl	8.32 ng/ µl	6.6 ng/ µl	4 ng/ µl	19.4 ng/ µl	21.4 ng/ µl	5.9 ng/ µl	0.2 ng/ µl

CCs1	CCs2	CCs3	CCs4	CCs5	CCs6	CCs7	CCs8	CCs9	CCs 10
0.27 ng/ µl	0.11 ng/ µl	0.26 ng/ µl	0.18 ng/ µl	22.3 ng/ µl	0.23 ng/ µl	1.9 ng/ µl	4.3 ng/ µl	0.44 ng/ µl	2.9 ng/ µl
CCs11	CCs12	CCs13	CCs14	CCs15	CCs16	CCs17	CCs18	CCs19	CCs 20
0.15 ng/ µl	1.1 ng/ µl	0.13 ng/ µl	0.15 ng/ µl	0.57 ng/ µl	1.3 ng/ µl	1.1 ng/ µl	0.46 ng/ µl	2.3 ng/ µl	6.3 ng/ µl

4.3.2. Nine-week-old (young adult) samples

The cDNA of oocytes and CCs measurements are shown in **Tale 4.3**

Table 4.3: cDNA concentrations of 9-week oocytes and CCs.

Oocyte1	Oocyte2	Oocyte3	Oocyte4	Oocyte5	Oocyte6	Oocyte7	Oocyte8
17.9 ng/ µl	19.1 ng/ µl	17.9 ng/ µl	15.6ng/ µl	22.3 ng/ µl	13.7 ng/ µl	20.37 ng/ µl	0.29 ng/ µl
Oocyte 9	Oocyte 10	Oocyte 11	Oocyte 12	Oocyte 13	Oocyte 14	Oocyte 15	Oocyte 16
10.8 ng/ µl	0.923 ng/ µl	9.7 ng/ µl	11.3 ng/ µl	9.7 ng/ µl	5.9 ng/ µl	4 ng/ µl	-0.33 ng/ µl

CCs1	CCs2	CCs3	CCs4	CCs5	CCs6	CCs7	CCs 8
0.68 ng/ µl	0.0.6ng/ µl	1.3 ng/ µl	3.3 ng/ µl	0.4 ng/ µl	0	0.19 ng/ µl	0.18 ng/ µl
CCs9	CCs 10	CCs 11	CCs 12	CCs 13	CCs 14	CCs 15	CCs 16
0.5 ng/ µl	0.15 ng/ µl	0.06 ng/ µl	8.19 ng/ µl	0	0.32 ng/ µl	0.05 ng/ µl	0.4 ng/ µl

4.3.3. Three-week-old (Pre-pubertal)

The cDNA measurements for three-weeks-old oocyte are shown in Table 4.4

Table 4.4: cDNA concentrations of 3-week oocytes and CCs.

Oocyte1	Oocyte2	Oocyte3	Oocyte4	Oocyte5	Oocyte6	Oocyte7	Oocyte8
2.1 ng/ µl	7.7 ng/ µl	9.2 ng/ µl	13 ng/ µl	11.4 ng/ µl	18.5 ng/ µl	16.8 ng/ µl	9.6 ng/ µl
Oocyte 9	Oocyte 10	Oocyte 11	Oocyte 12	Oocyte 13	Oocyte 14	Oocyte 15	Oocyte 16
11.5 ng/ µl	6.1 ng/ µl	9.8 ng/ µl	14.5 ng/ µl	5.1 ng/ µl	1.2 ng/ µl	14.1ng/ µl	0.1 ng/ µl

CCs1	CCs2	CCs3	CCs4	CCs5	CCs6	CCs7	CCs8
0.68 ng/ µl	0.06 ng/ µl	0	0.34 ng/ µl	1.55 ng/ µl	0.73 ng/ µl	0.71 ng/ µl	6.5 ng/ µl
CCs9	CCs 10	CCs 11	CCs 12	CCs 13	CCs 14	CCs 15	CCs 16
0.25 ng/ µl	0.25 ng/ µl	0.7 ng/ µl	0.8 ng/ µl	0.89 ng/ µl	0.53 ng/ µl	0.44 ng/ µl	0.26 ng/ µl

4.4. Sample selected for the subsequent step: Nextera XT library

By applying the criteria indicated in Figure 4.1, eight oocyte samples were selected from each age group in addition to their matched CCs where two samples are from same mother. Each cDNA sample was more than 0.2 ng/µl, and each two pairs were taken from one mouse. The final selection of samples for the Nextera XT library is shown in Table 4.5.

4.4.1. cDNA quality assessment

The quality of cDNA amplification was assessed using TapeStation (Agilent, UK). The selection of sample was firstly based on cDNA concentration then, cDNA quality. Therefore, 48 amplification profiles are listed as follow:

One-year old oocyte (Figure 4.3 [A], [B])

9-weeks-old oocytes (Figure 4.4 [A], [B])

3-weeks-old oocytes (Figure 4.5 [A], [B])

One-year old CCs (Figure 4.6 [A], [B])

9-weeks-old CCs (Figure 4.7 [A], [B])

3-weeks-old CCs (Figure 4.8 [A], [B])

Table 4.5: Final selection of cDNA samples of oocytes (Yellow) matched with their companion CCs. * indicates that for the 9-weeks samples, only one pair (oocyte and its CCs) from mouse 2 was chosen according to the selection criteria.

One-year-old							
1	2	3	4	5	6	7	8
Mouse1		Mouse2		Mouse3		Mouse4	
12.6 ng/ µl	9.2 ng/ µl	10.8 ng/ µl	12.5 ng/ µl	6.7 ng/ µl	4.7 ng/ µl	19.5 ng/ µl	21.5 ng/ µl
0.23 ng/ µl	0.26 ng/ µl	0.44 ng/ µl	0.29 ng/ µl	0.57 ng/ µl	1.3 ng/ µl	1.2 ng/ µl	0.46 ng/µl
9-week-old							
1	2	3	4	5	6	7	8
Mouse1			Mouse2*	Mouse3		Mouse4	
18 ng/ µl	18 ng/ µl	15.6 ng/ µl	22.3 ng/ µl	20.4 ng/ µl	10.9 ng/ µl	11.3 ng/ µl	5.9 ng/ µl
0.7ng/ µl	1.3 ng/ µl	3.4 ng/ µl	0.4 ng/ µl	0.19 ng/ µl	0.5 ng/ µl	8.2 ng/ µl	0.3 ng/ µl
3-weeks-old							
1	2	3	4	5	6	7	8
Mouse1		Mouse2		Mouse3		Mouse4	
2.1 ng/ µl	13.04 ng/ µl	11.4 ng/ µl	18.5 ng/ µl	16.7 ng/ µl	9.7 ng/ µl	9.8 ng/ µl	5.1 ng/ µl
0.68 ng/ µl	0.34 ng/ µl	1.6 ng/ µl	0.73 ng/ µl	0.71 ng/ µl	6.5 ng/ µl	0.7 ng/ µl	0.9 ng/ µl

4.5. Final quality assessment

4.5.1. TapeStation profiles for 1-year-old oocytes

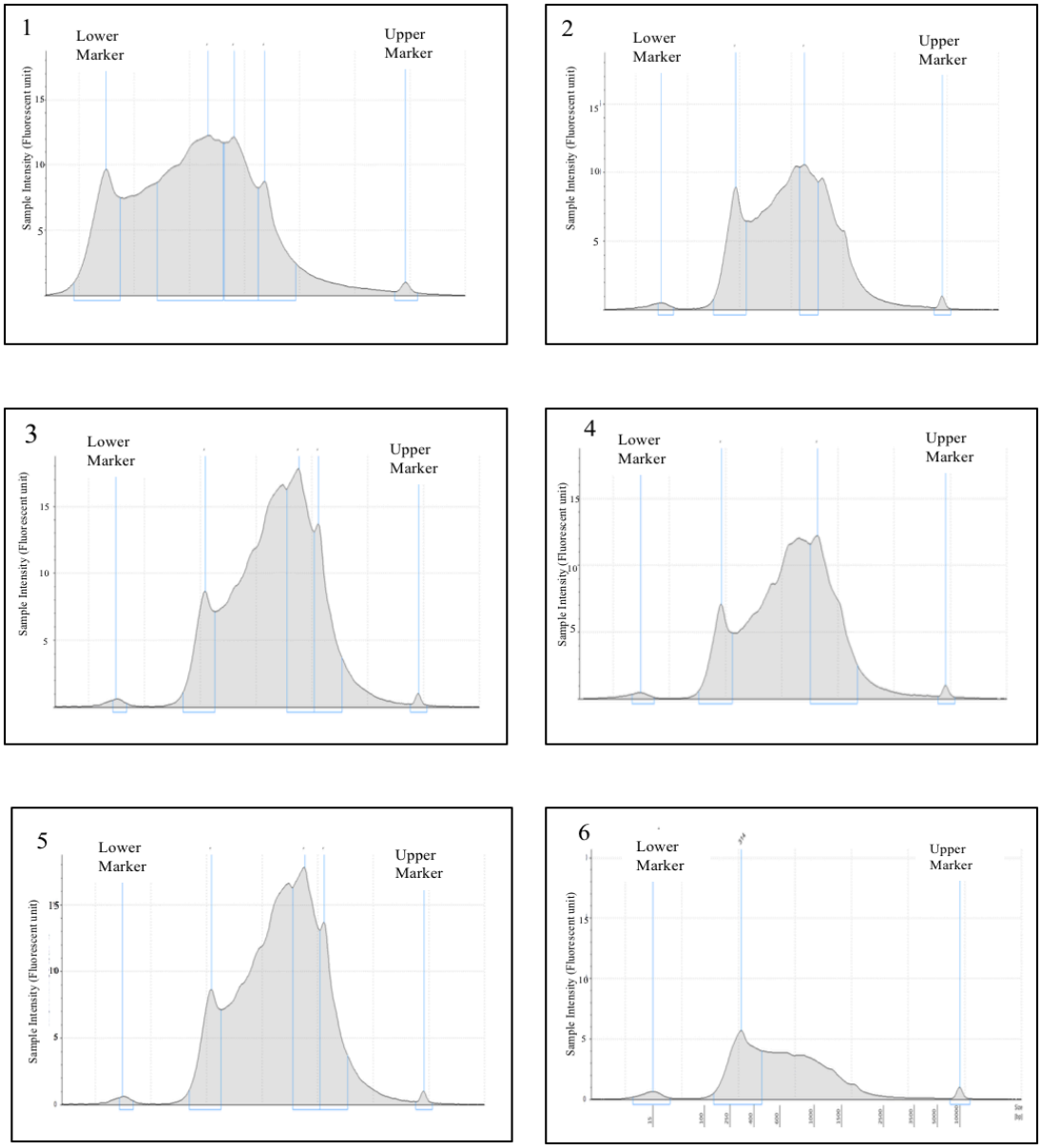


Figure 4.3 (A) TapeStation Profiles for 1-year-old. A quality assessment of amplification profiles for: 1) Oocyte1, 2) oocyte2, 3) Oocyte3, 4) Oocyte4, 5) Oocyte5 and 6) Oocyte 6.

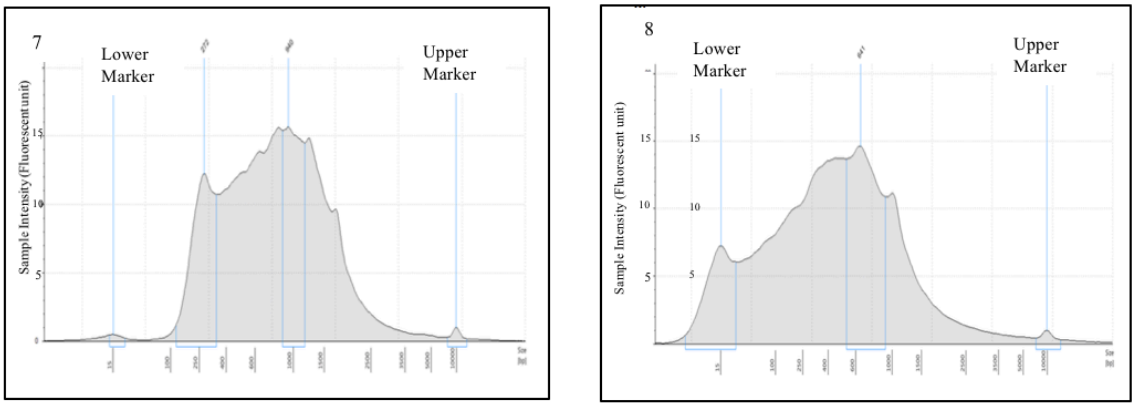


Figure 4.3 (B). TapeStation Profiles for 1-year-old oocytes: 7) Oocyte7 and 8) Oocyte8.

4.5.2. TapeStation profiles for 9-weeks-old oocytes

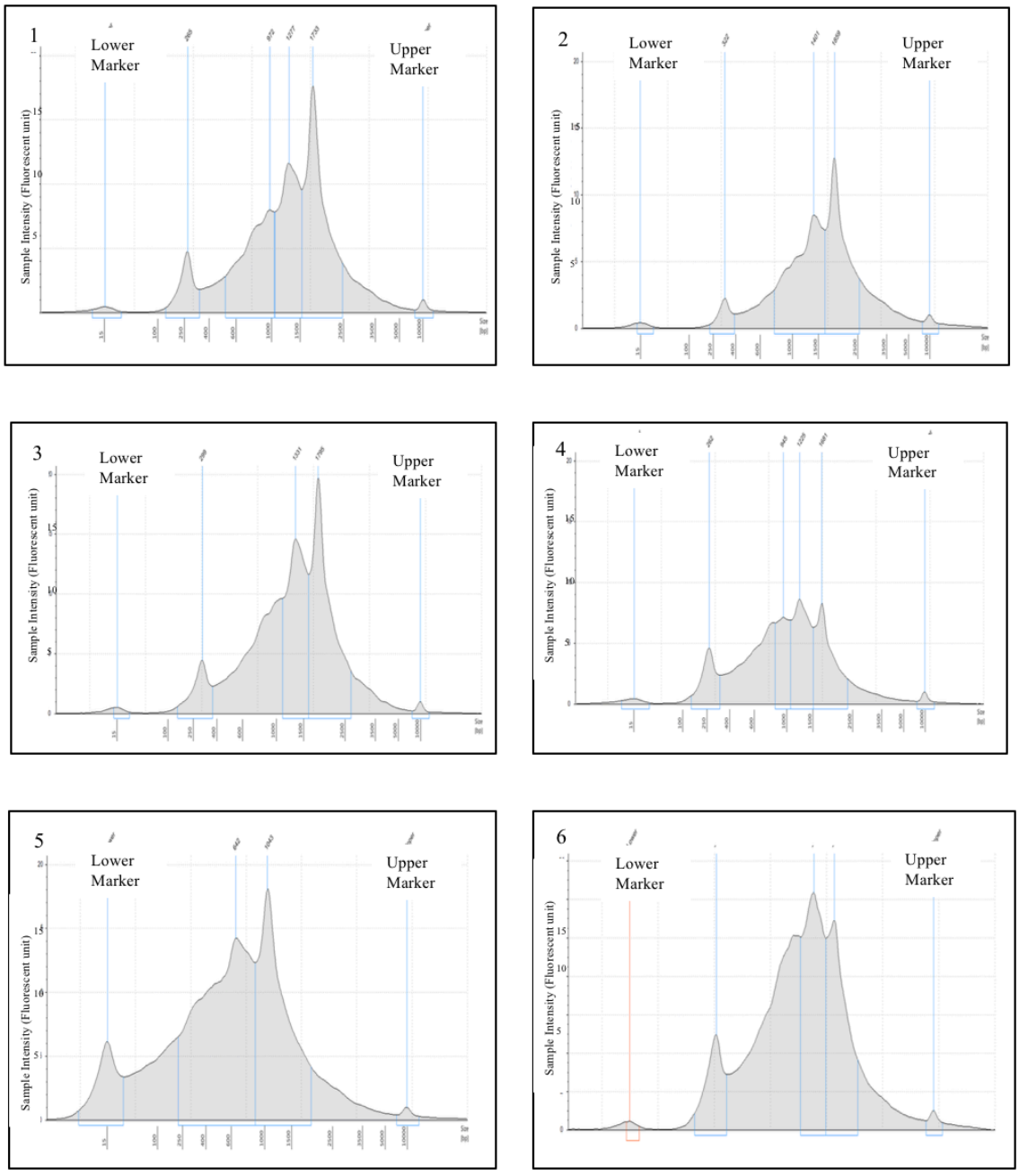


Figure 4.4 (A) TapeStation Profiles for 9-weeks-old. A quality assessment of amplification profiles for: 1) Oocyte1, 2) oocyte2, 3) Oocyte3, 4) Oocyte4, 5) Oocyte5 and 6) Oocyte 6.

§

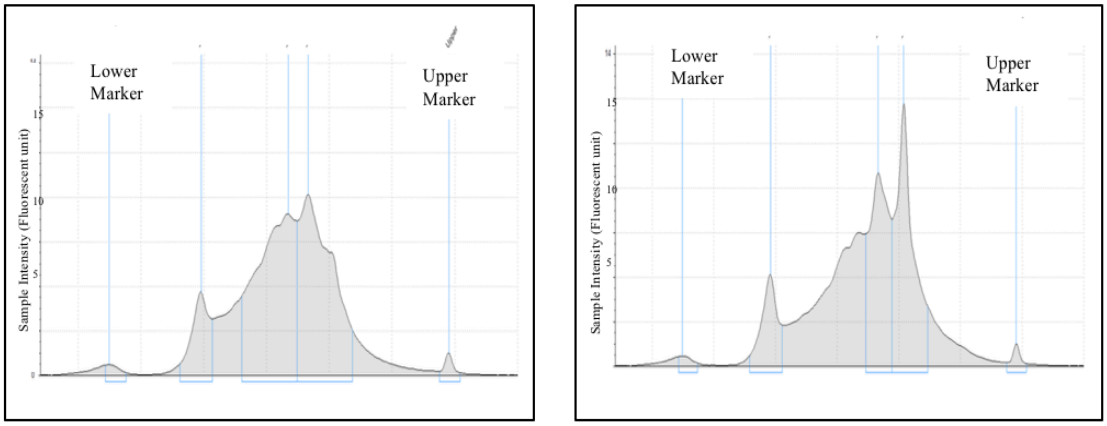


Figure 4.4 (B) TapeStation Profiles for 9-weeks-old. A quality assessment of amplification profiles for: 5) Oocyte5 and 6) Oocyte6.

4.5.3. TapeStation profiles for 3-weeks-old oocytes

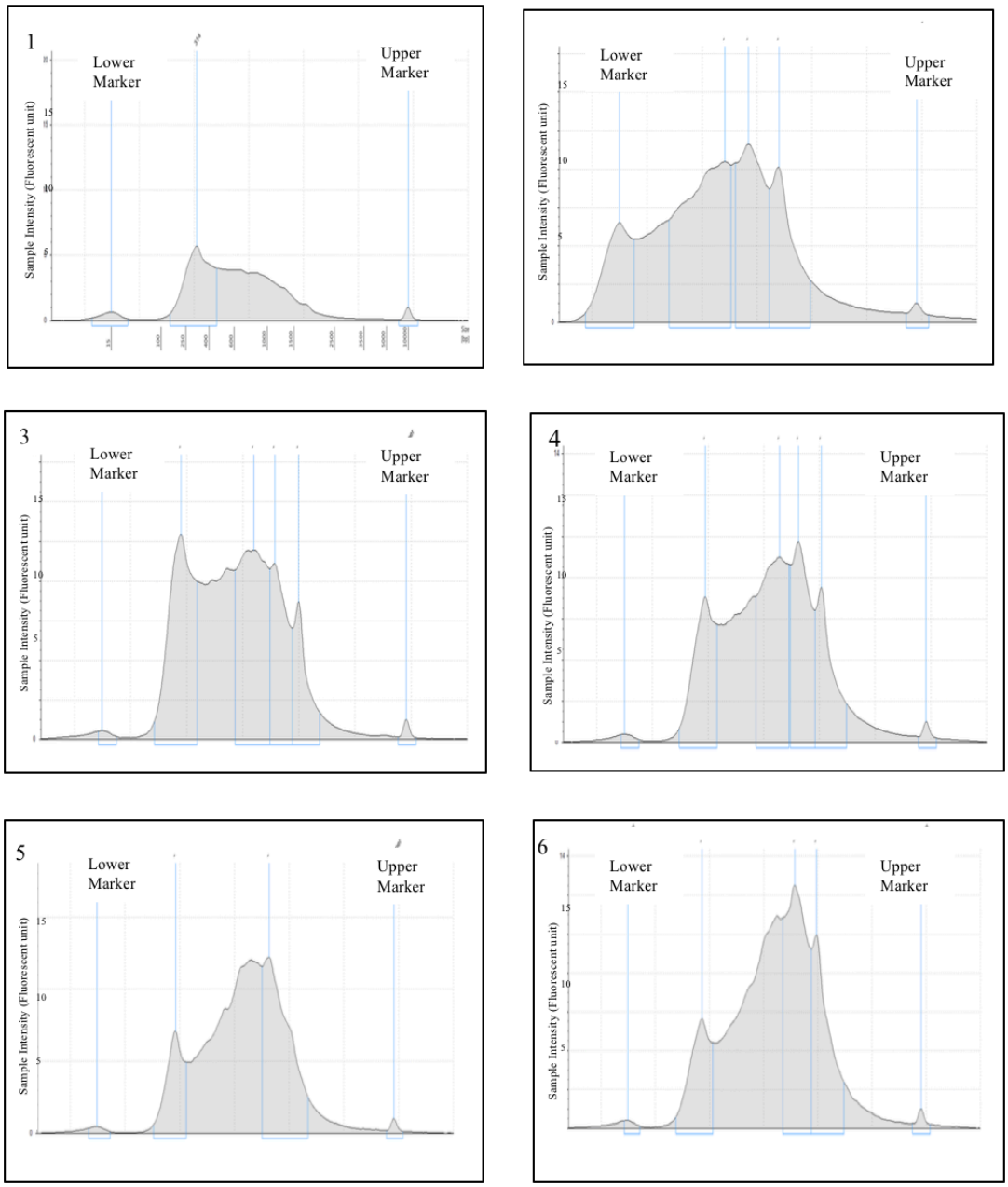


Figure 4.5 (A) TapeStation Profiles for 3-weeks-old. A Qulaity assessment of amplification profiles for: 1) Oocyte1, 2) oocyte2, 3) Oocyte4, 4) Oocyte5 and 6) Oocyte 6.

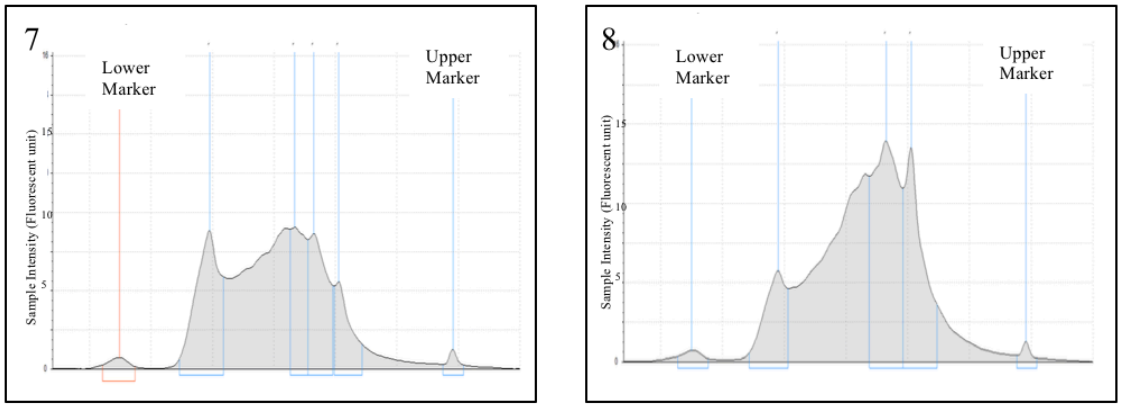


Figure 4.5 (B) TapeStation Profiles for 3-weeks-old. A quality assessment of amplification profiles for: 7) Oocyte7 and 8) Oocyte8.

4.5.4. TapeStation profiles for 1-year--old CCs

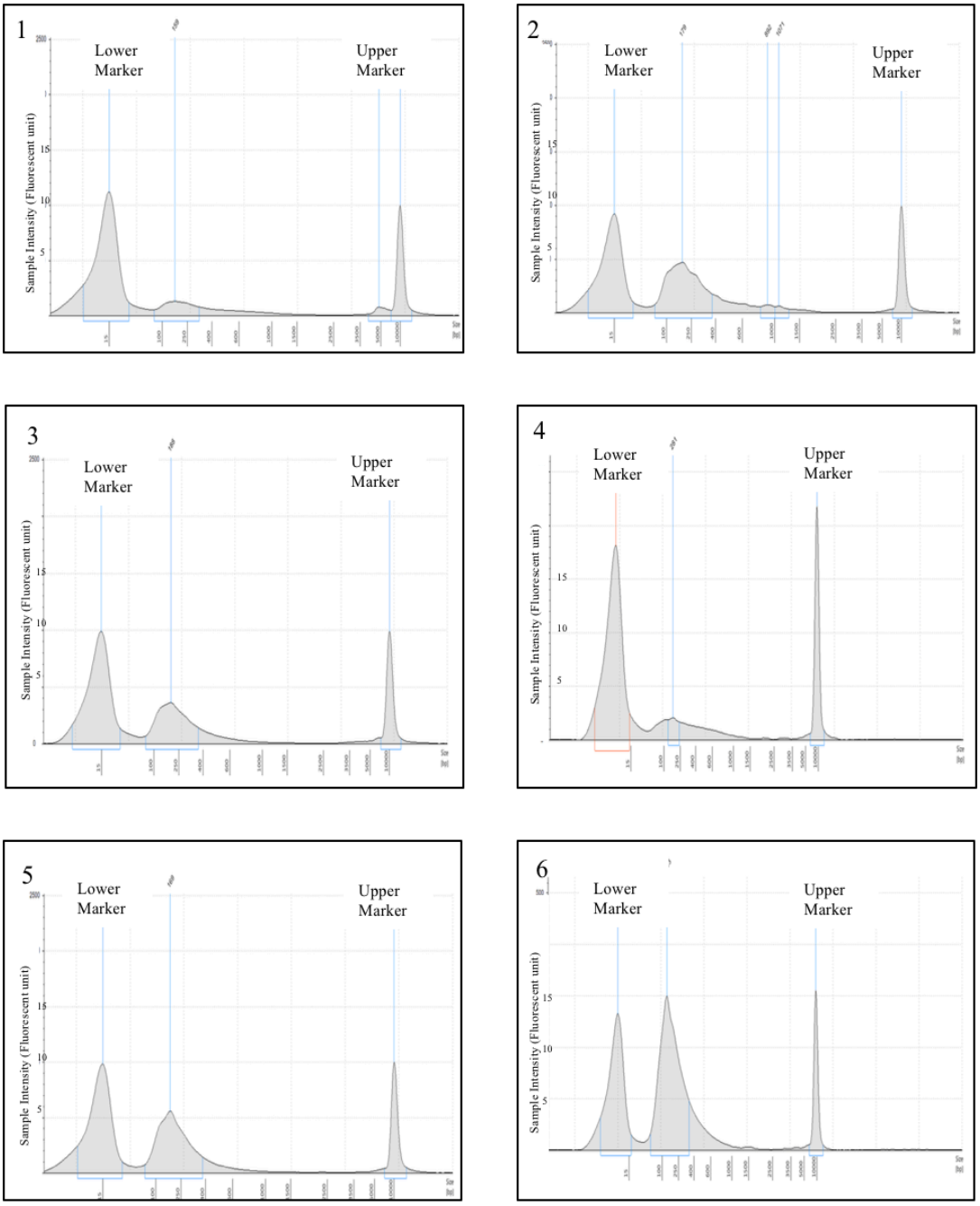


Figure 4.6 (A) TapeStation Profiles for 1-year-old Cumulus cells. A quality assessment of amplification profiles for: 1) CCs1, 2) CCs2, 3) CCs3, 4) CCs4 and 5) CCs5 and 6) CCs 6.

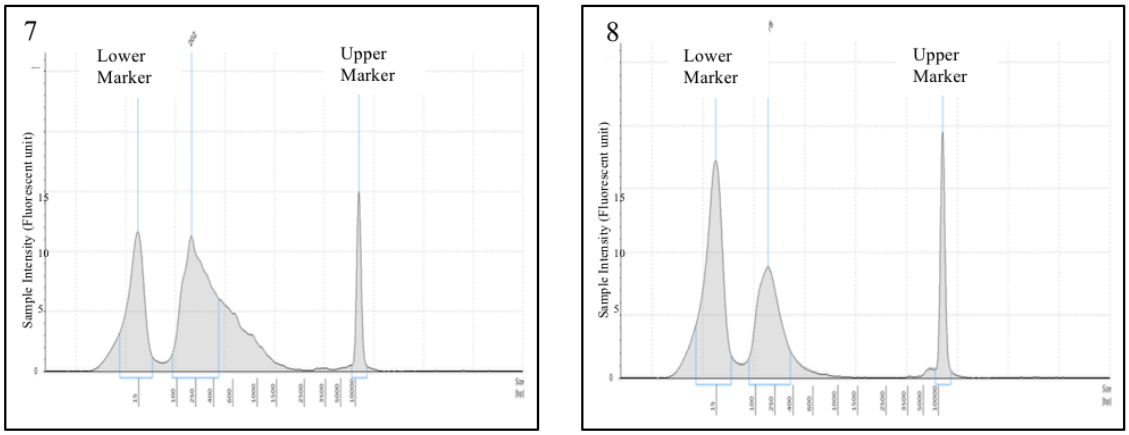


Figure 4.6 (B) TapeStation Profiles for 1-year-old Cumulus cells. A quality assessment of amplification profiles for: 7) CCs7 and 8) CCs8

4.5.5. TapeStation profiles for 9-weeks-old CCs

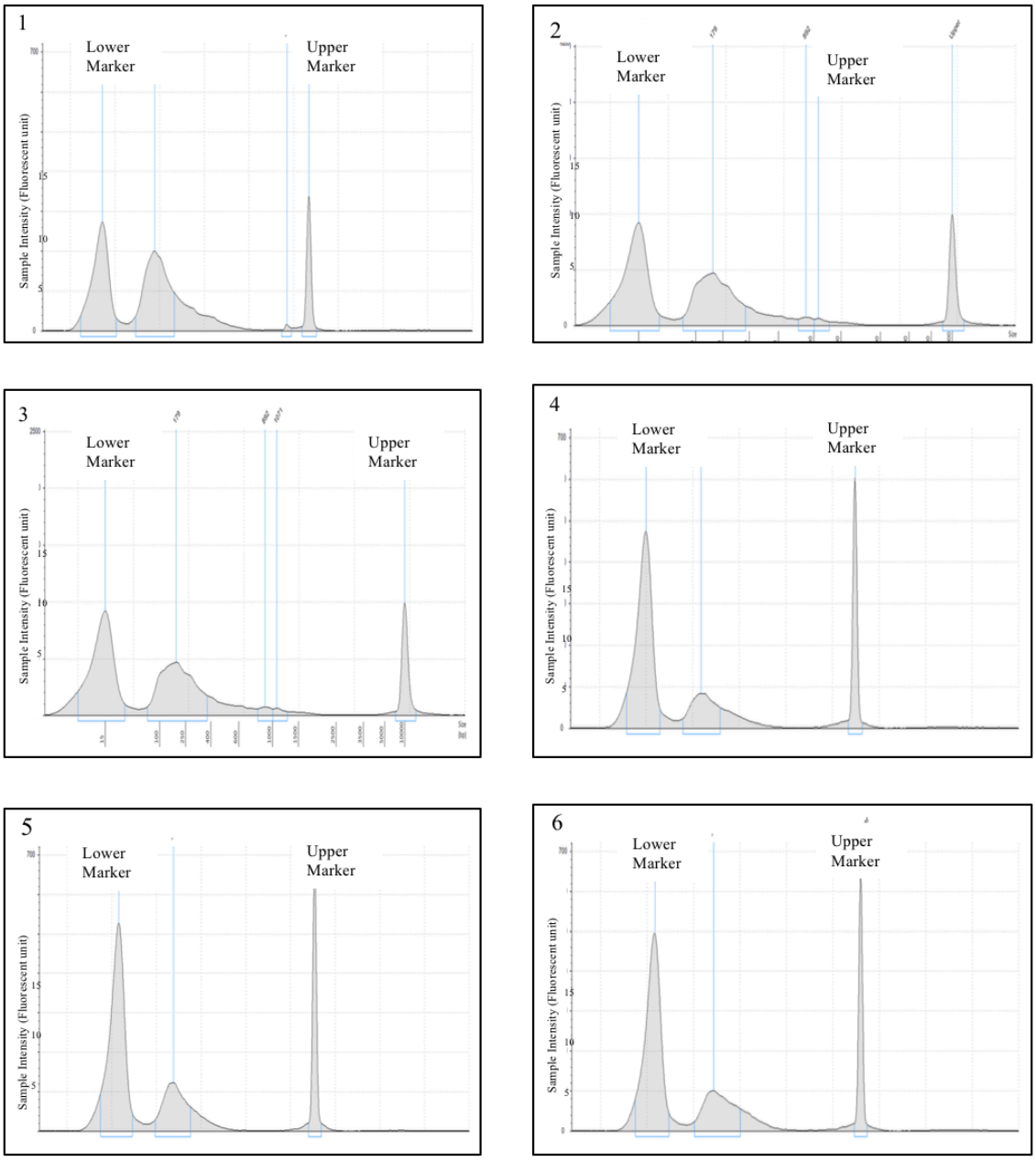


Figure 4.7 (A) TapeStation Profiles for 9-weeks-old CCs. A quality assessment of amplification profiles for: 1) CCs1, 2) CCs2, 3) CCs3, 4) CCs4, 5) CCs5 and 6) CCs 6.

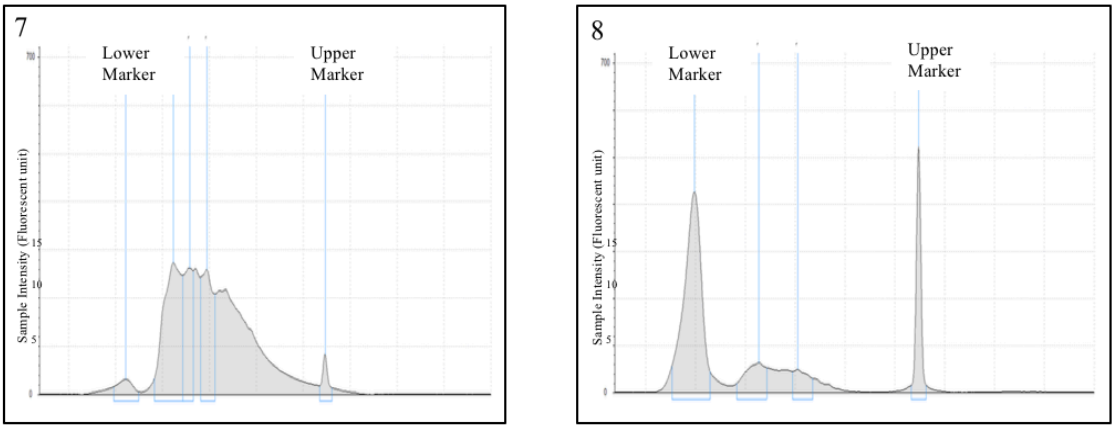


Figure 4.7 (B) TapeStation Profiles for 9-weeks-old CCs. A quality assessment of amplification profiles for: 7) CCs7 and 8) CCs8.

.

4.5.6. TapeStation profiles for 3-weeks-old CCs

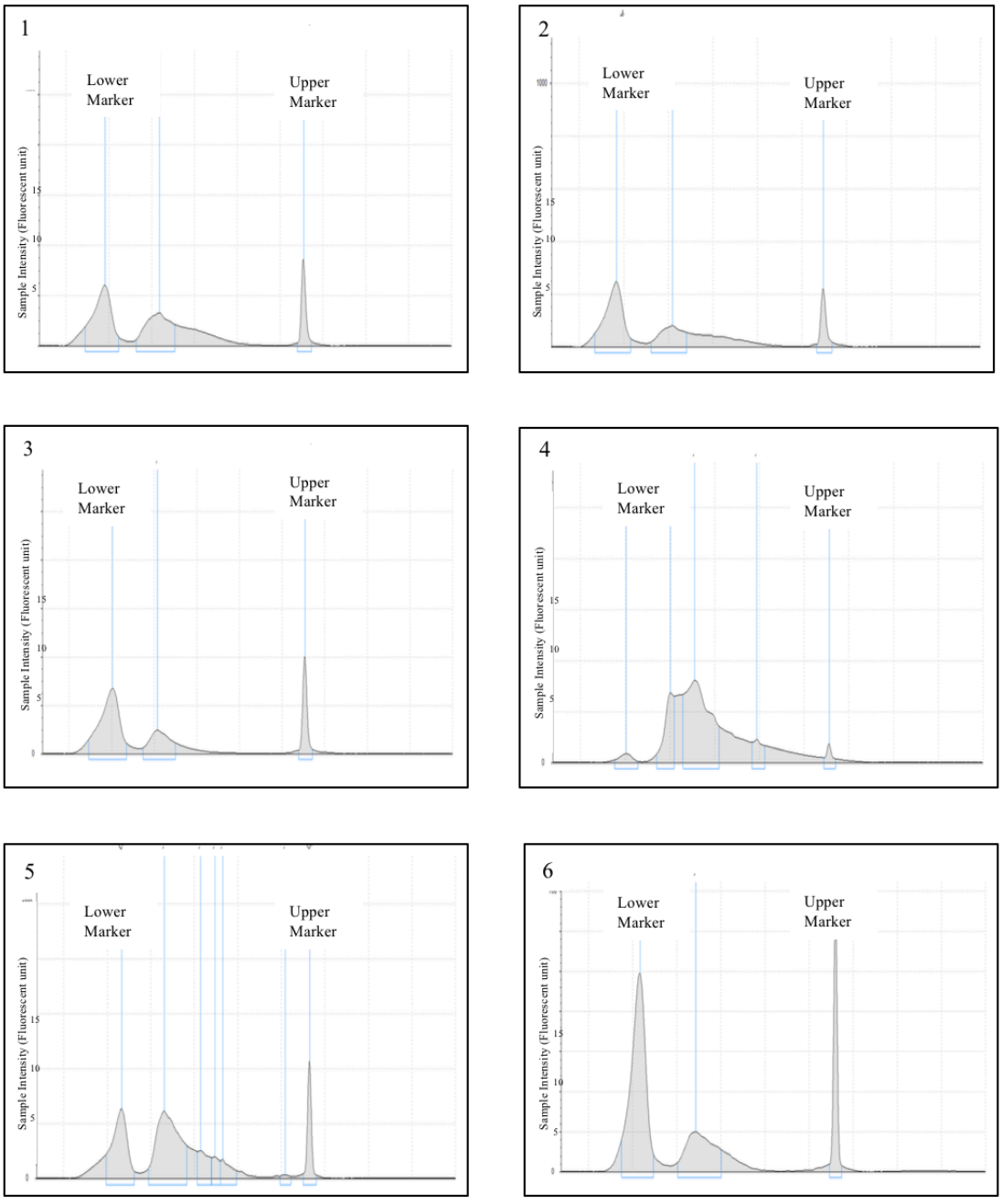


Figure 4.8 (A) TapeStation Profiles for 3-weeks-old CCs. A quality assessment of amplification profiles for: 1) CCs1, 2) CCs2, 3) CCs3, 4) CCs4, 5) CCs5 and 6) CCs 6.

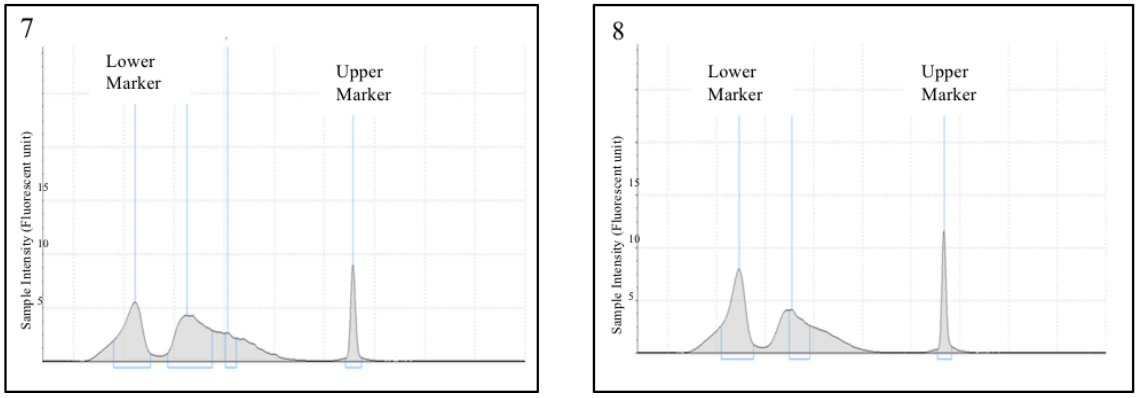


Figure 4.8 (B) TapeStation Profiles for 9-weeks-old CCs. A quality assessment of amplification profiles for: 7) CCs7 and 8) CCs 8.

4.6. Discussion

Although three replicates are the minimum requirement for estimating an accurate *P*-value and false discovery rate, it has been demonstrated that increasing the number of biological replicates from 2 to 5 improved the ability to detect differential expression (151). The power statics and the total number of differential expressed genes (DEG) increase when additional replicates or sequences are added (152); however, whether sacrificing sequencing depth for additional replications is beneficial was a concern while planning this project because it is not possible to estimate the amount of sequencing depth required to obtain DE genes.

As the aim was to apply a cost-effective strategy that guaranteed achieving the maximum power for detecting DEG, additional biological replicates with a decreased sequencing depth became an attractive strategy. A previous study compared the number of DE genes produced when two biological replicates were increased to three biological replicates at a 10 M depth, which yielded a 34,7% increase in the number of DE genes; however, an increase from six replicates to seven replicates yielded 26,3% more DE genes at this sequencing depth (152). Therefore, eight oocytes and their corresponding eight CCs were selected from each age group. The eight biological replicates were carefully selected for each age group considering that each two pairs of each COC are collected from one mouse. It could be argued that replicates number should be the number of mice used in this study, which is four female mice in each age group. However, in fact, it is well known that each oocyte has a different fate even those retrieved from the same mother. In our biroinformatic analysis; despite of being used 8 as the n number, it was still useful to have multiple animals as this can establish the robustness of the changes, allowing us to assess how the differences vary across animals, and makes generalisation to other animals possible.

It seems that additional replicates increase the number of DE genes, although a deeper sequencing may not affect the accuracy estimation of highly expressed genes. Low gene expression has been shown to have an approved estimation accuracy when there is deeper sequencing. Therefore, eight biological replicates were planned to be run for each age group and for both oocytes and CCs.

Sequencing depth, or coverage, is estimated as the number of total mapped sequences. Deeper sequencing for sc-RNA-seq improves the quantification of transcripts. Because genes are expressed at different levels in the transcriptome, a large number of sequences is required to measure the transcripts that are expressed at low levels (153). Therefore, increasing the sequencing depth is essential for an accurate measure of low gene expression. To achieve this, the aim was for the sequencing strategy to generate 30 million reads per single oocyte and per CCs sample.

Illumina sequencing was set to sequence at 75-bp paired read ends per sample. The longer the read, the more precise the mapping rate. Paired ends were chosen over single ends, as the sequencing of both ends of RNA fragments yields more information, especially for the detection of alternative splicing and chimeric transcripts. Paired-end sequencing is more efficient strategy for characterising and quantifying transcriptomes, and it increases the sensitivity and specificity of the detection of alternative splicing.

4.7. Summary

Overall, the success of sc-RNS-seq is relied on experimental design. Several factors need to be considered when planning a sc-RNA-seq experiment. Sequencing depth, library type, and number of replicates are important factors to determine.

Chapter 5

Transcriptomic Analysis

of Pre-Pubertal Oocyte

5.1.Introduction

At birth, oocytes are arrested at prophase of the first meiotic division. Oocytes are either naked or surrounded with a thin layer of few granulosa cells (154). Follicles then begin to grow during the first week of mouse life. It takes about 10 to 12 days until medium follicles are formed and the number of growing follicles increases until large follicles are found in the ovary of a juvenile mouse (during the third week). Between birth and early onset of maturity around week 5, the total number of oocytes decreases. At this point, follicle differentiation at different stages of development has already occurred. Follicle development does not proceed beyond this antral stage until the mouse reaches sexual maturation (3).

Oocytes remain in the diplotene stage during oocyte growth until shortly before ovulation (8), when they respond to the pre-ovulatory surge of Luteinising hormone (LH). Oocytes undergo important changes that determine their quality. Oocyte quality is mainly dedicated by both nuclear and cytoplasmic maturation which acquired during oocyte growth (155). While nuclear maturation includes Germinal vesicle breakdown (GVBD) and progression to meiosis 1 represented with the extrusion of the first polar body, cytoplasmic maturation refers to complicated events that result in acquisition of developmental competence. An oocyte's capability to attain successful fertilisation and undergo normal embryonic development is referred to developmental competence (156).

It is well documented that oocytes have limited developmental competence at a pre-pubertal age. A pre-pubertal oocyte is less capable than an adult oocyte of sustaining fertilisation and undergoing normal embryonic development (157). *In vitro* maturation studies of pre-pubertal oocytes have resulted in lower cleavage and lower blastocyst formation rates

(30,31). Many studies have addressed factors that are thought to be correlated with the competence of a pre-pubertal oocyte (140,158). Oocytes from large follicles have been found to have higher developmental potential than those from small follicles. Romaguera *et al* showed that blastocyst yield can increase from 5.45% to 18.07% when oocytes are recovered from large follicles. It is believed that oocyte competence is correlated with increasing oocyte diameter that expected to be fully grown oocyte (159).

Moreover, Previous study using nuclear transfer have confirmed that cytoplasmic maturation of pre-pubertal oocytes is comprised, which has led to conclude that nuclear and cytoplasmic maturation occur in a synchronised manner necessary for successful fertilization (160). This reduced competence is mainly attributed to inadequate cytoplasmic maturation (161). Oocytes build up the store of the essential factors including mRNA, organelles, proteins, which support meiotic maturation. They rely on stored transcripts to both complete meiosis and continue initial cleavage divisions after fertilisation, until the maternal zygotic transition (162). Thus, oocytes develop competence progressively as they approach puberty (163). However, it remains unclear at which period specifically development is completely acquired.

Despite the clinical demand of the using the pre-pubertal oocytes in infertility treatment, there has been limited progress in improving the maturation of pre-pubertal oocyte. The need for transcriptomic analyses in this area is essential to improve our understanding of the mechanisms underlying the developmental competence of immature oocyte. Single cell transcriptomics is a powerful method that allows to closely examine the transcriptional

activity of an oocyte at pre-pubertal age, and this investigation has the potential to develop ways in which oocyte quality can be manipulated and improved.

The aim of this chapter is to investigate the transcriptome differences between single oocytes from pre-pubertal female mice at 3 weeks of age and sexually mature mice age 9 weeks and gain insights into the regulatory pathways that contribute to the limited developmental competence of a pre-pubertal oocyte.

5.2. Method

5.2.1. Sample preparation

Individual COCs were recovered from 3 weeks old (pre-pubertal oocyte) and 9 weeks old (adult oocyte). Mice were stimulated with PMSG 48 hours prior to the collection of ovaries. Oocytes were from surrounding CCs, washed in PBS/BSA and kept in lysate buffer at -80°C. Ployadenylated RNA was isolated, amplified and processed for Illumina sequencing using Smart-seq2 approach (preparation details in section 2.4 and 2.5)

5.2.2. Bioinformatics analysis

The RNA sequencing data were processed using FastQC format. Duplicates and adaptors were removed. Then, only uniquely mapped reads were retained using STAR software. Differential gene expression was carried out using edgeR, the data were then enriched for pathways analysis using IPA, web based software, and Metacore Thompson (detailed in section 2.9).

5.2.3. Functional analysis

Pathway analysis and network identification were carried out using two software: Ingenuity Pathways Analysis (IPA) and Metacore. A total of 2912 differential expressed (DE) transcripts with their corresponding *P* values, and fold changes were imported into Ingenuity Pathways Analysis (IPA) (Ingenuity® Systems, http://www.ingenuity.com) to identify canonical pathways. The second software was Metacore Software used to analyse the molecular function and to determine the biological relevance.

5.3. Results

The result of 3 weeks oocyte transcriptomics is divided into four sections as follow:

1. Sequencing quality
2. Principle component analysis (PCA)
3. Differential genes expression
4. Functional analysis

   Note: This chapter discusses the comparison of 3-week-old oocytes relative to 9-week-old oocytes. However, since the sequencing data of the three ages (including 1-year-old transcriptome results in Chapter 6) were processed and generated together for technical purposes, the sequencing quality and PCA results show the three ages together

5.3.1. Sequencing quality

5.3.1.1. Mapping rate

Illumina sequencing resulted in 30 million reads per oocyte (*n=24*; 8 oocytes from each age: 3 weeks, 9 weeks and 1 year). As indicated in Figure 5.1, a total of three columns represent the mapping rate for each oocyte. From about ~ 30 million reads that were generated from each single oocyte, an average of 28 million reads were mapped, and about 23 million reads were mapped uniquely. About 70% of reads were uniquely mapped to mouse genome. The high mapping rate indicates to the high quality of sequencing data.

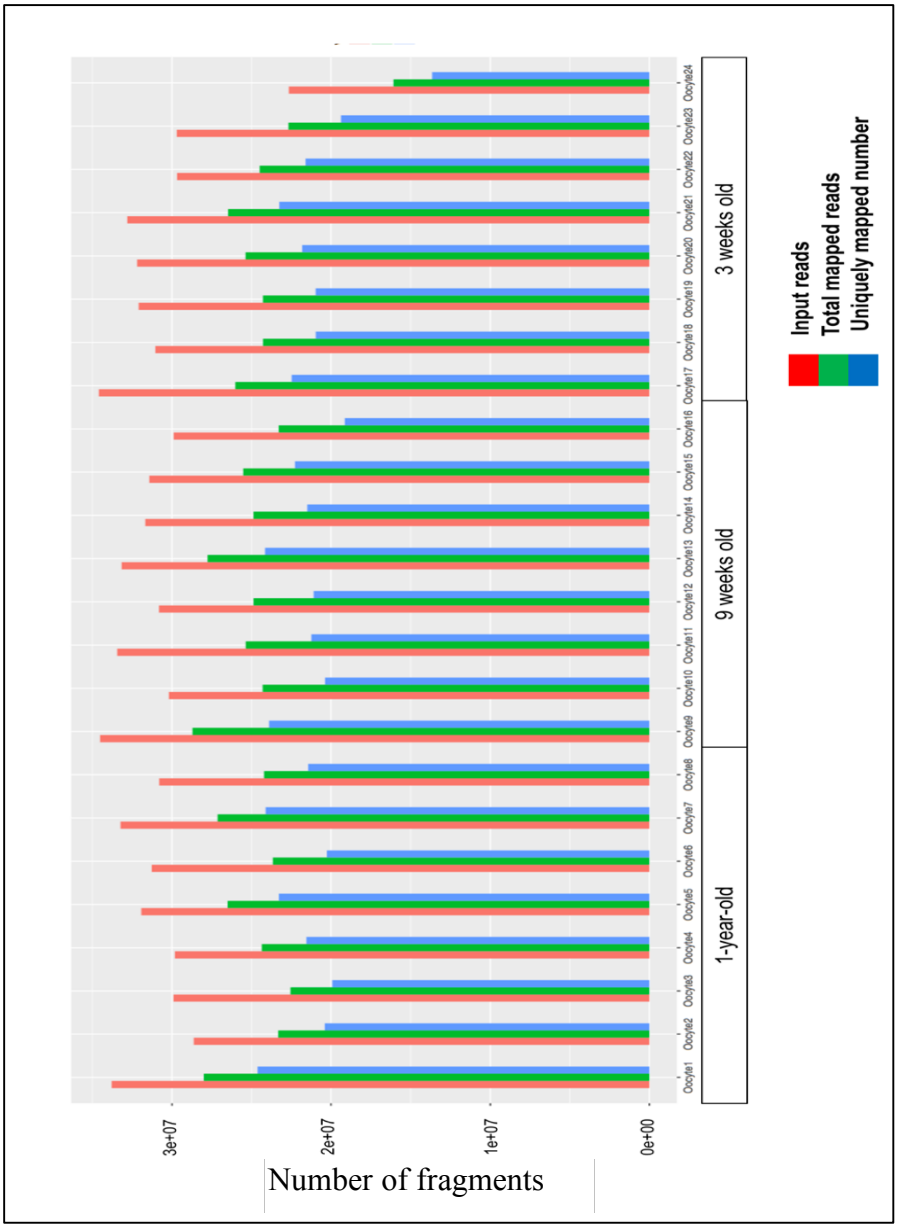


Figure 5.2 Mapping rate graph. Each oocyte (n=8) at 3 weeks, 9 weeks and 1 year of age generated around 30 million reads, around 28 million of which were mapped, and around 23 million of which mapped uniquely.

5.3.1.2. Oocyte-specific genes expression

To evaluate the validity of the generated sequencing data, a number of genes that are known to associated with oocyte function have been assessed to ensure that a high quality of sequencing was achieved. These genes include:

H1 Histone Family Member O Oocyte Specific (H1foo) (164), which is an oocyte-specific gene and is known to be essential in oocyte maturation. Growth differentiation factor-9 (Gdf9), is a member of the transforming growth factor-beta (TGFβ) and is known to be expressed in oocytes for ovarian folliculogenesis (165). Superfamily zona pellucida receptors (Zp2 and Zp3), which both mediate the initial interaction with sperm (166). Zygote arrest1 (*Zar1*) which is a novel maternal-effect gene critical for the oocyte-to-embryo transition (167). Bone marrow Protein (Bmp 15*)*, an *oocyte*-derived growth and differentiation factor, is a critical regulator of folliculogenesis and granulosa cells C activities (168). Wee1 homolog 2 (Wee2), acts as a key regulator of meiosis during both prophase I and metaphase II (169). Aurora kinase A (Aurka), which is a centrosome-localised serine/threonine kinase crucial for cell cycle control, is critically involved in centrosome maturation and spindle assembly (170). Enhancer of zeste homolog 2 (Ezh2), required for mouse oocyte meiotic maturation (171). Maternal embryonic leucine zipper kinase (Melk), it plays an important role in oocyte maturation and early embryonic development. F-box 43 (Fbxo43) protein coding gene is required to establish and maintain the arrest of oocytes at the second meiotic metaphase until fertilization (172). Cell division cycle family (Cdc25a, CDc25b and CDc25c) *are* members of the CDC25 family of phosphatases. *Cdc25a, which* is required for progression from the G1 phase to the S phase of the cell cycle *CDC25A*/B/C, are dual-specificity phosphatases and activate cyclin-dependent kinases (CDKs) that are essential for resumption of meiotic cell cycle (173). Sex

Comb On Midleg Like 2 (Scmi2) involved in DNA methylation. Phosphodiesterase Type 3A (*Pde3a),* which is essential for *oocyte* maturation and subsequent fertilisation (174). DNA (cytosine-5-)-methyltransferase 3 beta (Dnmt3b), is an enzyme required for DNA methylation (175).

The boxplot of the genes demonstrates the average gene expressions in the samples of each group (Figure 5.2). One example is the average of bmp15 expression in the eight 1-year-old oocytes. The range of expression, which is measured by Read Per Kilobase per Million (RPKM), for each gene indicates the consistency of the gene expression values in all samples at each age.

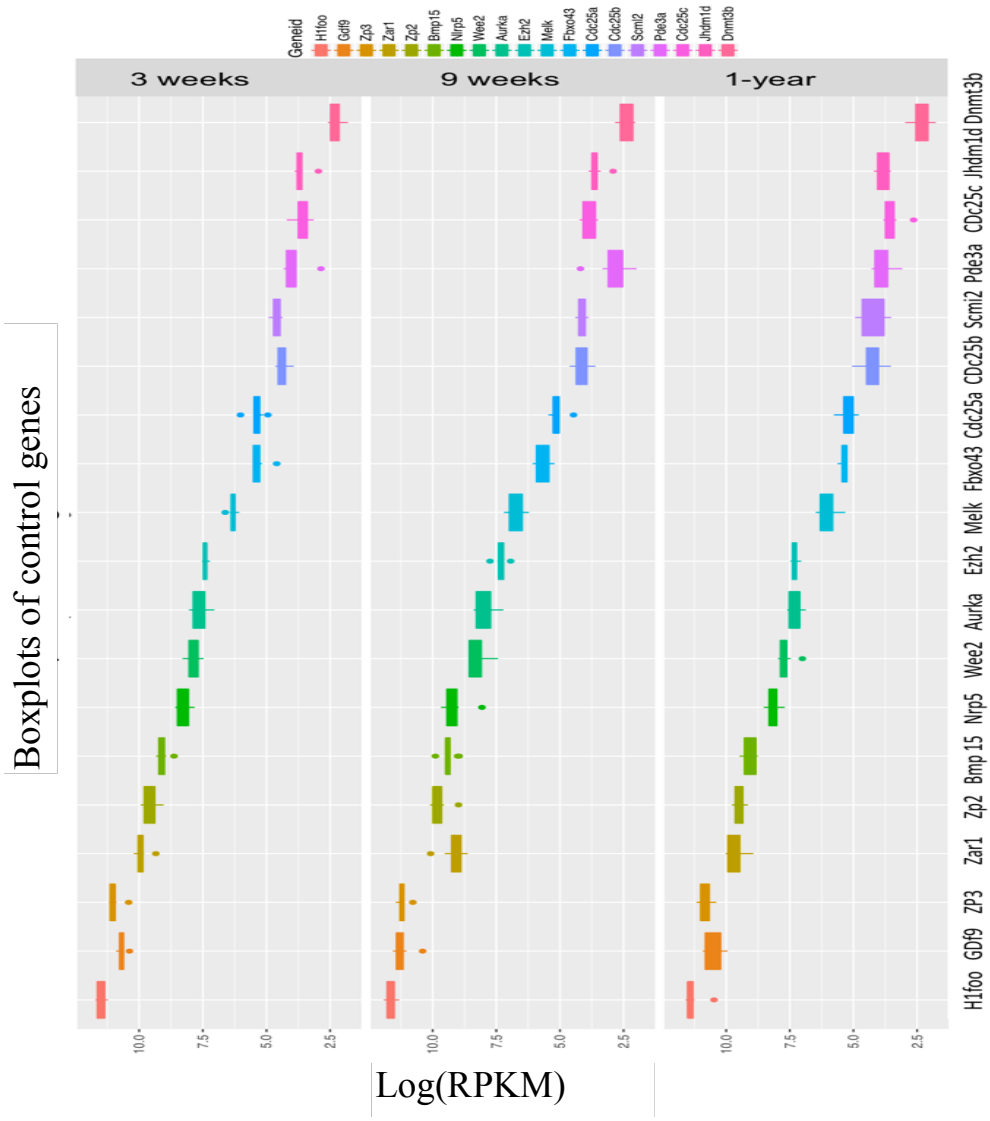


Figure 5.3 depicts a boxplot demonstrating the average expression of the gene-specific oocyte in the sequencing data. Boxes correspond the the first and third quartile; whiskers extend to 1.5 times the interquartile, and the dot indicates an outlier. Genes are sorted based from the highest to the lowest expression in each group.

5.3.1.3. Number of expressed genes (quality assessment)

Using the criteria that a gene is retained if expressed with a threshold of one count-per-million (cpm) for any given sample at each age group (Figure 5.3). The number of expressed genes is an important factor in assessing the complexity of sequencing library as a higher number indicates a deep sequencing depth was achieved. This is assessed by the number *o*f sequenced reads for a given sample (Figure 5.3).

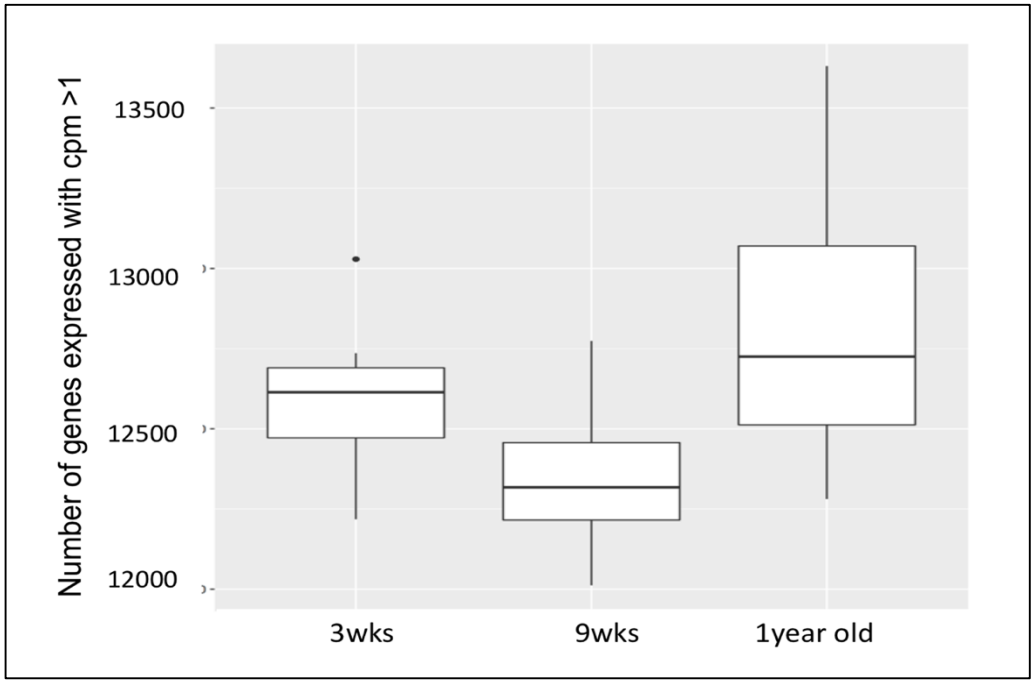


Figure 5.3 depicts a boxplot of the count per million (cpm), showing the number of genes expressed in oocytes in each age group. A threshold of 1 count-per-million (cpm) was used to deem a gene as expressed for any given sample. Boxes correspond to the first and third quartile. The horizontal line represents the median; whiskers extend to 1.5 times the interquartile range.

5.3.2. Principle Component Principle (PCA)

To investigate the gene expression pattern variations between oocytes within a given age, sample data were plotted along the first and second PCA to identify a single-cell sample with similar global gene expressions across different ages (Figure 5.4)

The generated sample clusters enable the evaluation of the similarity and the variability of the global transcriptome profiles between and within the age groups by revealing how principle components (each dot represents a sample) of all three ages relate to one another. The majority of oocytes clustered according to their developmental age: 3 weeks, 9 weeks and 1 year. Principle Component Analysis reveals that oocytes of a similar age have a distinct transcription pattern that separates them from the oocytes of the other two age groups. Clear transcriptomic differences were revealed between the oocytes collected from the different age groups. A large proportion of the variation of principal components is taken from two samples of the 9-week-old oocytes. These two oocytes are from the same mouse.

Although that each two oocytes were retrieved from the same mother, they do not seem to have a significant cluster apart from other samples in the same age group. This could support the assumption that each oocyte has a binary fate independent of other oocytes from the same mother. The oocytes from the same mother could be still sharing some features in comparison to other oocytes from different mice. But, samples in this PCA were scattered based on the most variable gene expressions that dominate the PCA, which apparently was age-related changes.

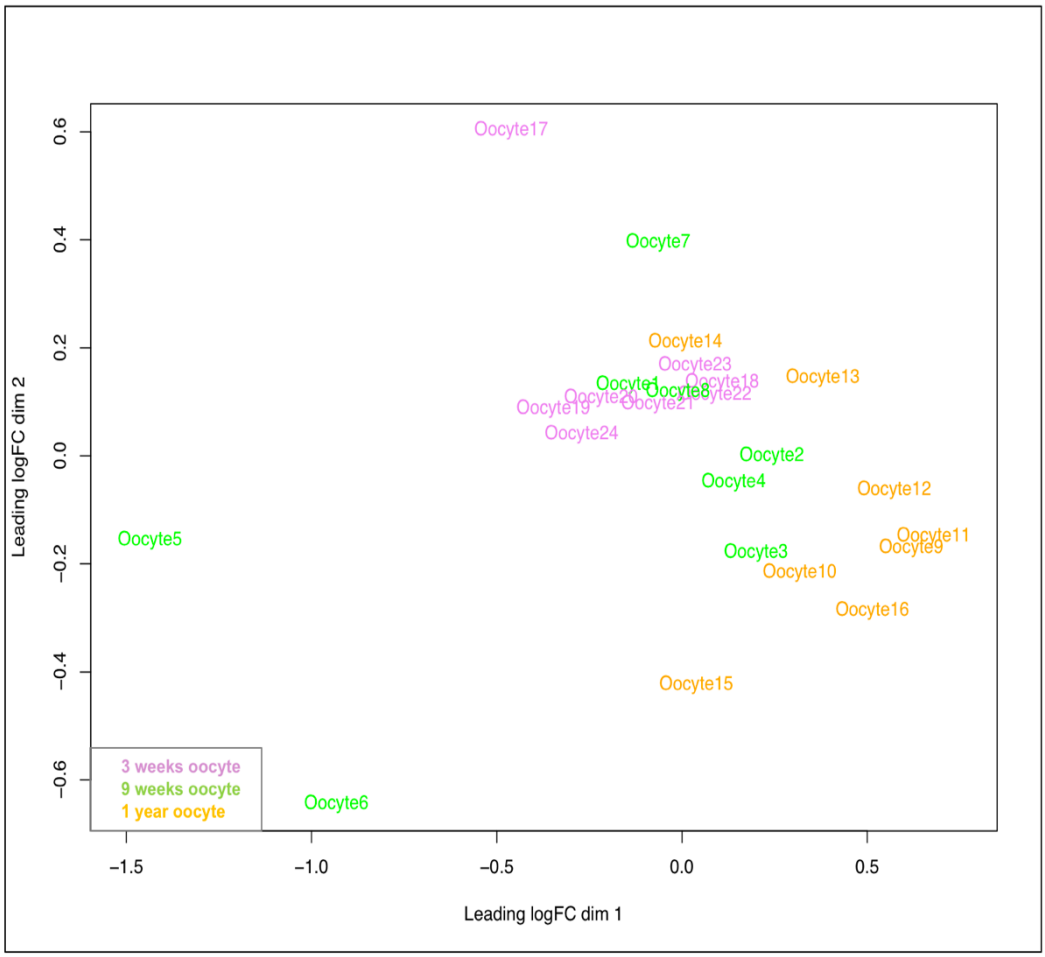


Figure 5.4 depicts the PCA of a mouse sc-RNA-seq transcriptome. Each point represents a single oocyte sample and is labelled according to age. Sample data were plotted along with first and second principle components.

5.3.2.1. Pre-pubertal versus adult oocytes (3 weeks vs 9weeks)

Oocytes at 3 weeks have clustered distinctly and closely together from those at the age of 9 weeks (Figure 5.5). The compact cluster of all the components of this group (the 3-week-old oocytes) that are not seen in other two clusters (9 weeks and 1 year) suggest that the oocytes have very similar transcription patterns before maturation takes place (before puberty). On the other hand, the variation in global gene patterns of oocytes become slightly obvious when oocytes reach 9 weeks and when the oocytes are completely mature, as shown in Figure 5.5.

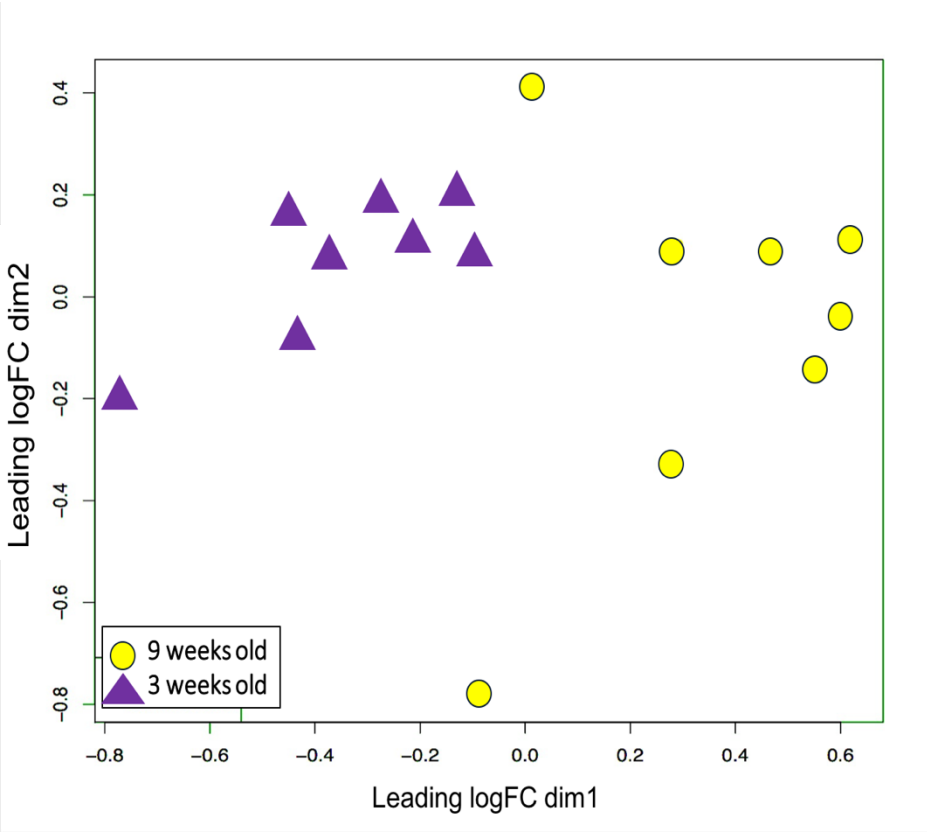
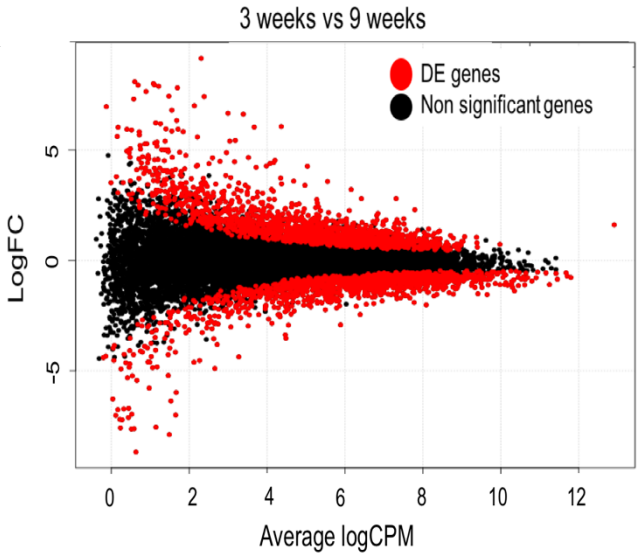


Figure 5.5 depicts the PCA of oocytes at 3 weeks of age (pre-pubertal) versus 9 weeks (adult). Oocytes from the immature 3-week-old mice cluster fairly separately to the oocytes of the 9-week-old mice.

5.3.3. Differential genes expression

A total of 2912 genes were identified as differentially expressed in the 3-week-old oocytes as compared to the oocytes from the 9-week-old mice. In the scatter plot, the x-axis is the average value of normalised counts (measured in cpm), and the y-axis shows the log2-fold change in expression. The significant differentially expressed genes ($P < 0.05$) are indicated in red in Figure 5.6. The range of the fold change is quite broad, showing many genes with log2-fold change up to 5 and several genes with a fold change greater than 5. Negative fold change values demonstrate downregulated genes. Additionally, the overlapping differentially expressed genes in both ages are colored in green (Figure 5.6).

A



B

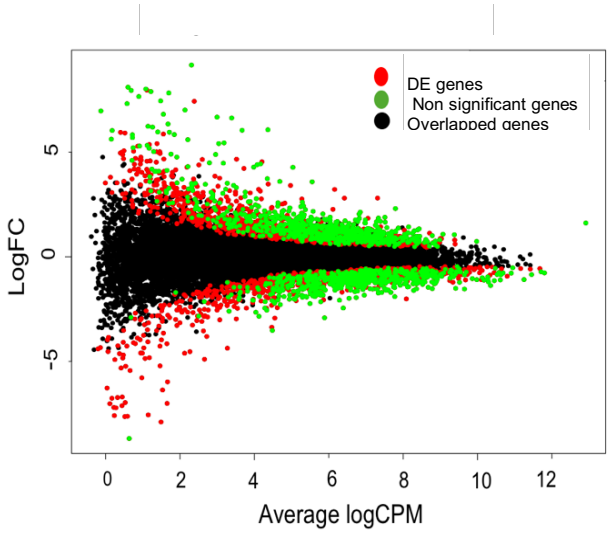


Figure 5.6 depicts a scatter plot of average counts (cpm) versus Log2 -fold changes for genes generated by DEseq. Genes were retained if they expressed a threshold of 1 count-per-million (cpm) for any given sample.

5.3.4. Overlapping genes between the three ages

About 1271 genes were commonly differentially expressed in both 3-week-old and 1-year-old oocytes, in comparison to a 9-week-old oocyte, in which 679 genes were upregulated and 592 were down-regulated. On the other hand, 893 and 136 upregulated genes were exclusive to 3-week-old and 1-year-old oocytes, and 748, 235 downregulated genes were exclusive to 3-week-old and 1-year-old oocytes, respectively (Figure 5.7)

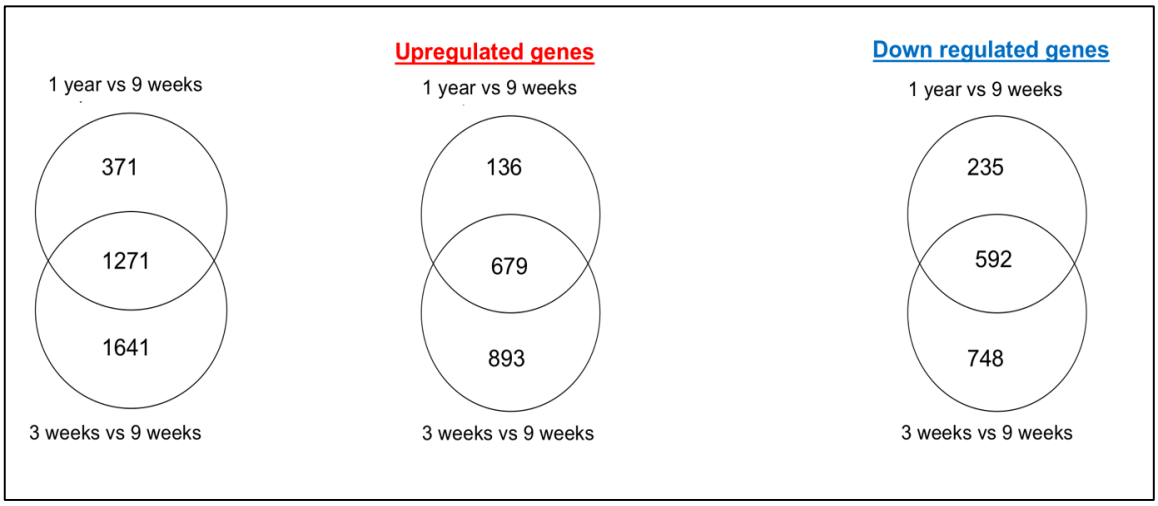


Figure 5.7 depicts a Venn diagram showing the number of genes differently regulated between the two comparisons.

5.3.5. Functional analysis

Ingenuity Pathway Analysis (IPA) was used to identify the canonical pathways and the regulatory network potentially involved in the regulation of immature oocyte development at a pre-puberty age. Using oocytes from 9-week-old mice as a baseline for conducting this comparison, differentially expressed genes were involved in 20 pathways at a *P* value 0.05. The canonical pathways were categorised based on the number of genes present from the dataset.

5.3.5.1. Canonical Pathways

Significant canonical pathways are shown in Figure 5.8. Activated metabolic pathways in the pre-pubertal oocyte indicate that oocyte transcriptome is undergoing global changes, which led to a different metabolism and gene expression pattern. This data represents a distinct transcriptional activity that is ongoing to acquire meiotic competence. Thus, changes are associated with the incomplete developmental competence of these oocytes in comparison with the adult oocytes at 9 weeks (Figure 5.8)

Additionally, a stacked bar chart was also generated to demonstrate the upregulated genes and downregulated genes in each pathway, along with the ratio (orange line) of the number of genes in a given pathway that met the cut-off criteria to the total number of the molecules involved in each pathway in the reference gene set (Figure 5.9).

A total of 579 differentially expressed genes with fold change more or less than 2. To give a clearer picture of the upregulated or downregulated genes in a specific canonical pathway, a list of DE genes and the top 20 canonical pathways are provided in Table 5.1.

Table 5.2: The top 20 enriched pathways in pre-pubertal oocytes with - log(P value)>2. Table demonstrates the number of Upregulated and downregulated genes in each pathways and the transcripts associated with each pathway, generated by Ingenuity Pathways Analysis (IPA)

Ingenuity Canonical Pathways	-log (p-value)	Down-regulated	Up-regulated	Molecules
Signaling by Rho Family GTPases	5.72	18/252 (7%)	36/252 (14%)	FGFR1,MRAS,SEPT14,MAPK10,PIP5KL1,GNB1,MAP3K10,ACTA2,MAP2K1,PAK5,ELK1, GNAI3,PIK3CD,ACTB,GNG4,RHOT1,ARPC3,RAF1,EZR,MYL9,GNB2,PTK2B,MAP3K21, MYLK, RHOU,MAP2K7,ACTR2,GNB4,TLR9,IRS1,SEPT12,VIM,CDH17,GNAI2,ARHGEF7,PTPN11,GRB2, GNA14, CDH1,MAPK8,MAPK3,ARHGEF18,FNBP1,MAP3K11,SLC9A1,PPP1R12C,SEPT2,RELA, MAP2K4,ARHGEF1,MAP2K2,SEPT9,CDC42EP2,PAK4
ILK Signaling	5.32	9/197 (5%)	36/197 (18%)	PPP2R5E,FLNB,FGFR1,AKT3,DSP,RSU1,MAPK10,RICTOR,TLR9,IRS1,PPP2R3A,ACTA2, GSK3A,PARVA,ILK,TESK1,DOCK1,VIM,PIK3CD,FN1,PTPN11,ACTB,GRB2,CREBBP,RHOT1,CDH1, MAPK8,ACTN4,PPP2R1B,MAPK3,PPP2CA,FNBP1,PPP2R5B,FLNA,EP300,ACTN1,MYL9,PARVB, RELA,ATF4,CREB3L4,MAP2K4,HIF1A,MAP2K6,RHOU
PXR/RXR Activation	4.18	2/65 (3%)	13/65 (20%)	PCK2,HMGCS2,AKT3,ALAS1,NCOA1,CPT1A,ABCB9,PRKACB,RELA,GSTM1,ALDH1A1,SCD,RXRA, FOXO1,CYP7A1

Ingenuity Canonical Pathways	-log (p-value)	Down-regulated	Up-regulated	Molecules
Xenobiotic Metabolism Signaling	3.8	12/290 (4%)	41/290 (14%)	PPP2R5E,NCOR2,FGFR1,MRAS,MAP3K10,MAP2K1,SCAND1,ALDH9A1,NDST2,PIK3CD, ALDH1A1,PRKCB,HS6ST1,RXRA,ARNT,PPP2R1B,NCOA1,ALDH7A1,RAF1,ALDH6A1, CAMK1D, UST,CITED2,UGT2B28,MAP2K6,MAOB,MAP2K7,MAPK14,TLR9,IRS1,PPP2R3A,GSTM1,PTPN11, GSTM5,GRB2,CREBBP,MAPK8,UGT8,CUL3,MAPK3,NRIP1,PPP2CA,Map3k7,PPP2R5B, KRAS, MAP3K5,MAP3K11,EP300,Sult1d1,RELA,MAP2K4,Fmo6,MAP2K2
RANK Signaling in Osteoclasts	3.51	7/102 (7%)	22/102 (22%)	FGFR1,MAP2K7,TRAF2,MAPK14,AKT3,MAPK10,TLR9,MAP3K10,IRS1, TAB2,MAP2K1,ELK1,BIRC3,PIK3CD,PTPN11,GRB2,XIAP,MAPK8,MAPK3,Map3k7, MAP3K5,MAP3K11,RAF1,RELA,Calm1 (includes others),MAP2K4,PTK2B,MAP2K2,MAP2K6
Type II Diabetes Mellitus Signaling	3.35	11/154 (7%)	23/154 (15%)	FGFR1,MAP2K7,TRAF2,AKT3,SOCS5,SOCS4,PRKAA1,MAPK10,TLR9,IRS1,CACNA1I,SOCS7,PIK3CD,PTPN11,CACNA2D3,GRB2,PRKCB,ADIPOR2,MAPK8,MAPK3 MAP3K5,SMPD1,SOCS2,TRADD,RELA,CACNA1H,CD36,MAP2K4,CEBPB,ACSBG1, PRKAA2,CACNA2D1,CACNB3,CACNG4
Type II Diabetes Mellitus Signaling	3.35	11/154 (7%)	23/154 (15%)	FGFR1,MAP2K7,TRAF2,AKT3,SOCS5,SOCS4,PRKAA1,MAPK10,TLR9,IRS1,CACNA1I,SOCS7,PIK3CD,PTPN11,CACNA2D3,GRB2,PRKCB,ADIPOR2,MAPK8,MAPK3 MAP3K5,SMPD1,SOCS2,TRADD,RELA,CACNA1H,CD36,MAP2K4,CEBPB,ACSBG1, PRKAA2,CACNA2D1,CACNB3,CACNG4
GP6 Signaling Pathway	3.25	2/134 (1%)	18/134 (13%)	FGFR1,AKT3,TLN1,COL6A1,COL3A1,TLR9,IRS1,COL1A2,COL4A3,GSK3A,COL1A1,LAMA5, Calm1 (includes others),PIK3CD,COL4A3BP,PTPN11,COL4A2,GRB2,PRKCB,LYN

Ingenuity Canonical Pathways	-log (p-value)	Down-regulated	Up-regulated	Molecules
LPS/IL-1 Mediated Inhibition of RXR Function	3.03	13/222 (6%)	17/222 (8%)	HMGCS2,MAP2K7,TRAF2,ACOX3,ALAS1,CPT1A,ALDH9A1,NDST2,GSTM1, ALDH1A1,GSTM5,HS6ST1,RXRA,IL1RAP,CPT2,FABP5,MAPK8,NCOA1,ALDH7A1 ,ABCB9,ALDH6A1,ABCG1,Sult1d1,UST,MAP2K4,Fmo6,IL18RAP,ACSBG1,MAOB,CYP7 A1
Germ Cell-Sertoli Cell Junction Signaling	2.93	9/173 (5%)	36/173 (21%)	FGFR1,MAP2K7,MAPK14,EPN2,TJP1,MRAS,MAPK10,TLR9,MAP3K10,IRS1,ACTA2, TUBA1C,MAP2K1,ILK,CTNND1,PAK5,ZYX,PIK3CD,PTPN11,ACTB,PLS1,GRB2,RHOT1 ,CDH1, MAPK8,ACTN4,MAPK3,FNBP1,Map3k7,CTNNA1,KRAS,MAP3K5,MAP3K11,ITGA6,AC TN1, EPN1,MAP2K4,TUBB2B,MAP2K2,BCAR1,TUBA1B,TUBA4A,PAK4,MAP2K6,RHOU
RhoA Signaling	2.89	6/124 (5%)	14/124 (11%)	PLEKHG5,ACTR2,ARPC3,SEPT14,PIP5KL1,EZR,ACTA2,MYL9,SEPT12,SEPT2,ABL2, PLXNA1,PFN1,ACTB,PTK2B,ARHGEF1,MYLK,SEPT9,CDC42EP2,ARHGAP9
Cholecystokinin/Gastrin-mediated Signaling	2.87	6/101 (6%)	16/101 (16%)	KRAS,ITPR2,MAPK14,MRAS,MAPK10,RAF1,MAP2K1,MEF2C,ELK1,SRF,MAP2K4, PTK2B,MAP2K2,GRB2,PRKCB,BCAR1,RHOT1,MAPK8,MAP2K6,RHOU,MAPK3,FNBP1

Ingenuity Canonical Pathways	-log (p-value)	Down-regulated	Up-regulated	Molecules
Sertoli Cell-Sertoli Cell Junction Signaling	2.84	13/178 (7%)	33/178 (19%)	CLDN10,MAP2K7,MAPK14,EPN2,AKT3,TJP1,MRAS,MAPK10,CLDN8,MAP3K10,ACTA2, GSK3A,TUBA1C,MAP2K1,PRKACB,ILK,ELK1,TJAP1,ACTB,PLS1,MPP6,CDH1,MAPK8, ACTN4, CLDN20,MAPK3,Map3k7,CTNNA1,KRAS,JAM2,OCLN,MAP3K5,MAP3K11,RAF1,ACTN1,CGN, EPN1,MAP2K4,CLDN9,TUBB2B,MAP2K2,BCAR1,TUBA1B,TUBA4A,SYMPK,JAM3
SAPK/JNK Signaling	2.79	6/104 (6%)	15/104 (14%)	KRAS,FGFR1,MAP2K7,TRAF2,MAP3K5,MAP4K2,MAP3K11,HNRNPK,MRAS,MAPK10, GNB1,TLR9,MAP3K10,IRS1,TRADD,ELK1,PIK3CD,PTPN11,MAP2K4,GRB2,MAPK8
Mechanisms of Viral Exit from Host Cells	2.78	0/41 (0%)	9/41 (22%)	SH3GL1,SNF8,LMNB1,SH3GLB1,ACTB,CHMP4B,PRKCB,ACTA2,LMNB2
Hepatic Fibrosis / Hepatic Stellate Cell Activation	2.67	5/187 (3%)	14/187 (7%)	FGFR1,COL6A1,COL3A1,CTGF,IGFBP5,COL1A2,COL4A3,ACTA2,MYL9,COL1A1,SMAD4, IGF2,RELA,EGF,FN1,COL4A3BP,IL18RAP,COL4A2,IL1RAP
Gα12/13 Signaling	2.64	6/135 (4%)	23/135 (17%)	FGFR1,MAP2K7,AKT3,MRAS,MAPK10,TLR9,IRS1,MAP2K1,ELK1,CDH17,F2R,PIK3CD, PTPN11,GRB2,CDH1,MAPK8,F2,MAPK3,KRAS,MAP3K5,RAF1,MYL9,MEF2C,RELA,MAP2K4,PTK2B,ARHGEF1,MAP2K2,RASA1

Ingenuity Canonical Pathways	-log (p-value)	Down-regulated	Up-regulated	Molecules
Integrin Signaling	2.6	13/219 (6%)	40/219 (18%)	FGFR1,MRAS,ACTA2,MAP2K1,ILK,PAK5,ZYX,PIK3CD,ACTB,ARF3,RHOT1,ACTN4,ARHGAP26, NCK1,RAP2A,ARPC3,TLN1,RAF1,ACTN1,MYL9,ABL1,MYLK,BCAR1,ARF6,RHOU,TNK2,AKT3, ACTR2,TLR9,IRS1,PARVA,Arf2,DOCK1,ARHGEF7,PFN1,PTPN11,GRB2,GIT1,MAPK8,RAPGEF1, CAPNS1,MAPK3,CAPN2,FNBP1,KRAS,MAP3K11,ITGA6,PARVB,ITGA9,TSPAN3,MAP2K4, MAP2K2,PAK4
B Cell Receptor Signaling	2.59	9/191 (5%)	40/191 (21%)	FGFR1,MRAS,MAP3K10,INPPL1,MAP2K1,ELK1,PIK3CD,INPP5B,PRKCB,RPS6KB1,RAP2A, RAF1,ABL1,Calm1 (includes others),CREB3L4,PTK2B,BCL2L1,MAP2K6,POU2F2,LYN,MAP2K7,MAPK14,AKT3, TLR9,IRS1,GSK3A,INPP5D,PTPN11,GRB2,CREBBP,MAPK8,PTPN6,FOXO1,SYNJ1,MAPK3,Map3k7,KRAS,TCF3,MAP3K5,MAP3K11,EP300,EGR1,MEF2C,RASSF5,RELA,ATF4, MAP2K4,CSK,MAP2K2
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	2.54	8/194 (4%)	26/194 (13%)	PPP2R5E,FGFR1,MAP2K7,MAPK14,AKT3,MAPK10,TLR9,MAP3K10,IRS1,PPP2R3A,MAP2K1,PPP1R3D, PIK3CD,PPP1R10,PTPN11,GRB2,PRKCB,CREBBP,RHOT1,MAPK8,PTPN6,PPP2R1B,MAPK3,PPP2CA, FNBP1,Map3k7,PPP2R5B,MAP3K5,MAP3K11,MPO,RELA,MAP2K4,ARG2,RHOU

Table 5.2: demonstrates genes involved in signaling by Rho of GTPases, generated by IPA

Symbol	Entrez Gene Name	Expression Log Ratio	False Discovery Rate (q-value)
ACTA2	actin, alpha 2, smooth muscle, aorta	6.633	0.000182
CDH17	cadherin 17	-2.197	0.00248
ELK1	ELK1, ETS transcription factor	2.085	0.0224
EZR	Ezrin	2.174	0.00185
FNBP1	formin binding protein 1	-2.331	0.0489
IRS1	insulin receptor substrate 1	4.18	0.0101
MAP3K10	mitogen-activated protein kinase kinase kinase 10	3.327	0.000807
MAP3K11	mitogen-activated protein kinase kinase kinase 11	2.211	0.00000932
MAP3K21	mitogen-activated protein kinase kinase kinase 21	-2.822	0.0112
MAPK10	mitogen-activated protein kinase 10	-2.143	0.0275
MYL9	myosin light chain 9	9.163	0.000123
MYLK	myosin light chain kinase	4.277	0.00227
PIP5KL1	phosphatidylinositol-4-phosphate 5-kinase like 1	-4.089	0.0362
PPP1R12C	protein phosphatase 1 regulatory subunit 12C	4.428	0.0126
PTK2B	protein tyrosine kinase 2 beta	2.183	0.0225
RHOU	ras homolog family member U	3.294	0.0164
SEPT9	septin 9	4.189	0.0108
TLR9	toll like receptor 9	2.095	0.00851
VIM	Vimentin	6.043	0.000353

Table 5.3: demonstrates genes involved in ILK signaling pathway.

Symbol	Entrez Gene Name	Expression Log Ratio	False Discovery Rate (q-value)
ACTA2	actin, alpha 2, smooth muscle, aorta	6.633	0.000182
DSP	Desmoplakin	3.597	0.0000558
FLNA	filamin A	2.323	0.0444
FNBP1	formin binding protein 1	-2.331	0.0489
GSK3A	glycogen synthase kinase 3 alpha	2.596	0.000654
IRS1	insulin receptor substrate 1	4.18	0.0101
MAP2K6	mitogen-activated protein kinase kinase 6	-2.139	0.0287
MAPK10	mitogen-activated protein kinase 10	-2.143	0.0275
MYL9	myosin light chain 9	9.163	0.000123
PARVA	parvin alpha	4.567	0.0096
PARVB	parvin beta	3.715	0.00667
PPP2R1B	protein phosphatase 2 scaffold subunit Abeta	-2.125	0.000364
PPP2R5B	protein phosphatase 2 regulatory subunit B'beta	2.121	0.00305
RHOU	ras homolog family member U	3.294	0.0164
TLR9	toll like receptor 9	2.095	0.00851
VIM	Vimentin	6.043	0.000353

Table 5.4: demonstrates genes involved in PXR/RXR Activation signaling pathway.

Symbol	Entrez Gene Name	Expression Log Ratio	False Discovery Rate (q-value)
ALDH1A1	Aldhyde dehydrogenase	2.063	0.00243
CPT1A	Cartnitine	3.588	0.0048
CYP7A1	Cytochrome P450 family 7	-4.272	0.0002
GSM1	Glutathione S-transferase	2.557	0.0031
HMGC2	3-hydoxy-3-	2.861	0.0026
PCK2	Phoshoenolpyruvate	7.894	0.0001
RXRA	Retinoid X receptor alpha	2.404	0.0001
SCD	Stearoyl-CoA desaturase	4.077	0.0002

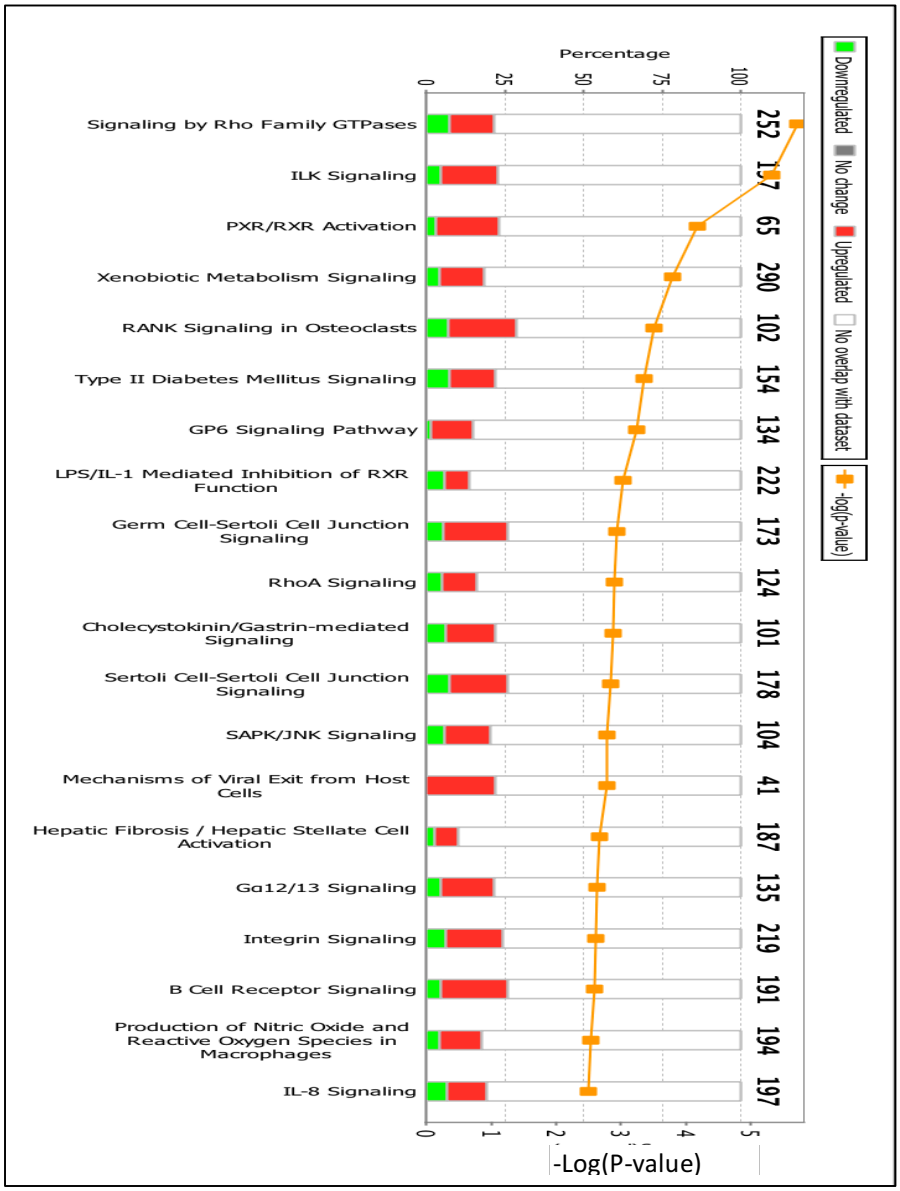


Figure 5.8 depicts a stacked bar chart of the top 20 canonical pathways in a pre-pubertal oocyte, compared to an adult oocyte. The upregulated genes (red colour) and downregulated genes in each pathway are demonstrated, along with the ratio (orange line), which is the number of genes in a given pathway that meet the cut-off criteria in proportion to the total number of molecules involved in each pathway in the reference gene set.

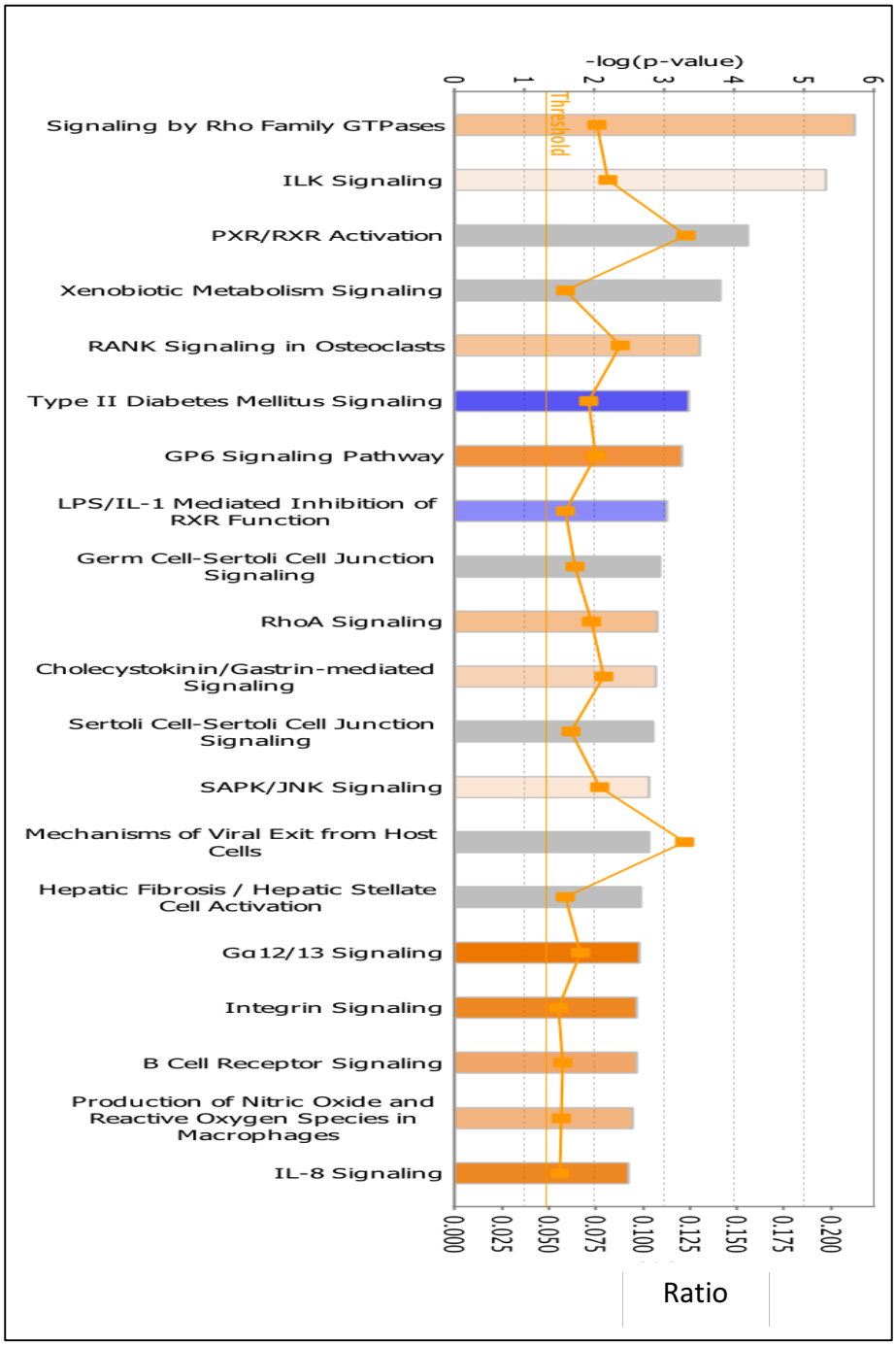


Figure 5.9 The 20 top significant pathways activated in a pre-pubertal oocyte, as compared to an adult oocyte

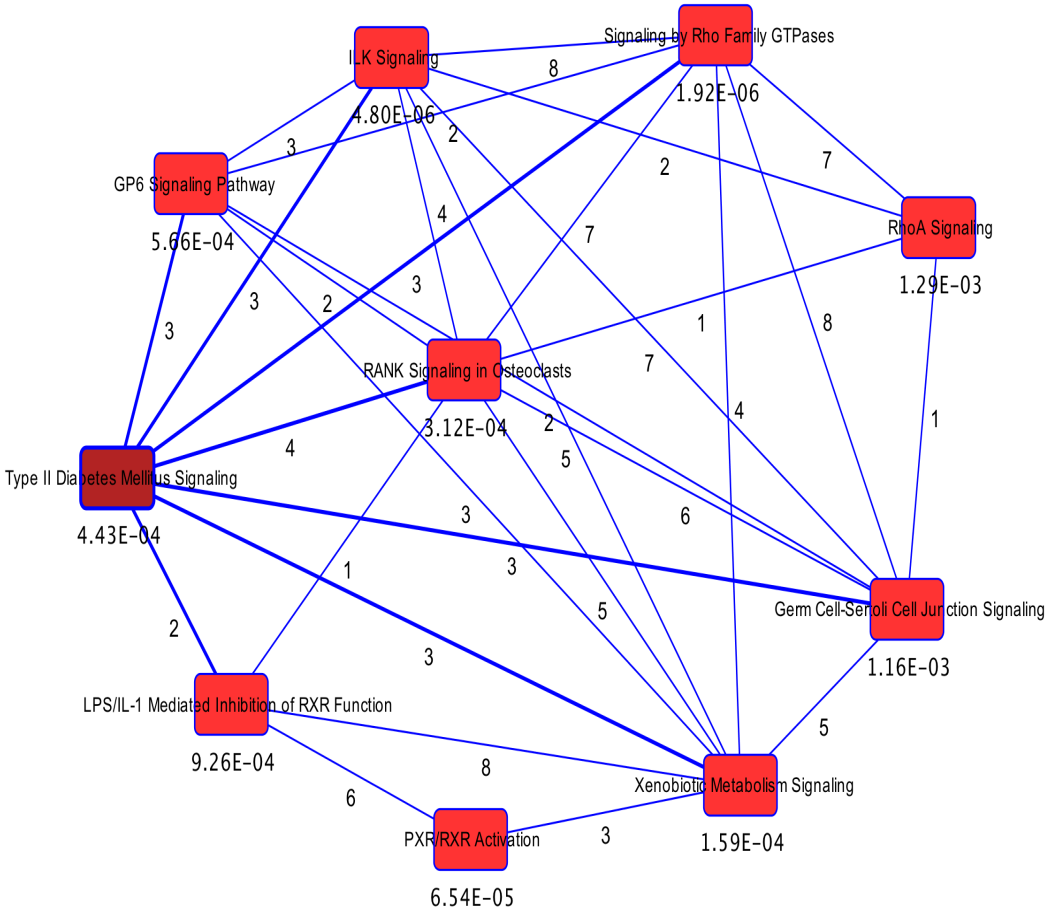


Figure 5.10 Overlapping of the top ten canonical pathways generated by IPA. P values and the number of common genes between the two pathways. This represents the interconnectivity among activated signaling pathways.

.

5.3.6. Upstream regulators

Based on the functional analysis of the general data set, IPA was used to predict the upregulator effect of the activated pathways. Transforming growth factor beta family (TGF-β) is the candidate regulator, of which the most significant expressed genes are upregulated (red colour) or downregulated (green). Figure 5.11 shows the direction of the relationship between gene and the TGFβ .

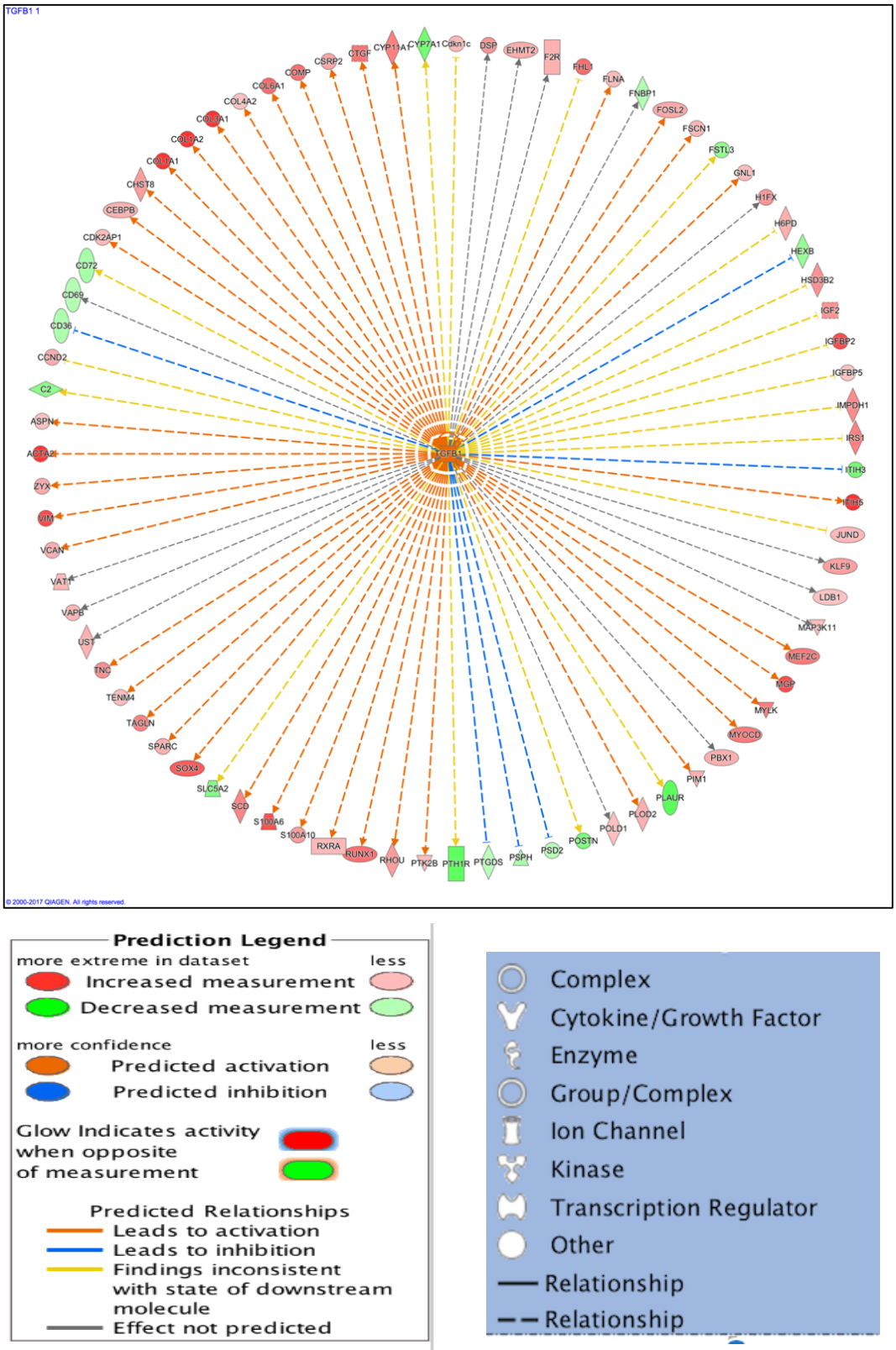


Figure 5.11 TGFβ family is predicted as the upregulator effect in pre-pubertal oocyte.

5.3.7. Networks identification

Using IPA, differential genes were grouped based on their molecule connectivity and cellular functions. The top five networks were identified as shown in Table 5.2.

Table 5.5: Top five networks involved in pre-pubertal oocyte

Network	Score	Node	Molecules
Developmental Disorder, Organismal Injury and Abnormalities, Reproductive System Disease	44	28	AMHR2,ATP4A,CAPN2,CEBPB,CENPI,Cg,CHST8,COLEC12,CTDNEP1,CUEDC1,CYP11A1,Eif4g, FAM189B,FHL1,FOXL2,FSH,FSTL3,HSD3B2,KCNT2,Lh,LIMCH1,Mitochondrial complex 1,MNT,NDUFAF1,NR5A1,Pka,PTPase,SMARCA1,SMARCD3,SOX4,STEAP1,SUN2,TENM4,TMEM255A, Tmsb4x (includes others)
Cellular Assembly and Organisation, Cardiovascular System Development and Function	26	39	ARHGAP11A,BASP1,CD72,CDK5R2,Cr3,CTTNBP2,ERK,FLNA,FXR2,GK,Glycogen synthase,H1FX,Hist1h1b,Hist1h1e,Histone H1,KIF1A,LGALS1,MYO5C,PTEFb,PFKFB1,PP2A,PPP1R9B,Ppp2c,PPP2R1B,PPP2R5B,PURB,Rnr,RPSA,Rsk,SEPT9,SPSB1,SYNPO,USF2,ZFP36L1,ZNF358
Inflammatory Disease, Skeletal and Muscular Disorders, Cellular Development	25	37	Akt,ALDH1A1,Ap2 alpha,ATP1B2,Atrial Natriuretic Peptide,Collagen,Alpha1,COMP,CPT1,CPT1A,CYP19,DDAH1,DDAH2,DIP2A,EIF4EBP2,GRB14,Igf,Igfbp,IGFBP2,IGFBP5,INSM1,Marcks,METRNL,NDRG2,PBXIP1,PHLDB1,PLAC1,POSTN,PRKAA,PTGDS,RBP1,SCD,SCGB3A1,Smad2/3-Smad4,TNC,VCAN
Cell Cycle, Dental Disease, Developmental Disorder	24	35	143,CaMKII,Creb,CRIM1,DLG2,DYNC1H1,EGLN,EID1,FOSL2,IRF2BP1,JUND,KLC2,Mapk,MAZ,MED16,mediator,Mlc,MLLT1,MLLT6,MYL9,MYLK,NHS,Notch,PDE5A,Pka catalytic subunit,PP1 protein complex group,PTGR2,SIRPB1,SOX12,SYNCRIP,TJAP1,trypsin,VARS,ZBED6,ZC3H10
Cell Death and Survival, Cellular Assembly and Organization, Cellular Compromise	19	25	ARHGDIG, atypical protein kinase C,BCR (complex),Caspase 3/7,CD69,CDKL4,DEPDC1B,EFNA1,Filamin,ISG20L2,Jnk,KIF26A,LARS2,MAP3K10,MAP3K11,MAP3K21,MAPK10,ME3,MLK,Mt1,Mt2,NCK,Pak,PLC gamma,Ptk,RAP2C,Rho gdi,RHOU,S100A6,Sapk,Sos,SYDE1,SYK/ZAP,TCF,VAV

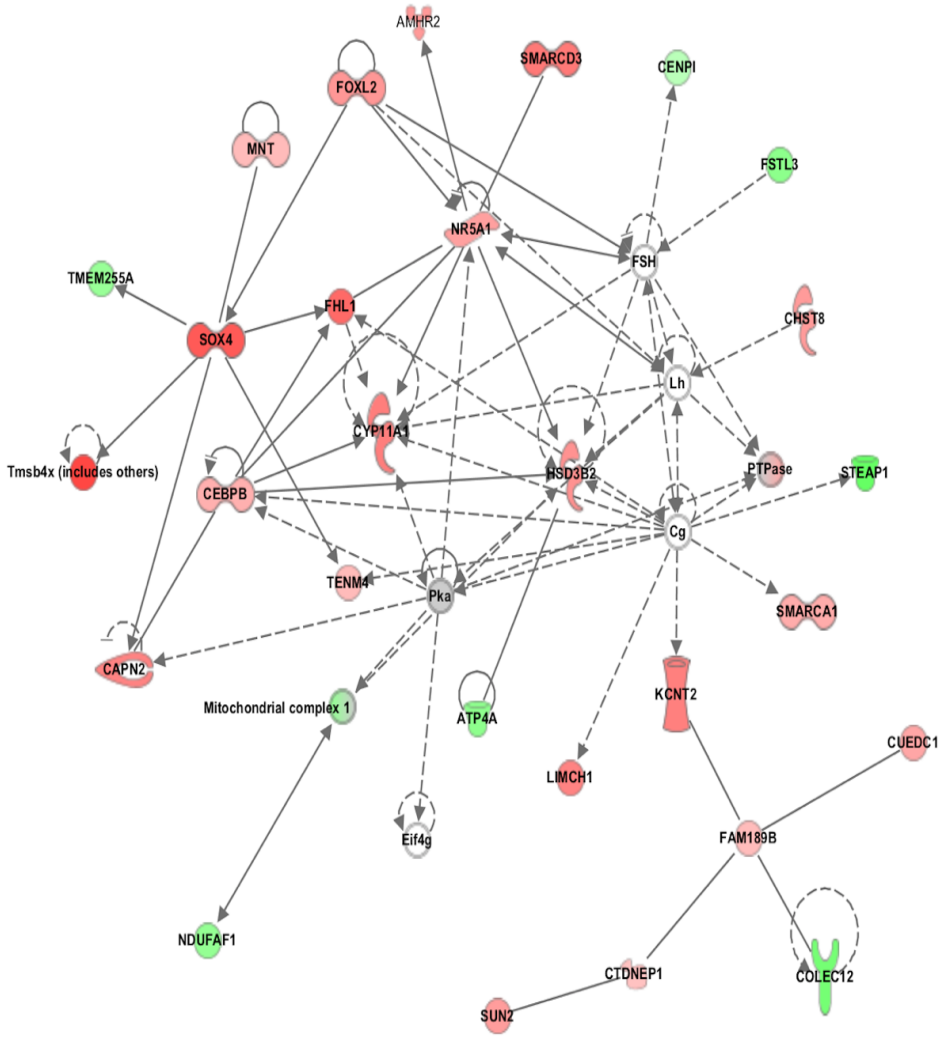


Figure 5.12 Network 1: Development Disorder, organismal injury and abnormalities, reproductive system disease. One of the networks involved in pre-pubertal oocyte

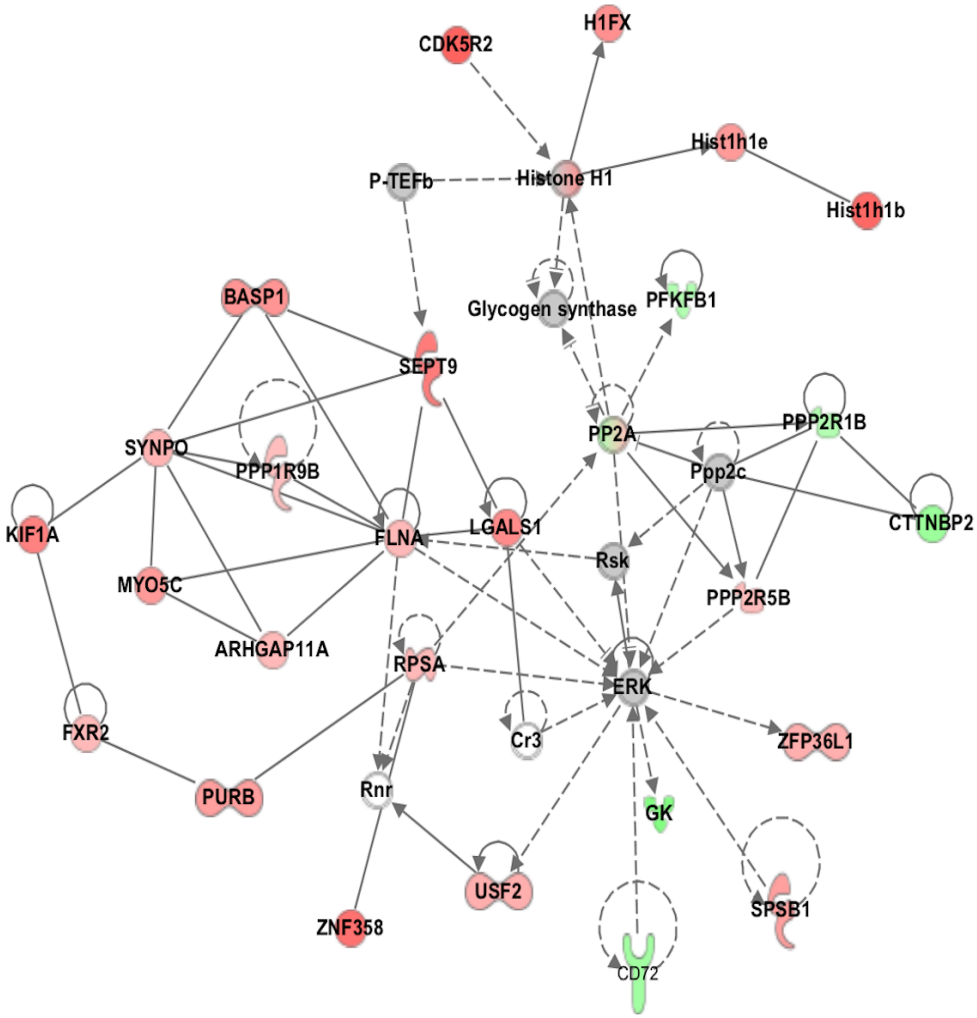


Figure 5.13 Second network involved in pre-pubertal oocyte: Cellular assembly and organisation, cardiovascular system development and function.

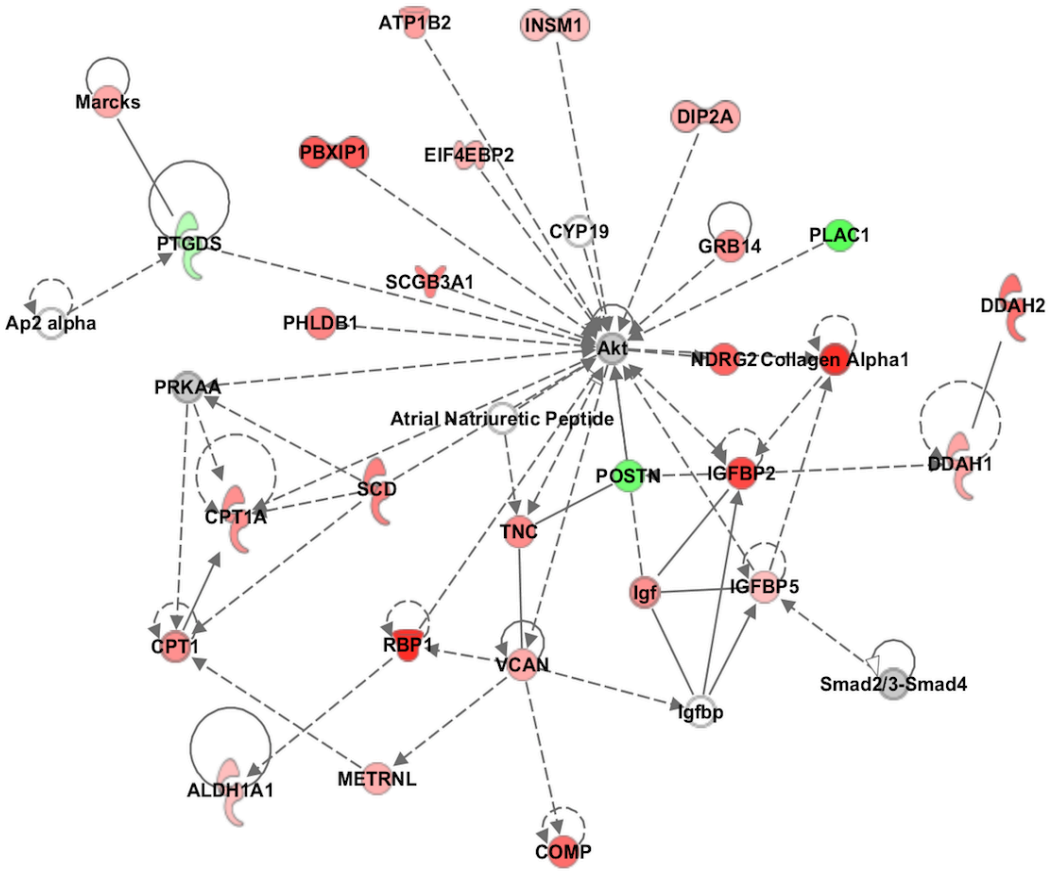


Figure 5.14 Third network involved in pre-pubertal oocyte: Inflammatory Disease, skeletal and muscular disorders, cellular development network.

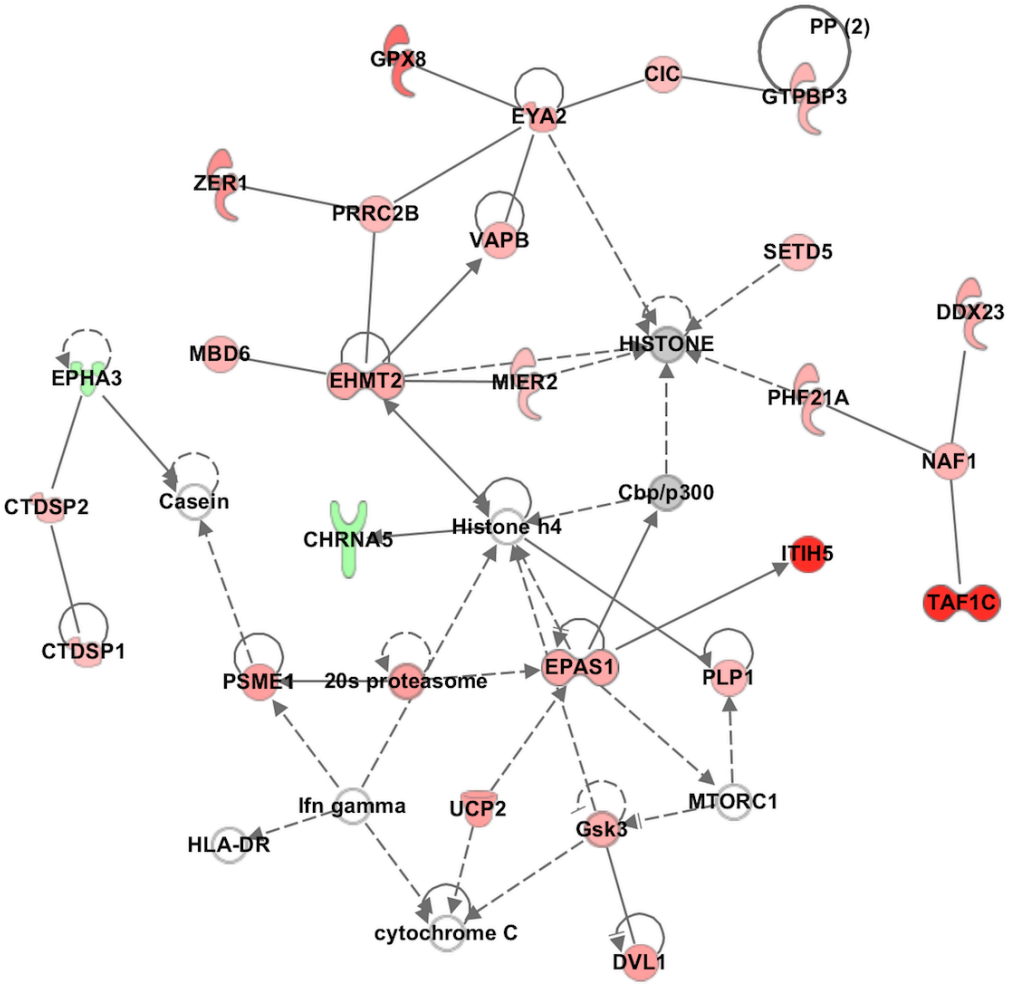


Figure 5.15 Fourth network involved in pre-pubertal oocyte:: Cell cycle, dental disease, developmental disorder.

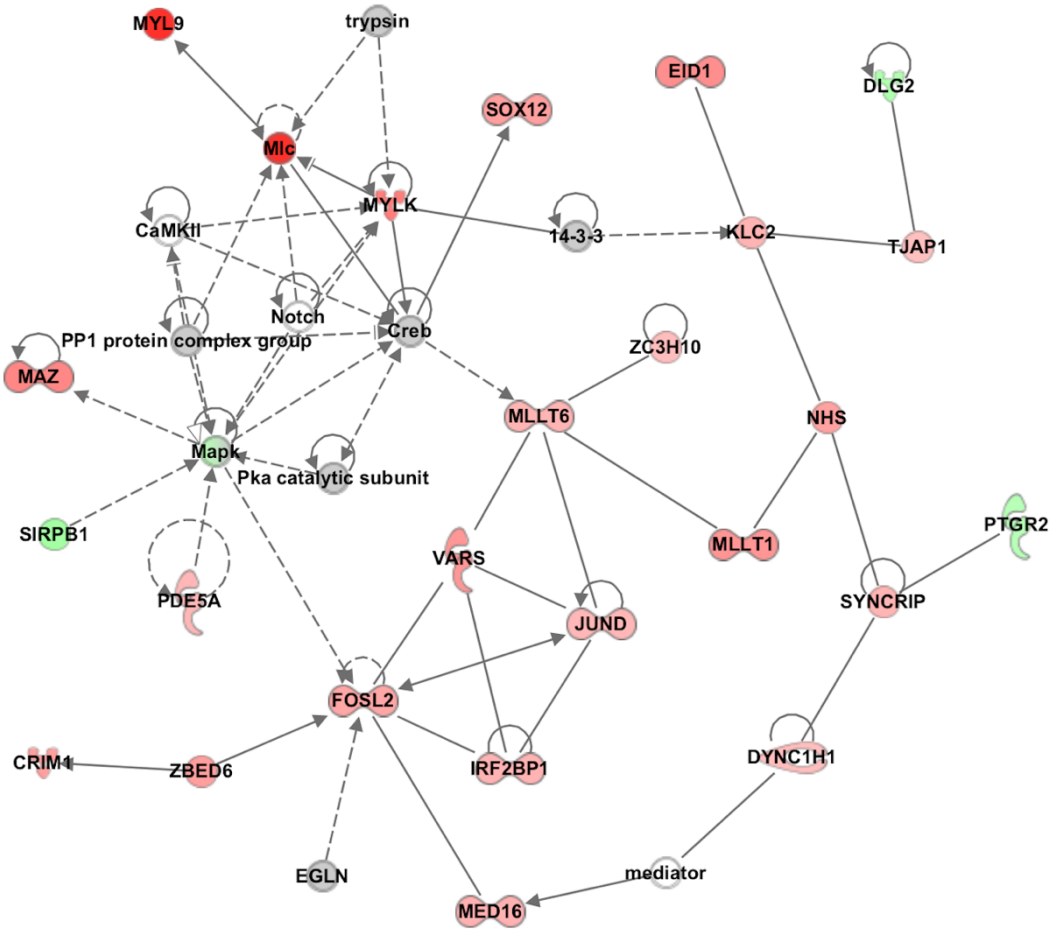


Figure 5.16 Fifth network involved in pre-pubertal oocyte: Cell death and survival, cellular assembly and organisation, cellular compromise.

5.4. Discussion

Single-cell RNA –seq was performed to characterise the transcriptome profiling of the pre-pubertal oocytes recovered from mice at 3 weeks of age and compared with oocytes of sexual mature mice at 9 weeks old. Single-cell transcriptomics was used for the first time to shed light on the molecular pathways and the cellular mechanisms that determine the competence of pre-pubertal oocyte. Knowing the molecular mechanisms underlying oocyte competence enables potential improvement of assisted reproductive technologies.

A limited focus on the ongoing transcriptional activity during the oocyte at pre-meiotic stages (before puberty) has led to slow progress in improving the competence of the pre-pubertal oocyte. Previous studies have investigated transcript abundance in oocytes derived from pre-pubertal goats, pigs and cattle and made comparisons to their adult counterparts (40,176). While these studies have revealed a number of candidate genes related to developmental competence, it is implausible to evaluate the gene expressions as the *in vitro* maturation environment could be deleterious to the level of gene expressions. Oocytes acquire cytoplasmic maturation after a long and complex process that involves the transcription and translation of genes during maturation, which is followed with a massive destruction of transcripts after maturation. Therefore, gene expression could have been altered.

The developmental competence of oocytes from the pubertal mice is inferior to that of the sexually mature mice (30,31,34,157,161). Using 9-week-old oocytes, which are at their prime reproductive age, as models of high developmental competence has allowed comparison the quantities of metabolites between the two ages. Single cell transcriptomics

combined with bioinformatics analysis was able to reveal new intracellular pathways that appear to be determinant for the cytoplasmic maturation and involved in the the developmental potential of the oocytes.

The findings showed that deep RNA sequencing of single mouse oocytes is achievable. Approximately 30 million 75 bp reads were generated per oocyte, of which 70% uniquely mapped to the mouse genome reference. The high mapping rate, the number of reads achieved and the number of expressed genes that were detected with the analysis demonstrate the ability to successfully produce an Illumina library from a single mouse oocyte.

The high number of differential genes expression of immature oocytes revealed that an oocyte at the age of 3 weeks has a distinct transcriptome profile as compared to oocyte from adult mouse. The PCA plot illustrated a distinguished cluster of the pre-pubertal oocytes from the 9-week age group, indicating substantial differences in gene expression before and after puberty. More importantly, PCA has allowed the assessment of the global quality of the RNA-seq dataset, as it enables one to examine the reproducibility among replicates (eight samples from each age) as well as the possible batch effects of the sequencing lane.

Single cell-RNA- sequencing revealed novel regulatory molecules associated with the maturation of oocytes at the pre-puberty stage. Canonical pathways analyses with IPA identified pathways that were significantly associated with the maturation of pre-pubertal oocytes. The top ten significant signalling pathways were mainly involved in: Rho Family GTPases pathway, Integrin Linked Kinase (ILK) pathways, Pregnane X receptor /Retinoid

X receptor (PXP/RXR) pathway, Xenobiotic Metabolism Signaling pathway, RANK Signaling in Osteoclasts pathway, Type II Diabetes Mellitus Signaling pathway, LPS/IL-1 Mediated Inhibition of RXR Function pathway, Germ Cell-Sertoli Cell Junction Signaling pathway, GP6 signalling pathway, Interlukin-8 signalling pathway. Most of these are relevant to oocyte development (Figure 5.8, 5.9).

Interestingly, most of the canonical pathways are closely related (overlapping pathways in Figure 5.10). Many of the differentially expressed genes are present in multiple pathways, which also relate to network processes (Table 5.1). Some pathways have actually included similar overlapped genes. For instance, the Rho Family by GTPases signalling pathway includes all the genes that associated with RANK Signaling in Osteoclasts pathways as well as its own genes. Therefore, the pathways that are biologically relevant to ovarian development were selected for analysis.

The most activated canonical pathway was signalling by Rho family GTPases family, which includes 19 transcripts (FC>2 or <2) as shown in Table 5.2. The genes present in this pathway dominate the data with the highest expression (together with ILK signaling pathway) and are involved mainly with cytoskeleton organisation.

The Rho of GTPases family are well-known regulators of the actin cytoskeleton, which is involved in many important processes, such as cell adhesion, regulation of cell shape and intracellular processes (177). The Ras homolog family (Rhou), is a member of the GTPases family and plays multiple roles in many aspects of cell function, was highly expressed in pre-pubertal oocyte (Table 5.2) Rhou was recently revealed to regulates cytoskeleton dynamics by affecting actin filament assembling and spindle formatting (177). Interestingly,

Rhou was recently reported as upstream for the Mitogen activated protein kinase10 (Mapk10), a cytoplasmic factor that has a fundamental role in initiating meiosis resumption. Mapk10 was diminished two fold in pre-pubertal oocyte relative to the 9 weeks oocyte (Mapk10, FC= -2). Mapk10 is a serine/threonine-protein kinase involved in various processes such as cell proliferation, differentiation and programmed cell death. Cytoplasmic maturation is regulated by kinase involved in the initiation and progression of meiosis. Since Mapk10 plays a crucial role in the resumption of oocyte maturation (178). The reduced expression level of Mapk10 in a pre-pubertal oocyte indicates that activity of Mapk10 has not been triggered yet, negatively influencing the initiation of meiosis resumption. The lower Mapk10 is consistent with previous observations that demonstrated the lower activity of Mapk10 in an immature calf oocyte, which correlated with the incomplete cytoplasmic maturation that led to poor results (37).

Interestingly, Mapk10 is found to be an upstream regulator for myosin light kinase (Mylk) that phosphorylate Myosin light chain 9 (Myl9)(179). Both genes were highly expressed with FC= 9.16, 4.2, respectively. Both enables myosin-II organisation into bipolar thick filaments that stabilise spindle. Additionally, Vimentin (Vim) with FC more than 6.04, has been identified in different species, but not detected in mice, and is reported to be required for successful nuclear reprogramming in porcine oocytes (160,180)

VIM is one type of intermediate filament distributed throughout the cytoplasm and functionally connected to the nuclear matrix. Vim has a vital function in the maintenance of cellular integrity and acting as a network that binds and regulates the activity of several effector proteins. Together with the higher expression of Acta, Alpha-actin-2 (Acta2, FC= 6.63) indicates that these cytoskeleton molecules accumulate in large quantities in the

cytoplasm of pre-pubertal oocyte to prepare the oocyte for meiosis because transition in the meiotic cycle has to be coordinated with the function of the spindle and cortical cytoskeleton. The increased expression of Septin (Sep9), which regulates spindle assembly and microtubule, indicating that Sep9 is key player in cytoskeletion remodeling that taking place during this stage of oocyte development. This supports the fact that cytoplasmic maturation involves the organelles reorganisation during oocyte maturation. Collectively, the actomyosin cytoskeleton undergoes significant changes to prepare oocyte for meiosis I. The actin-cytoskeleton family seemed to crucially participate in supporting the onset of maturation during week 3 of oocyte age. This is based on the fact that oocytes rely on the storage of essential factors that oocytes accumulate during the period before GVBD (161). At this point, it seems that oocytes at this age have not reached enough cytoplasmic maturation required for gaining competence.

Integrin-linked Kinase pathway was the the second canonical pathway (Table 5.3). Integrins are adhesion receptors that allow cells to interact to micro-environmental signals encoded by the extracellular matrix. Integrin pathways have been reported as contributing during oocyte fertilisation at adhesion and are mainly involved in cell cycle regulation (181,182). Integrin-mediated signals allow cells to progress from G1 to S phase, and are important as temporal growth factor signals. The genes included eight that also act across the Rho family of GTPases pathway (Acta2, Mapk10, Myl9, Rhou, im, Tlr9, Irs1, Ppp2rib), with large overlap between the two pathways indicating that these genes have intersected functions that govern mechanisms in the cytoplasm of pre-pubertal oocyte. This data reveals the strong communication between Rho family of GTPases and the Integrin signaling pathway, of

which both support oocyte communication with the extracellular matrix and allow receiving growth factors from the microenvironment.

A toll-like receptor (Tlr9), which is shown to be participated with most of significant pathways, was reported as expressed in cumulus oocyte complex (COC) expansion, acting as a receptor during ovulation (183). A recent study demonstrated that TLR is not specific for cumulus cells, but it is also expressed during embryonic development (184). It could therefore be suggested that Tlr9 has a novel regulatory role in mediating the interaction with the extracellular matrix (ECM) during the maturation at the pre-pubertal stage. Also, since TLR is one of the most important elements required for the induction of insulin resistance in different organs, it is therefore expected that the Insulin growth substrate (IGFS) has been affected by Tlr9, which seems to have a novel role in oocyte maturation.

The integrin receptor allows the transduction of a signal via cytoskeletal proteins that involve Actin, Tubulin, Lamin B, all encoded by genes that have been expressed at high levels in pre-pubertal oocytes. This confirms the significant role of ILK signaling pathways in regulating the mechanisms that potentially associated with the development of competence.

Ingenuity Pathway Analysis has highlighted that PXR/RXR activation as a novel pathway seems to be implicated in the transcriptome dynamics of the pre-pubertal oocyte. Retinoic Acid (RA), the active metabolite of vitamin A, mediates many physiological functions, including embryogenesis and reproduction (185). Retinoic acid acts through RA receptors (RARs) and retinoid X receptors (RXRs), where the later was upregulated two fold in pre-

pubertal oocytes (Table 5.4). The data additionally demonstrated that Aldehyde dehydrogenase (ALdh1a1), which is the principle enzyme in in the final step of RA synthesis, was highly expressed along with the RXR receptor (Table 5.4). A recent study detected the expression of RA in early foetal ovarian development, and it is suggested that germ cells differentiation is induced by RA to express the pre-meiotic marker, Steroidgenic acute regulatory protein (Star), indicating that meiosis was triggered by Aldh1a1 (186). This is evidenced by the study that showed that germ cells in foetal- deficient vitamin A were unable to initiate meiosis (187). It is therefore clear that activation of the RXR pathways drives the initiation of meiosis, which indicates that RA is necessary for the support of oocytes to gain the competence.

Since Aldh1 induces the expression of Star,This may also indicate the important role of RA in mediating other vital functions, such as steroidogenesis. The critical observation that RA influenced Star gene, provides en explanation for the activation of Cholesterol side-chain cleavage enzyme (Cyp11a1) in pre-pubertal oocyte, an enzyme involved in steroid hormone metabolism, which is part of lipid metabolism

Moreover, the level of Cholesterol 7 -hydroxylase (Cyp7A1) was diminished. Cyp1171 has been previously observed be repressed by RA in hepatocyte. Cyp1171 is an enzyme which plays a key role in maintaining lipid metabolism and cholesterol synthesis that are important for the formation of membranes required for repeated cell division to form an embryo. Other molecules associated with fatty acid metabolism, such as Carnitine (Cpt1a), shuttle of the long chain fatty acid oxidation mediated in mitochondrial matrix and Stearoyl-CoA desaturase (Scd1), are rate-limiting enzymes in the biosynthesis of fatty acids and are involved in the regulation of reproduction by maintaining (influencing) energy homeostasis (188). Both were significantly upregulated in pre-pubertal oocytes. The higher expression

of these Cpta1 and Scd1 indicate that oocytes at this stage contain a large lipid reserve for energy production that is required for subsequent fertilisation. Scd1 expression was found to be regulated by retinoic acid (189). This finding demonstrates the RA serve as critical regulators in a wide range of physiological processes which include lipid metabolism and steroidogenesis.

In line with previous findings, the results showed that PXR/RXR pathway is regulatory pathway that have impacted the activation in wide range of cellular functions and modulate oocyte metabolism.

Ingenuity Pathways Analysis was used to predict the upstream regulatory molecules and associated mechanisms underlying the causes of the gene expression pattern. An upstream regulator is an analytic tool that provides biological insight into the observed expression changes of the pre-pubertal oocyte. It identified the TGFβ signaling pathway as the upstream regulator of most of the differential genes in the dataset. TGFβ family is very well known for its central role in follicle development by mediating signal transduction of the essential growth factors required for oocyte growth. IPA enabled the prediction of the direction of TGFβ (Figure 5.11), suggesting that the TGF beta family is a potent indirect regulator of the meiotic competence of pre-pubertal oocytes.

To further confirm the molecular functions of the differential genes in our data, network analysis agrees with the function of the canonical pathways. With regard to organism development and reproductive system network was identified as consistent with Rho family of GTPases and ILK signaling, which were both implicated in oocyte development. This network includes many genes associated with steroidogenesis and ovarian development.

However, it may be more advantageous to focus on the canonical pathways, rather than focusing on a single gene as the aim of this study is to gain insight into what regulates developmental competence during pre-puberty. Apart from this, the high interconnectivity between genes and correlated genes in other pathways may imply the biological significance of these genes.

The study demonstrated that a high proportion of genes expressed in a 3-week-old oocyte decreased in abundance in a 9-week-old oocyte. The decreased level of abundance in mature oocytes is consistent with global reduction of total RNA while the oocyte proceeds to maturation. Synthesis and storage of maternal transcripts prior to the onset of transcriptional silencing, at the time of GVBD, are essential for the oocyte to reach developmental competence.

The activated pathways represented suggest that a pre-pubertal oocyte recovered from pre-ovulatory follicle is still undergoing a series of changes while approaching puberty. More importantly, it can be claimed that signalling pathways were manipulated as a response to the PMSG used to stimulate ovaries. Hormonal stimulation was used, aiming to synchronise follicle development. A previous study compared blastocyst formation rate from primed and non-primed pre-pubertal goats and showed that gonadotropins' treatment has not improved the competency of the resulting embryo from the primed goats but has only increased the number of oocytes (190). In addition, some studies have claimed that the exogenous hormone (ecG) could have increased the level of some metabolites in immature porcine oocytes. This assertion can be ruled out as the transcriptome data revealed a profound alteration of these gene levels in the pre-pubertal oocytes as compared with mature oocytes,

despite the use of same dose of PMSG. This confirmed these global changes are a consequence of alteration in transcriptional activity of the immature oocyte, independent of external hormonal stimulation.

Single cell-RNA analysis has elucidated the molecular mechanisms that govern gradual cytoplasmic maturation. The sc-RNA-seq approach used in this study enabled the identification of novel pathways that were clearly activated during the period shortly before maturity. Key genes were also identified as having major roles in equipping the oocyte with the essential functions necessary to develop competence. An oocyte actually is not quiescent as at this stage. On the contrary, GV transcriptome is undergoing dynamic changes that take place during oocyte growth, which is potentially associated with the acquisition of competence.

5.5. Conclusion

Single cell-RNA sequencing analysis was successfully performed in a single mouse oocyte. The use of high throughput sequencing of a single oocyte, combined with bioinformatics, provided insights into the regulatory signalling pathways that contributed to the developmental competence of pre-pubertal oocytes

Pre-pubertal oocytes seem to have different metabolisms, which are potentially associated with the acquisition of developmental competence. Identifying oocytes with the greatest potential for producing viable embryo would be of great benefit to pre-pubertal patients who undergo chemotherapy where the *in vitro* maturation of the pre-pubertal oocyte is a promising treatment.

.

Chapter 6

Transcriptome Analysis of Aged Oocyte

6.1. Introduction

It is well recognized that female fertility declines with advanced age. Ovarian function deteriorates with ageing as the ovary exhibits a faster rate of ageing than other organs (191). A gradual depletion in the number of ovarian follicles coincides with diminished quality of oocytes. Poor oocyte quality is principally attributed to reduced developmental competence as the aged oocytes have less capacity to sustain fertilisation and continue development (192). Therefore, oocyte quality determines the embryo's developmental potential and the success of *in vitro fertilisation* (IVF) treatment (41).

Indeed, successful completion of fertilisation is deeply rooted in two components: nuclear and cytoplasmic maturation (155). Most studies on ageing have emphasised the importance of chromosomal abnormalities, recombination errors and spindle assembly as leading causes of compromised oocyte competence (116,193–195). While these issues address nuclear maturation in aged oocytes that can be in a part assessed and visualized by the extrusion of the polar body, unfortunately, poor oocyte quality is intrinsically related to cytoplasmic maturation (155). Events related to oocyte cytoplasmic maturation are less well defined than those of nuclear maturation. Due to the fact that aged oocytes have a normal morphological appearance indistinguishable from young oocytes, fertilisation failure and cell division arrest in developing embryos from aged oocytes remains idiopathic, which is most likely related to compromised cytoplasmic factors.

Previous ageing studies have conducted transcriptome analysis on pooled oocytes, which are likely to be heterogeneous in terms of competence (52,53). The oocytes appear to vary in their ability to develop into a healthy embryo, thus, analysing them as a group masks

individual variability between oocytes. Aside from this, previous studies were conducted on oocyte at metaphase II stage, which is an oocyte's fully mature stage (196). It is known that the transcriptome undergoes changes during oocyte maturation, beside that *in vitro* culture period may have altered gene expression (197). Therefore, a clear understanding of when and how age-related molecular alterations arise in oocyte senescence remains elusive.

Recent developments in single cell transcriptomics now allow the precise, deep coverage of the transcriptome, which enhances the ability to explore the regulatory mechanism of oocyte ageing more deeply. The single oocyte transcriptome has not previously been studied using a high throughput technique, which involves the entire transcriptional activity being examined on a single oocyte. This chapter demonstrates the first use of single cell-RNA sequencing (sc-RNA-seq) to investigate the transcriptional activity of a single aged oocyte at germinal vesicle stage (GV), providing novel insight into the molecular pathways implicated in oocyte senescence. Here, we use a one-year-old mouse as a model of ageing that resembles pre-menopausal women.

The aim of this chapter is: To analyse the sc-RNA-seq data of a single aged oocyte, isolated from a one-year-old mouse and compare it with the transcriptomic profile of a young adult mouse at nine weeks old, aiming to uncover new molecular pathways that cause impaired quality of aged oocytes.

6.2. Method

To get a better understanding of the age associated molecular changes, Ingenuity Principle Analysis (IPA) was used to identify canonical pathways and the top regulatory network involved in one-year oocyte using an oocyte from a nine-week-old mouse as a baseline in conducting this comparison. Differential expressed genes were imported into IPA with their fold changes and *P* value. Metacore was used to further analyses the molecular functions of the genes.

6.3. Results

The analysis of the one-year-old oocyte transcriptomics are divided into four sections:

A. Sequencing quality: The sequencing data of the one-year-old oocyte were processed with those of the nine- and three-week-old oocytes, as discussed in Chapter 5, section 5.3
B. Principle component analysis (PCA)
C. Differential gene expression;
D. Functional analysis.

6.3.1. Principle Component Analysis

Principle component analysis plotting revealed that oocytes from one-year-old mice (green squares in Figure 6.1) exhibit a distinguished cluster from the oocytes taken from nine-week-old mice (yellow circles in Figure 6.1). Principle Component Analysis demonstrated that the

transcription patterns of aged oocytes are affected by advancing age. Interestingly, the distance between the object sample clusters of the nine-week-old and one-year-old oocytes are not separated as completely as they are in comparisons of pre-pubertal oocytes and mature oocytes (Chapter 5, Figure 5.4 and Figure 5.5). This could perhaps be expected, as oocytes from both young adult and older adult samples should not exhibit the level of variation in gene expression as seen between pre-pubertal and adult oocytes. It is notable, however, that although aged oocytes have spread within the same direction in the PCA, there was a heterogeneity seen within this group in the level of gene expression between individual oocytes (aged oocyte samples are more scattered). This may reflect individual differences in the global gene expression distinguished by sc-RNA-seq analysis. This individual variability of aged oocytes may support the conclusion that each oocyte, even those from the same ovary, have their own binary fate.

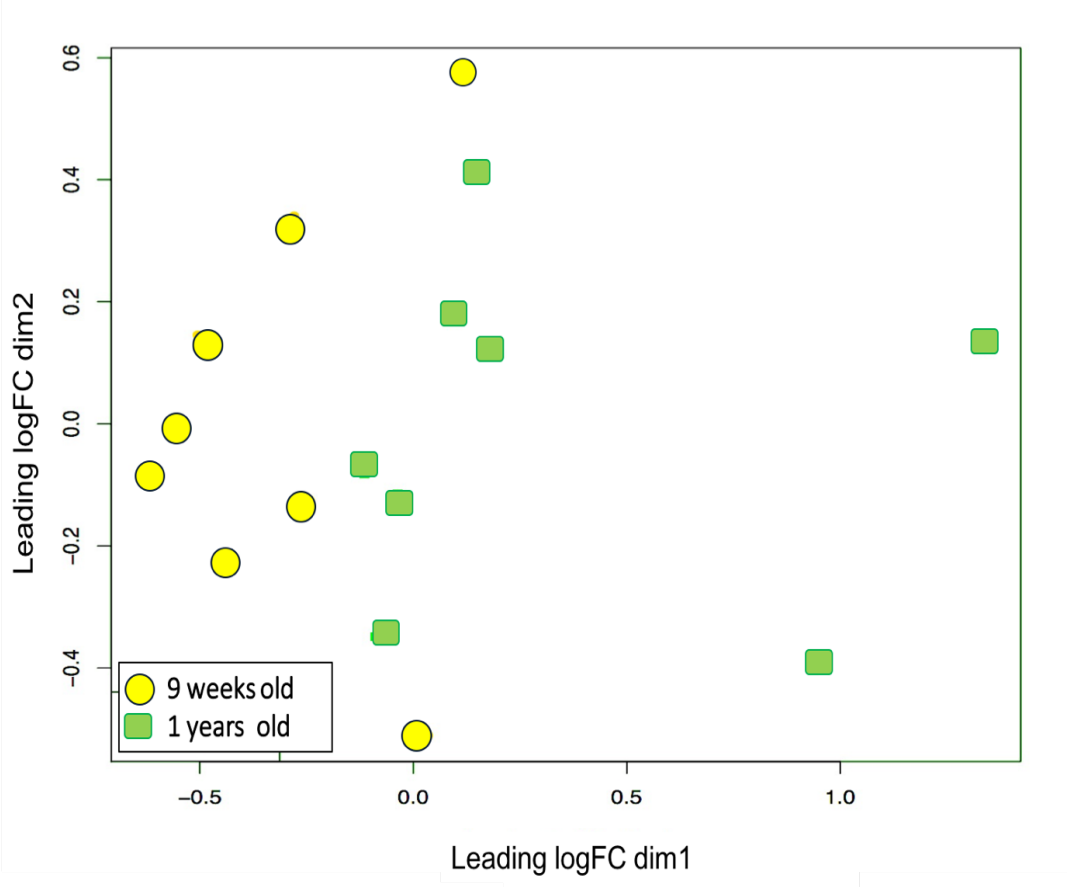


Figure 6.1 Principle component analysis of oocytes at one year of age (aged) versus nine weeks of age (young adult). Oocytes from the aged mice cluster fairly separately from the oocytes of the young adult mice.

6.3.2. Differential Gene Expression

The scatter plot of the comparison between the one-year-old oocyte and the nine-week-old oocyte depicts differentially expressed genes in red (Figure 6.2, adjusted *P* value<0.05). sc-RNA-seq detected a total of 1,642 genes found to be differentially expressed in the one-year-old oocyte compared to the nine-week-old oocyte. The pronounced difference in gene expression indicates an aberrant transcriptome profile of the aged oocytes, demonstrating a major alteration caused by age. The range of the fold change is quite broad, showing many genes with a log2-fold change of up to five and several genes with a fold change of greater than five. Negative fold change values represent downregulated genes.

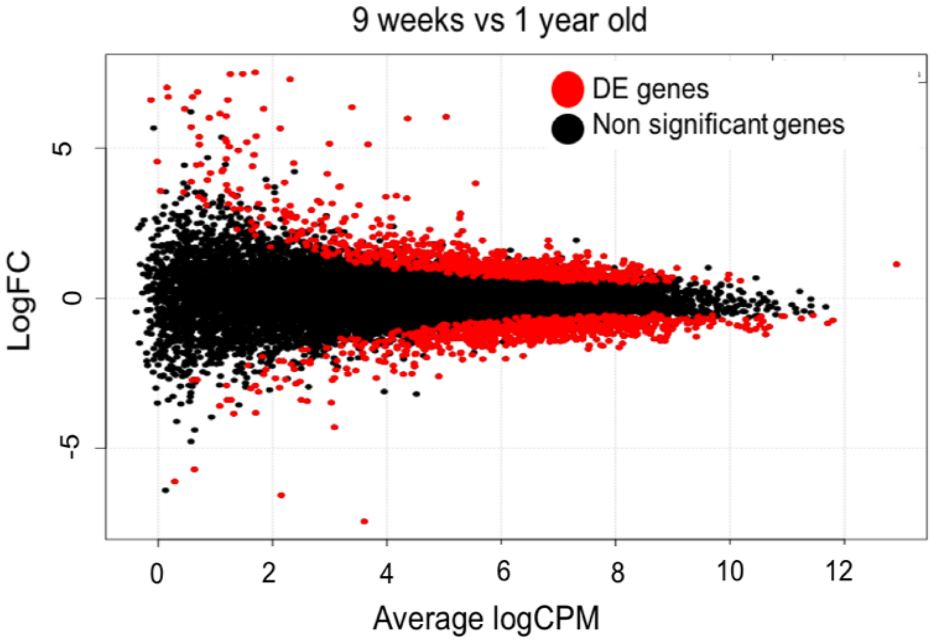


Figure 6.2 Scatter plot of the differentially expressed genes in 1-year-old oocyte compared to 9 weeks' oocyte. A scatter plot illustrates the average counts (cpm) versus Log2 -fold changes for genes generated by DEseq. Genes were retained if they expressed a threshold of 1 count-per-million (cpm) for any given sample.

6.3.3. Functional Analysis

6.6.3.1. Canonical Pathways

Core analysis was carried out using Ingenuity Pathways Analysis (IPA), which revealed pathways. Significant canonical pathways with their log (-*P* value) are shown in Table 6.1. Activated metabolic pathways in the aged oocyte indicate that the oocyte transcriptome is undergoing global changes, which lead to a different metabolism and gene expression pattern. These data represent a distinct transcriptional activity that is on-going at GV stage confirming that aging-related molecular changes arise before maturation. In turn, altered cytoplasmic maturation resulted in the declined developmental competence of these oocytes in comparison with the adult oocytes at nine weeks of age (Figure 6.3).

Additionally, a stacked bar chart was also generated to demonstrate the upregulated genes and downregulated genes in each pathway, along with the ratio (orange line) between the number of genes in a given pathway that met the cut-off criteria and the total number of molecules involved in each pathway in the reference gene set (Figure 6.4).

To give a clearer picture of the upregulated or downregulated genes in a specific canonical pathway, a list of differentially expressed genes and the top canonical pathways are provided in Table 6.1. Of the 1642 differentially expressed genes, a total of 179 with a fold changes of more or less than two is listed in Table 6.2. The top significant signalling pathways including the differential genes and their fold change are detailed as follow: retinoic acid receptors activation (RAR) pathway (Table 6.3), mitochondrial L-carnitine shuttle pathway (Table 6.4), biosynthesis, type II diabetes mellitus signalling pathway

(Table 6.5), sertoli cell-sertoli cell junction signalling pathway (Table 6.6),signalling by rho family GTPases pathway (Table 6.7), GP6 signalling pathway (Table 6.8).

Table 6.1: The top 20 enriched signalling pathways in one-year-old oocytes compared with adult young oocyte with – log (P value)>1.8. The table demonstrates the the transcripts associated with each pathway.

Ingenuity Canonical Pathways	-log (pvalue)	Ratio	Molecules
RAR Activation	4.16	0.0421	NCOR2,DHRS3,ADH1C,MAPK10,RBP1,PRKCB, MAPKAPK2,RARB
Mechanisms of Viral Exit from Host Cells	3.68	0.0976	LMNB1,CHMP4B,PRKCB,ACTA2
Mitochondrial L-carnitine Shuttle Pathway	3.64	0.176	SLC27A2,CPT1A,ACSBG1
Sertoli Cell-Sertoli Cell Junction Signaling	2.74	0.0337	MAPK10,MAP3K10,ACTA2,GSK3A,SYMPK,EPN1
Retinoate Biosynthesis I	2.74	0.0882	DHRS3,ADH1C,RBP1
Type II Diabetes Mellitus Signaling	2.64	0.0391	SLC27A2,MAPK10,CEBPB,ACSBG1,PRKCB
Signaling by Rho Family GTPases	2.64	0.0278	VIM,MAPK10,SEPT9,MAP3K10,ACTA2,MYL9,CDH8
GP6 Signaling Pathway	2.55	0.0373	COL3A1,PRKCB,COL1A2,GSK3A,COL1A1
ILK Signaling	2.52	0.0305	VIM,DSP,MAPK10,ACTA2,GSK3A,MYL9
Intrinsic Prothrombin Activation Pathway	2.47	0.0714	COL3A1,COL1A2,COL1A1
Fatty Acid Activation	2.42	0.154	SLC27A2,ACSBG1
γ-linolenate Biosynthesis II (Animals)	2.19	0.118	SLC27A2,ACSBG1
Melatonin Degradation III	2.14	1	MPO
Dermatan Sulfate Biosynthesis	2.06	0.0508	NDST2,UST,CHPF
The Visual Cycle	2.05	0.1	DHRS3,RBP1
Hepatic Fibrosis / Hepatic Stellate Cell Activation	1.95	0.0267	COL3A1,COL1A2,ACTA2,MYL9,COL1A1
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	1.88	0.0258	PPP1R14A,MAPK10,MPO,PRKCB,MAP3K10
IL-6 Signaling	1.87	0.0312	MAPK10,CEBPB,COL1A1,MAPKAPK2

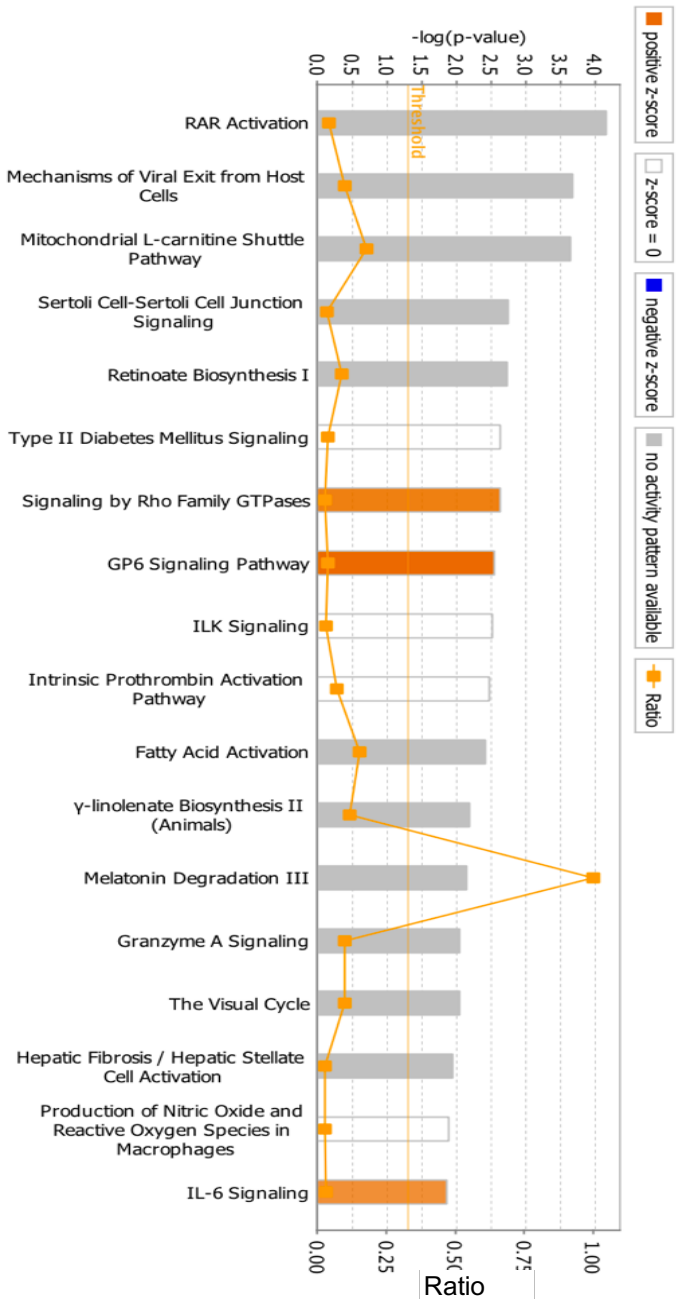


Figure 6.3 Enriched significant pathways in an aged oocyte, as compared to an adult young oocyte. The bar colours indicate prediction of the Z score, where orange represents activation and grey indicates that inhibition is implicated.

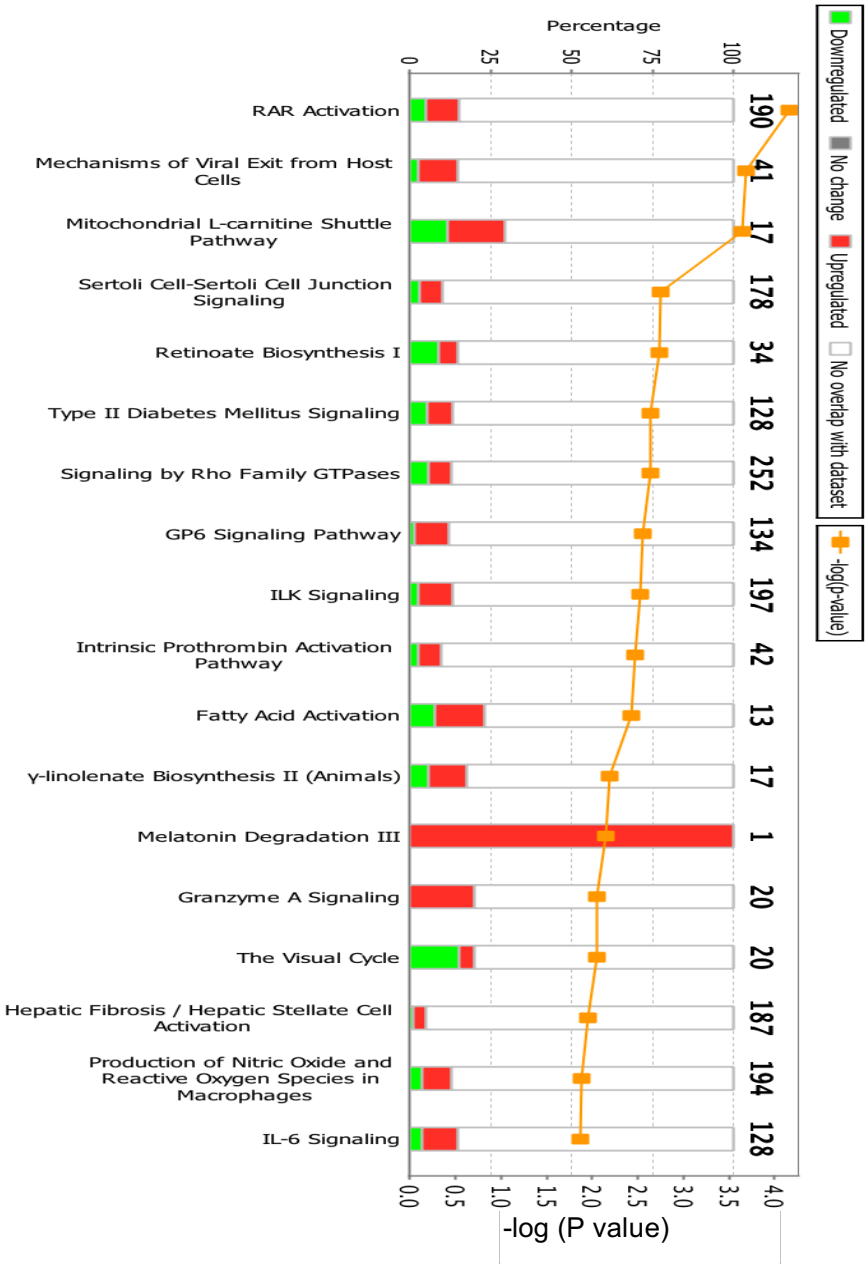


Figure 6.4 A stacked bar chart of the top enriched canonical pathways in an aged oocyte compared to a young adult oocyte. The upregulated genes (red) and downregulated genes in each pathway are demonstrated, along with the ratio (orange line).

Table 6.2: List of differential expressed genes in a aged oocyte (one-year-old) relative to young oocyte (9 weeks old), generated by IPA

Gene ID	Entrez Gene Name	Log Ratio	FDR
1110038B12Rik	RIKEN cDNA 1110038B12 gene	2.706	0.0171
2510002D24Rik	RIKEN cDNA 2510002D24 gene	5.397	0.017
2900060B14Rik	RIKEN cDNA 2900060B14 gene	2.099	0.0478
4932416H05Rik	RIKEN cDNA 4932416H05 gene	3.287	0.021
4933405O20Rik	RIKEN cDNA 4933405O20 gene	-5.699	0.0323
Acsbg1	acyl-CoA synthetase bubblegum family member 1	3.706	0.0103
Acta2	actin, alpha 2, smooth muscle, aorta	6.361	0.000782
Adh1	alcohol dehydrogenase 1C (class I), gamma polypeptide	4.497	0.0277
Ankrd52	ankyrin repeat domain 52	2.226	0.0191
Anp32a	acidic (leucine-rich) nuclear phosphoprotein 32 family, member A	2.245	0.00249
Armcx2	armadillo repeat containing, X-linked 2	7.469	0.00003
Btnl1	butyrophilin-like 1	-2.909	0.00327
2310022B05Rik	chromosome 1 open reading frame 198	2.141	0.000501
E330014E10Rik	expressed sequence C87414	-2.714	0.0499
Cald1	caldesmon 1	5.146	0.0014
Camk2n2	calcium/calmodulin dependent protein kinase II inhibitor 2	4.441	0.00854
Casc3	cancer susceptibility 3	2.266	0.01
Ccny	cyclin Y	2.066	0.0019
Cdh8	cadherin 8	-4.296	0.000461
Cebpb	CCAAT/enhancer binding protein beta	2.454	0.00127
Cenpb	centromere protein B	3.375	0.00328
Chmp4b	charged multivesicular body protein 4B	2.717	0.0421
Chpf	chondroitin polymerizing factor	2.129	0.00111
Chst5	carbohydrate sulfotransferase 6	-2.358	0.01
Col1a1	collagen type I alpha 1 chain	5.7	0.0219
Col1a2	collagen type I alpha 2 chain	5.197	0.0151
Col3a1	collagen type III alpha 1 chain	6.87	0.00941
Cpt1a	carnitine palmitoyltransferase 1A	4.64	0.0117
Crip1	cysteine rich protein 1	3.248	0.0447
Ctsz	cathepsin Z	2.558	0.0329
Cttnbp2	cortactin binding protein 2	-3.386	0.0259
Cyp11a1	cytochrome P450 family 11 subfamily A member 1	6.038	0.000984
Dazap1	DAZ associated protein 1	2.363	0.000811
Ddx60	DExD/H-box helicase 60	-2.727	0.00812
Defb41	defensin beta 41	2.684	0.0386
Dhrs3	dehydrogenase/reductase 3	-2.043	0.00404
Dnajc27	DnaJ heat shock protein family (Hsp40) member C27	6.063	0.0000761
Dsp	Desmoplakin	2.037	0.0371
Dync1h1	dynein cytoplasmic 1 heavy chain 1	2.156	0.00354

Gene ID	Entrez Gene Name	Log Ratio	FDR
Dync1i1	dynein cytoplasmic 1 intermediate chain 1	-3.476	0.0188
Ehmt2	euchromatic histone lysine methyltransferase 2	2.674	0.0269
Eid1	EP300 interacting inhibitor of differentiation 1	3.262	0.0255
Enpp2	ectonucleotide pyrophosphatase/phosphodiesterase 2	-2.6	0.0014
Epm2aip1	EPM2A interacting protein 1	3.381	0.0331
Epn1	epsin 1	2.476	0.0196
Evc	EvC ciliary complex subunit 1	3.936	0.0406
Fbxo32	F-box protein 32	2.935	0.013
Flt3	fms related tyrosine kinase 3	6.603	0.0173
Fndc4	fibronectin type III domain containing 4	5.375	0.0227
Foxl2	forkhead box L2	2.706	0.0459
Gca	Grancalcin	-2.847	0.0353
Gm10280	predicted gene 10280	-2.082	0.0149
Gm14085	predicted gene 14085	-2.272	0.0167
Gm19276	predicted gene, 19276	-2.092	0.0469
Gm4224	predicted gene 4224	-2.115	0.0421
Gm8221	apolipoprotein L 7c pseudogene	-2.13	0.00325
Gm8817	predicted gene 8817	-7.436	0.0329
Gsk3a	glycogen synthase kinase 3 alpha	2.152	0.00948
H1fx	H1 histone family member X	2.666	0.0037
H6pd	hexose-6-phosphate dehydrogenase/glucose 1-dehydrogenase	2.814	0.00501
Hars2	histidyl-tRNA synthetase 2, mitochondrial	5.117	0.0442
Hist1h1b	histone cluster 1, H1b	4.14	0.0000877
Hist1h1e	histone cluster 1, H1e	2.712	0.0202
Hnrnpul1	heterogeneous nuclear ribonucleoprotein U like 1	2.019	0.00826
Hsd3b1	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 2	3.827	0.00513
Isl1	ISL LIM homeobox 1	-2.657	0.0191
Kcnh5	potassium voltage-gated channel subfamily H member 5	-2.339	0.0109
Kctd17	potassium channel tetramerization domain containing 17	2.26	0.00265
Kif26a	kinesin family member 26A	3.854	0.00076
Lbh	limb bud and heart development	4.46	0.00752
Lgals1	galectin 1	3.327	0.00152
Limch1	LIM and calponin homology domains 1	3.631	0.0476
Lingo2	leucine rich repeat and Ig domain containing 2	-2.027	0.00003
Lmnb1	lamin B1	2.531	0.00071
Map3k10	mitogen-activated protein kinase kinase kinase 10	2.818	0.00849

Gene ID	Entrez Gene Name	Log Ratio	FDR
Mapk10	mitogen-activated protein kinase 10	-2.178	0.0389
Mapkapk2	mitogen-activated protein kinase-activated protein kinase 2	2.194	0.031
Marcksl1	MARCKS like 1	3.154	0.0417
Mdga2	MAM domain containing glycosylphosphatidylinositol anchor 2	-3.43	0.00152
Mertk	MER proto-oncogene, tyrosine kinase	-2.113	0.00964
Mettl7a2	methyltransferase like 7A3	-2.017	0.000246
Mgp	matrix Gla protein	4.922	0.00961
Mill1	MHC I like leukocyte 1	-2.076	0.014
Mpo	Myeloperoxidase	3.08	0.0172
Mt1	metallothionein 1	2.668	0.0008
Mt2	metallothionein 2	3.175	0.0493
Myl9	myosin light chain 9	7.291	0.00388
Myocd	Myocardin	3.41	0.0396
Mypop	Myb related transcription factor, partner of profiling	2.29	0.00637
Naf1	nuclear assembly factor 1 ribonucleoprotein	2.069	0.0284
Ncor2	nuclear recept§or corepressor 2	2.536	0.00883
Ndst2	N-deacetylase and N-sulfotransferase 2	4.224	0.0496
Nrtn	Neurturin	2.98	0.021
Ntn1	netrin 1	2.279	0.0166
Ntn4	netrin 4	7.017	0.00869
Nudt17	nudix hydrolase 17	-3.006	0.0464
Olfr1164	olfactory receptor 1164	-3.118	0.0376
Olfr111	olfactory receptor family 5 subfamily V member 1	-2.727	0.0395
Olfr1104	olfactory receptor family 8 subfamily I member 2	-3.39	0.0077
Ostm1	osteopetrosis associated transmembrane protein 1	-2.379	0.0376
Otud7a	OTU deubiquitinase 7A	6.699	0.0112
Pank1	pantothenate kinase 1	7.52	0.000213
Pbxip1	PBX homeobox interacting protein 1	4.169	0.0274
Pck2	phosphoenolpyruvate carboxykinase 2, mitochondrial	6.603	0.0119
Pcolce	procollagen C-endopeptidase enhancer	4.283	0.031
Phf2	PHD finger protein 2	2.125	0.0014
Phldb3	pleckstrin homology like domain family B member 3	3.874	0.0122
Pim1	Pim-1 proto-oncogene, serine/threonine kinase	2.297	0.025
Ppp1r14a	protein phosphatase 1 regulatory inhibitor subunit 14A	5.213	0.0442
Ppp1r9b	protein phosphatase 1 regulatory subunit 9B	2.022	0.000984
Prkcb	protein kinase C beta	2.589	0.0414
Prr7	proline rich 7, synaptic	2.409	0.0028
Psg17	pregnancy specific glycoprotein 18	-2.53	0.000813

Gene ID	Entrez Gene Name	Log Ratio	FDR
Psme1	proteasome activator subunit 1	2.035	0.0423
Pycard	PYD and CARD domain containing	6.147	0.0176
R3hcc1	R3H domain and coiled-coil containing 1	2.88	0.0432
Rab24	RAB24, member RAS oncogene family	2.693	0.00748
Ralyl	RALY RNA binding protein-like	-2.243	0.00061
Rarb	retinoic acid receptor beta	-2.371	0.0103
Rasl10a	RAS like family 10 member A	4.391	0.00998
Rbms1	RNA binding motif single stranded interacting protein 1	-2.192	0.0151
Rbp1	retinol binding protein 1	5.656	0.0118
Renbp	renin binding protein	5.295	0.00513
Rmdn2	regulator of microtubule dynamics 2	-2.666	0.0277
Rnf144b	ring finger protein 144B	3.587	0.0147
Rnf149	ring finger protein 149	2.454	0.016
Runx1	runt related transcription factor 1	3.717	0.0299
S100a1	S100 calcium binding protein A1	2.957	0.0206
S100a10	S100 calcium binding protein A10	3.098	0.00824
S100a6	S100 calcium binding protein A6	6.306	0.00228
Sbk1	SH3 domain binding kinase 1	2.108	0.000213
Scrib	scribbled planar cell polarity protein	3.151	0.0327
Sdsl	serine dehydratase like	7.462	0.00276
Sema3e	semaphorin 3E	-2.537	0.0183
Sept9	septin 9	5.046	0.00383
Serpinb3a	serpin family B member 4	-2.657	0.00286
LOC100038947	signal regulatory protein beta 1	-3.386	0.0499
Slc27a2	solute carrier family 27 member 2	3.575	0.0308
Slc4a10	solute carrier family 4 member 10	-2.095	0.0392
Snora81	small nucleolar RNA, H/ACA box 81	2.971	0.0181
Sox4	SRY-box 4	3.723	0.0174
Sparc	secreted protein acidic and cysteine rich	2.622	0.0187
St6galnac3	ST6 N-acetylgalactosaminide alpha-2,6-sialyltransferase 3	-2.412	0.0014
Sympk	Symplekin	2.397	0.00193
Synpo2l	synaptopodin 2 like	4.547	0.0484
Taf10	TATA-box binding protein associated factor 10	2.291	0.00000409
Taf1c	TATA-box binding protein associated factor, RNA polymerase I subunit C	6.698	0.00965
Tcf19	transcription factor 19	6.008	0.0171
Tmc3	transmembrane channel like 3	3.385	0.0433
Tmeff2	transmembrane protein with EGF like and two follistatin like domains 2	-6.559	0.00263
Tmem176a	transmembrane protein 176A	3.779	0.0326
Tmem69	transmembrane protein 69	-6.102	0.0265
Tmppe	transmembrane protein with metallophosphoesterase domain	2.404	0.0254
Tmsb4x	thymosin, beta 4, X chromosome	5.985	0.000777
Trmt61a	tRNA methyltransferase 61A	3.467	0.0165

Gene ID	Entrez Gene Name	Log Ratio	FDR
Trpc5	transient receptor potential cation channel subfamily C member 5	-2.975	0.0461
Tubgcp4	tubulin gamma complex associated protein 4	-3.821	0.00824
Uap1l1	UDP-N-acetylglucosamine pyrophosphorylase 1 like 1	3.135	0.0152
Ubqln2	ubiquilin 2	2.104	0.00003
Ust	uronyl 2-sulfotransferase	2.819	0.00381
Vars	valyl-tRNA synthetase	2.753	0.00558
Vim	Vimentin	5.123	0.00388
Vmn2r26	vomeronasal 2, receptor 26	-3.846	0.0014
Zbed6	zinc finger BED-type containing 6	2.546	0.0488
Zcchc24	zinc finger CCHC-type containing 24	2.922	0.0114
Zfp119a	zinc finger protein 119a	-2.905	0.0441
Zfp871	zinc finger protein 871	4.772	0.00404
Zfp113	zinc finger protein 3	-2.007	0.00874
Zfp317	zinc finger protein 317	-2.703	0.0487
Zfp418	zinc finger protein 416	-2.078	0.0101
Zfp579	zinc finger protein 579	3.532	0.0462
Zfp580	zinc finger protein 580	2.135	0.0279

Table 6.3: RAR Activation *P* value= 6.98E-05

Symbol	Entrez Gene Name	Log Ratio	FDR (q-value)	Location
Mapkapk2	mitogen-activated protein kinase-activated protein kinase 2	2.194	0.031	Nucleus
Prkcb	protein kinase C beta	2.589	0.0414	Cytoplasm
Ncor2	nuclear receptor corepressor 2	2.536	0.00883	Nucleus
Rbp1	retinol binding protein 1	5.656	0.0118	Extracellular Space
Adh1	alcohol dehydrogenase 1C (class I), gamma polypeptide	4.497	0.0277	Cytoplasm
Dhrs3	dehydrogenase/reductase 3	-2.043	0.00404	Cytoplasm
Mapk10	mitogen-activated protein kinase 10	-2.178	0.0389	Cytoplasm
Rarb	retinoic acid receptor beta	-2.371	0.0103	Nucleus

Table 6.4: Mitochondrial L-carnitine Shuttle Pathway. *P* value= 2.28E-04

Symbol	Entrez Gene Name	Log Ratio	FDR (q-value)	Location
Acbg	acyl-CoA synthetase bubblegum family member 1	3.706	0.0103	Cytoplasm
Cpt1a	carnitine palmitoyltransferase 1A	4.64	0.0117	Cytoplasm
Slc27a2	solute carrier family 27 member 2	3.575	0.0308	Cytoplasm

Table 6.5: Type II Diabetes Mellitus *P* value= 2.31E-03

Symbol	Entrez Gene Name	Log Ratio	FDR (q-value)	Location
Acsbg1	acyl-CoA synthetase bubblegum family member 1	3.706	0.0103	Cytoplasm
Cebpb	CCAAT/enhancer binding protein beta	2.454	0.00127	Nucleus
Mapk10	mitogen-activated protein kinase 10	-2.178	0.0389	Cytoplasm
Prkcb	protein kinase C beta	2.589	0.0414	Cytoplasm
Slc27a2	solute carrier family 27 member 2	3.575	0.0308	Cytoplasm

Table 6.6: Sertoli Cell-Sertoli Cell Junctions Signalling *P* value= 1.81E-03

Symbol	Entrez Gene Name	Log Ratio	FDR (q-value)	Location
Acta2	actin, alpha 2, smooth muscle, aorta	6.361	0.000782	Cytoplasm
Epn1	epsin 1	2.476	0.0196	Plasma Membrane
Gsk3a	glycogen synthase kinase 3 alpha	2.152	0.00948	Nucleus
Mapk10	mitogen-activated protein kinase 10	-2.178	0.0389	Cytoplasm
Map3k10	mitogen-activated protein kinase kinase kinase 10	2.818	0.00849	Cytoplasm
Sympk	Symplekin	2.397	0.00193	Cytoplasm

Table 6.7: Signalling by Rho Family GTPases *P*-value= 2.31E-03

Symbol	Entrez Gene Name	Log Ratio	FDR (q-value)	Location
Acta2	actin, alpha 2, smooth muscle, aorta	6.361	0.000782	Cytoplasm
Cdh8	cadherin 8	-4.296	0.000461	Plasma Membrane
Map3k10	mitogen-activated protein kinase kinase kinase 10	2.818	0.00849	Cytoplasm
Mapk10	mitogen-activated protein kinase 10	-2.178	0.0389	Cytoplasm
Myl9	myosin light chain 9	7.291	0.00388	Cytoplasm
Sept9	septin 9	5.046	0.00383	Cytoplasm
Vim	vimentin	5.123	0.00388	Cytoplasm

Table 6.8: GP6 Signalling Pathway *P* value = 2.81E-03

Symbol	Entrez Gene Name	Log Ratio	FDR (q-value)	Location
Col1a1	collagen type I alpha 1 chain	5.7	0.0219	Extracellular Space
Col1a2	collagen type I alpha 2 chain	5.197	0.0151	Extracellular Space
Col3a1	collagen type III alpha 1 chain	6.87	0.00941	Extracellular Space
Gsk3a	glycogen synthase kinase 3 alpha	2.152	0.00948	Nucleus
Prkcb	protein kinase C beta	2.589	0.0414	Cytoplasm

The top 10 canonical pathways were connected into a single network showing the interaction between these pathways and the number of genes that overlapped between each of two pathways, as shown in the figure below.

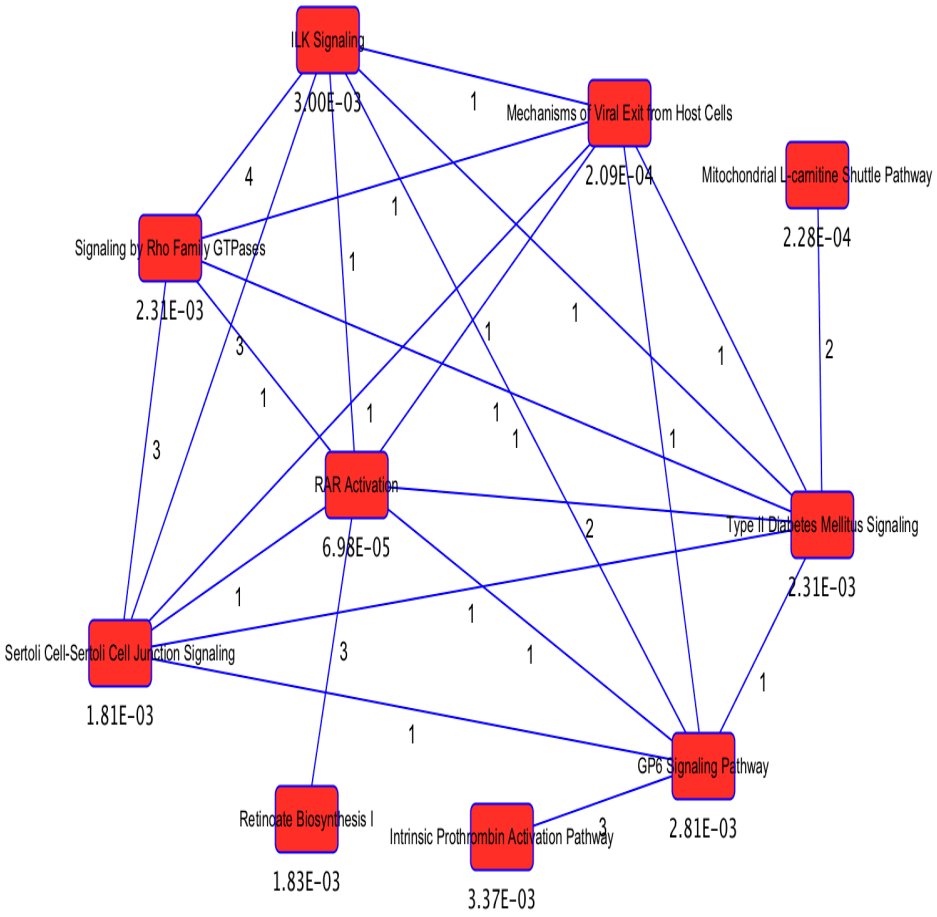


Figure 6.5 Overlapping of the top 10 canonical pathways generated by IPA. *P* values and the number of common genes between the two pathways. This represents the interconnectivity among activated signalling pathways.

6.3.2. Network Identification

Using IPA, differential expressed genes were grouped based on their molecular connectivity and cellular functions. The top four networks were identified and listed below (Red color represents up-regulated genes; green color represents down- regulated genes)

A) Embryonic Development, Organismal Development (Figure 6.6).

B) Cell Death and Survival, cellular Assembly and organization (Figure 6.7).

C) Organismal injury and and abnormalities (Figure 6.8).

D) Lipid Metabolism, small molecules chemistry, infectious disease (Figure 6.9).

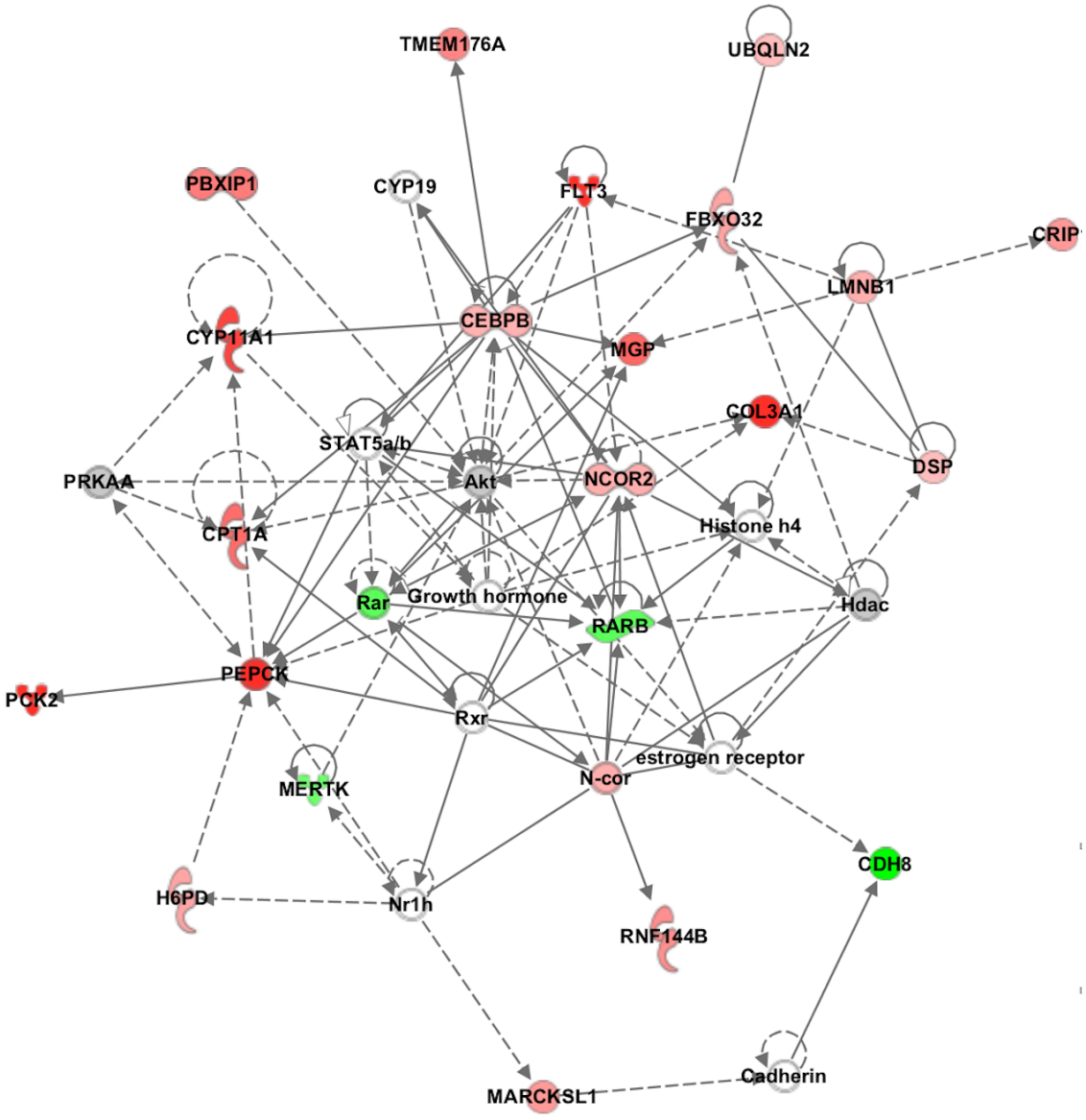


Figure 6.6: Network 1: Embryonic Development, Organismal Development. One of the top network affected in 1-year oocyte.

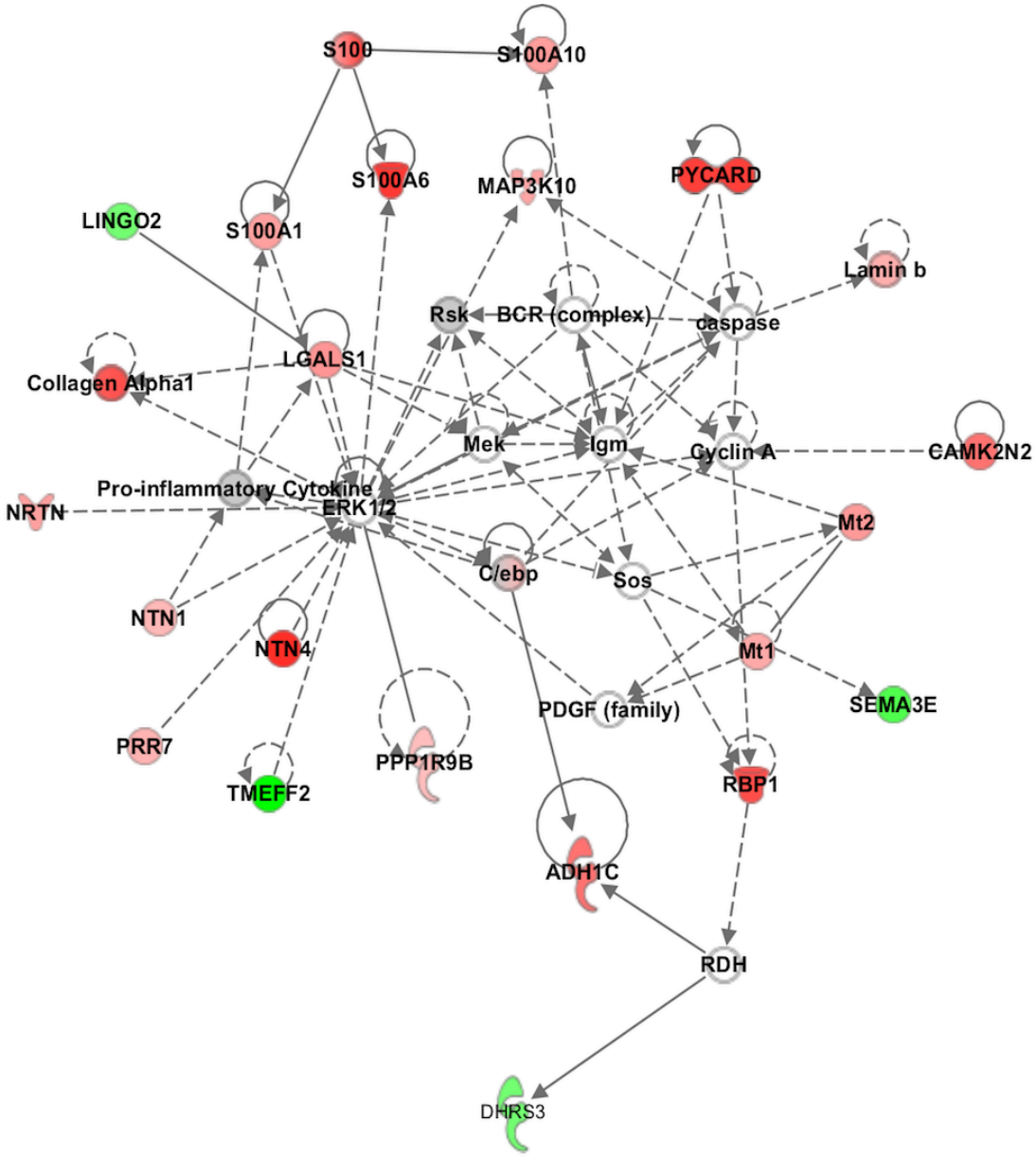


Figure 6.7 Network 2: Organismal injury and abnormalities.

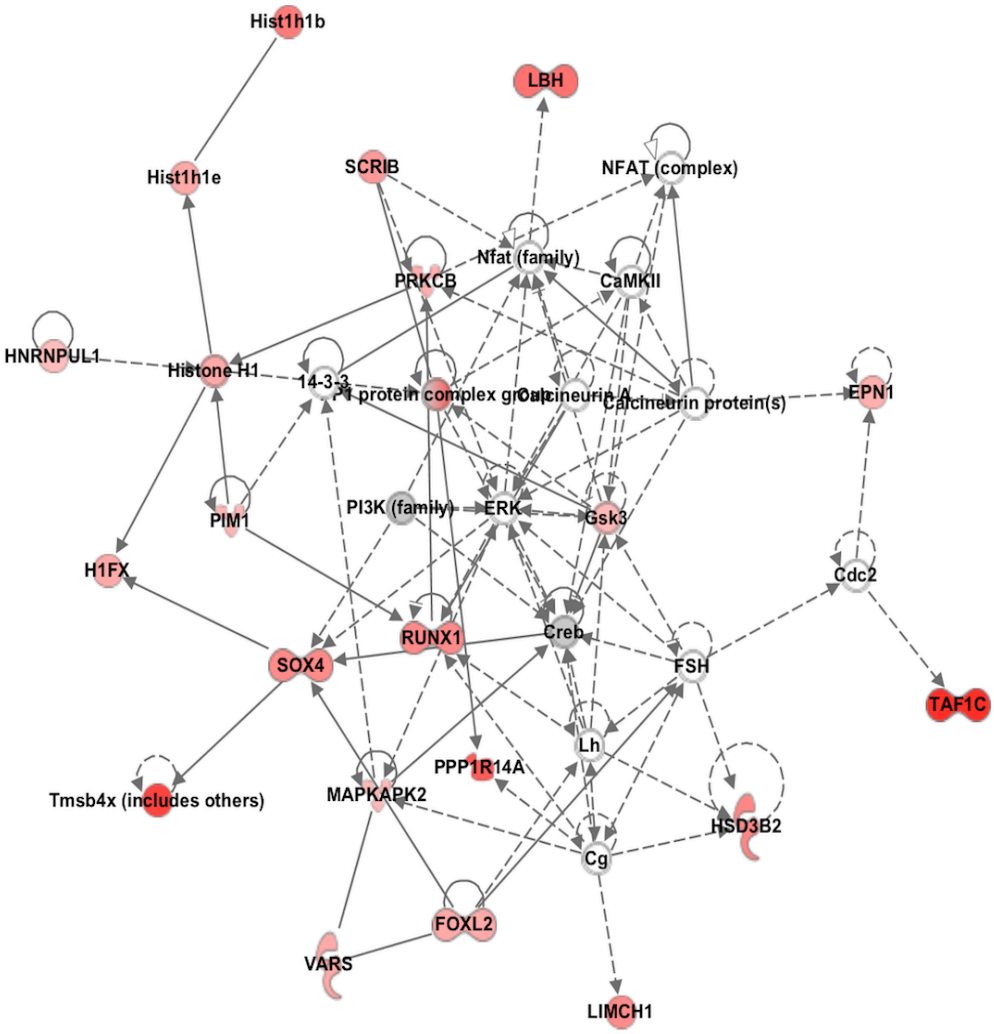


Figure 6.8 Network 3: Cell Death and Survival, Cancer.

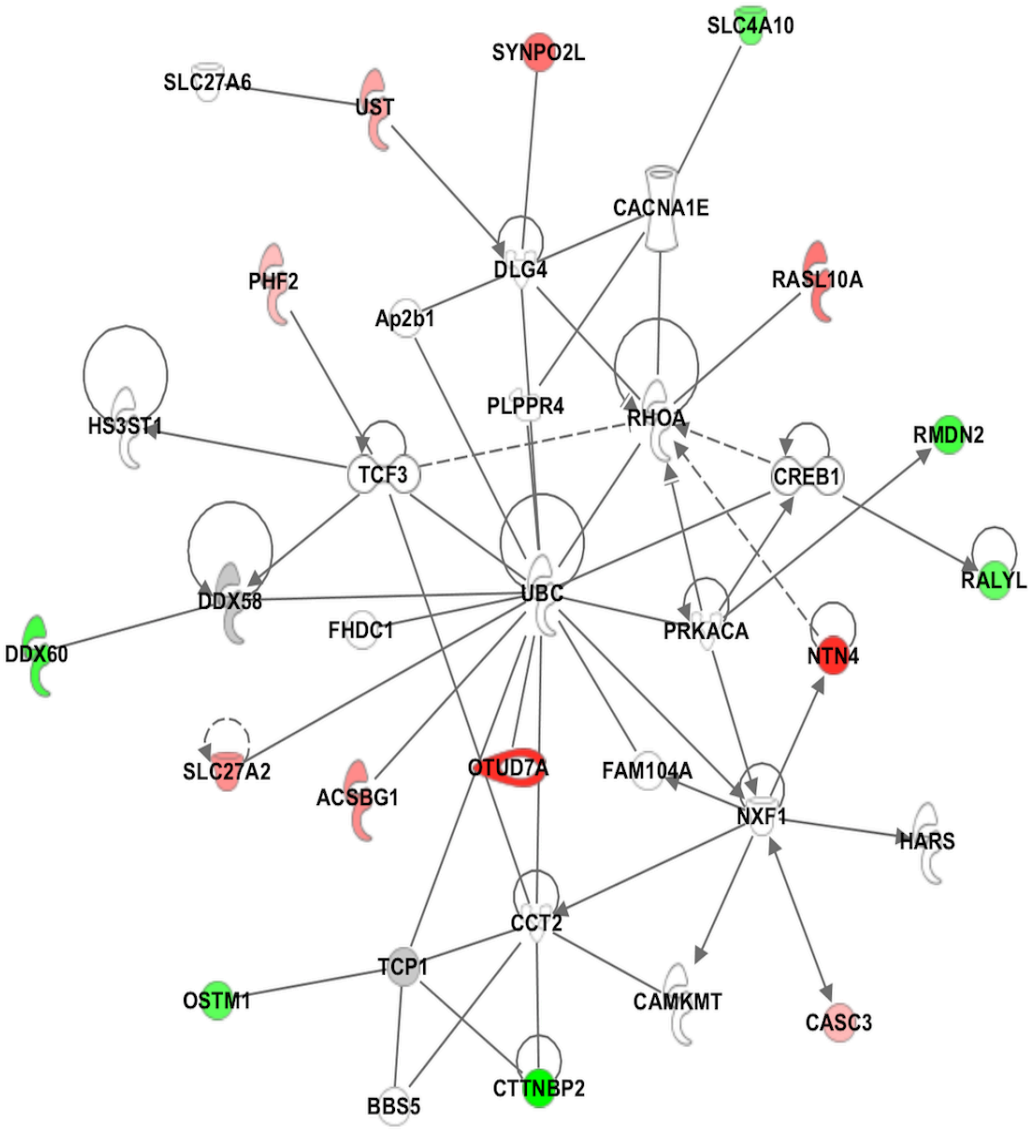


Figure 6.9 Network 4: Lipid metabolism, small molecule Biochemistry.

6.4. Discussion

Ovarian aging is believed to be dominated by a gradual decrease in the number of ovarian follicles, accompanied by diminished quality of oocytes. Characterization of changes to the transcriptome related to age has seldom been determined using high throughput technologies. This study employed sc-RNA-seq for the first time to characterize the transcriptome profile of aged oocytes, aiming to uncover novel pathways and regulatory transcripts that underlie the the compromised competence of aged oocytes.

Using one-year-old female mice as a model for ageing, sc-RNA seq detected an average of 12,500 to 13,000 expressed genes in a single oocyte, of which 1,642 genes were differentially expressed in a single aged oocyte compared to a young oocyte. Thus, sc-RNA-seq provided a deep investigation on the single cell level. The amplification strategy employed in this study has generated highly reproducible results that exhibit a high degree of accuracy and concordance between samples. The Smart seq2 approach has greater sensitivity and coverage compared with previous microarrays studies. Pan *et al.* (2008) have found 904 genes that differentially expressed between young and aged mice oocytes. Of these differential genes, only 47 changed more than 2 folds (52) Furthermore, Hamatani *et al.* (2004) conducted a microarray study on a group of mice oocytes at the metaphase II stage (MII). This study found 530 genes to be differentially expressed in aged oocytes relative to young oocytes (113) Using sc-RNA-seq has revealed much more pronounced differences, through which 1,670 differentially expressed genes were revealed. It is also important to emphasize that although microarray studies revealed age-related alteration in aged oocytes(113), these differences in fold changes for several genes do not reflect the actual

expression of transcripts in each oocyte. Because microarrays were conducted on pooled oocytes, it rather demonstrates the average of the changes that may not be well represented when analysing one oocyte.

sc-RNA seq combined with bioinformatics analysis using IPA revealed that oocyte ageing involves dysregulation of multiple interrelated biological functions. Canonical pathway analyses with IPA identified pathways that were significantly associated with the maturation of aged oocytes. The top 10 significant signalling pathways were: retinoic acid receptors activation (RAR) pathway, mitochondrial L-carnitine shuttle pathway, sertoli cell-sertoli cell junction signalling pathway, retinoid biosynthesis, type II diabetes mellitus signalling pathway, signalling by rho family GTPases pathway, GP6 signalling pathway and integrin linked kinase (ILK) pathways. Most of these pathways are relevant to oocyte development (Figures 6.3 and 6.4).

The most activated canonical pathway was RAR activation, including eight genes that are all involved in retinoid metabolism (Table 6.3). Retinoid has a vital role in cell growth and differentiation, oocyte maturation and embryo development (198,199). Retinoid activity during embryo development is reliant on the presence of two nuclear retinoic acid receptors: RAR isotypes (α, β, and γ) and RXR isotypes (α, β, and γ). Rarβ, which encodes retinoic acid receptor beta and localizes to the cytoplasm close to subnuclear compartments, was found to be decreased in one-year-old oocytes (Table 6.3). Given that Rarβ expression was detected during pre-implantation (200), it is highly likely that the decreased abundance of Rarβ leads to depletion in retinoid activity in aged oocytes, which may be related to the cleavage arrest seen in embryos from aged oocytes.

Additionally, it is important to highlight the contribution of retinoid signalling in meiosis resumption in oocytes (201,202). As the data show a decreased Mapk10 (FC= -2.18) level, which is a key factor for initiating meiosis, this finding supports a previous observation of decreased Mapk10 activity in aged oocytes (203). This suggests a relationship between retinoic acid activity and Mapk10, which may be reduced accordingly.

Therefore, retinoid levels must be continuously regulated in order to maintain normal oocyte and embryo development. Retinoid Acid (RA) synthesis is mainly controlled by two sequential steps: alcohol dehydrogenase (Adh1) or short-chain dehydrogenase/reductase (SDR)-mediated oxidation, which produce retinaldehyde (RAL). Retinaldehyde is subsequently oxidized into RA, which is mediated by an enzyme called aldehyde dehydrogenases (Aldha1) (189).

Alcohol dehydrogenase exhibits a four fold increased in aged oocytes (Table 6.3). Alcohol dehydrogenase is an essential component that acts on the excess amount of retinol by degrading it (204). Excess of all-trans retinoic acid (ATRA), which is one of the derivatives of retinoic acid, was reported to be accumulated in an Adh1-knockout mouse (205). Based on the increased Adh1 in aged oocytes, it can be proposed that Adh1 has not been metabolized due to the decreased uptake of retinol in the aged oocyte. Potentially supporting this hypothesis is that data showed a diminished level of retinaldehyde reductase (Dhrs, Table 6.3), which acts as negative feedback regulating the level of RA (206). In a Dhrs3 deficient mouse model, lack of Dhrs3 led to a 40% increase in the levels of ATRA and a 60% decrease in the levels of retinol, with the result that Dhrs3$^{-/-}$ embryos die late in gestation and display severe developmental defects and malformation (207). In aged oocytes, the decreased level of Dhrs appears to adversely contribute to unbalanced retinoid

concentration. The highly critical role of retinol and its precursors in post-implantation embryo development constitutes strong evidence that oocytes are extremely sensitive to alteration in retinol concentration.

Retinol-binding protein 4 (Rbp4), a fat-derived adipokine, is a carrier protein responsible for retinol transport to tissues (200). Retinol-binding protein 4 is elevated in the serum of polycystic ovaries, and is positively correlated with type II diabetes and obesity (208). More interestingly, is it also elevated in menopausal women. Apparently, Rbp4 concentration is increased in endocrine disorders where insulin resistance, glucose tolerance, and lipid metabolism are impaired (209), all of which have shown to be altered in aged oocytes (Figure 6.3 and 6.4). Thus, the increasing indications that this binding protein is correlated with impaired metabolic functions in aged oocytes suggest that raised levels of Rbp4 in aged oocytes could be a key marker for oocyte quality.

Taken together, as RA signalling is a key determent in embryogenesis, and deviation from normal levels results in multiple, sometimes severe developmental malformations (185) This single oocyte RNA-seq analysis has revealed a new role of endogenous retinoid metabolism in oocyte competence. Evidently, one-year oocyte lacks retinoid that composes essential maternal factors required to support further embryonic development prior to activation of the embryonic genome. Dysregulation of the RAR activation pathway could be a major source of poor quality in aged oocytes. Consistent with this conclusion, *in vitro* maturation studies have demonstrated that using a Retinoid supplement have a positive impact on oocyte maturation and enhance embryonic survival (210–212). In light of these

results, dysregulation of RAR activation pathway can be, in part, result in a compromised oocyte quality.

The mitochondrial L-carnitine shuttle pathway was the second pathway involving three genes (Table 6.4) that is primarily involved in fatty acid oxidation. Fatty acid is considered to be a major source for energy production required during oocyte maturation (24). Oocytes store lipids in droplet organelles in the cytoplasm that re-localize during maturation; these droplets are closed to mitochondria (213). The fatty acids are metabolized by ß-oxidation in the mitochondria for the production of Adenosine triphosphate (ATP). Acyl-CoA synthetase (Acsbg1) is an enzyme that converts fatty acid to Fatty-CoA which is transported into the outer mitochondrial membrane through Carnitine (cpt1). Fatty acyl-carnitine is then transported to the mitochondria matrix by diffusion, where free cpt1 is then transferred back to be used for the next cycle.

Carnitine is therefore the rate-limiting step in fatty acid oxidation in the mitochondria. Changes in the lipid metabolism profile may be a result of disruption of mitochondrial functions in aged oocytes. The high expression of both enzyme Acsbg1 and cpt1 in aged oocytes (Table 6.4) may indicate excessive ß-oxidation, which leads to excessive cleavage of fatty acid utilized to generate ATP. It is very important to maintain the balance between the uptake and oxidation of fatty acid.

Excessive ß-oxidation could be indicating a higher activity of mitochondria, which is deleterious to oocyte quality especially as the need for ATP production becomes more necessary towards maturation. Particularly, at the blastocyst stage, where the growing

embryo needs more ATP for implantation and further development. It therefore could be hypothesized that an aged oocyte has excessively overused its energy, meaning that a fertilised oocyte would have inadequate lipid reserves and low ATP levels that consequently lead to developmental arrest or abnormal cell division within blastomers (214).

Hamatani *et al.* (2004) indicated that mitochondrial dysfunction was the most prominent functional group altered in aged mouse oocytes. Genes related to energy pathway showed decreased expression, suggesting a decline in ATP level and reduced mitochondria activity in aged oocytes (113). While these findings agree with the results of this study to some extent in that mitochondrial function and energy production were impaired, with regard to whether ATP is elevated or not, the Hamatani study used microarrays on a group of 500 MII oocytes, whereas a single oocyte at GV stage was used in this study. The stage of maturation may explain the conflicting findings regarding energy levels in aged mouse oocytes, as ATP production decreases after maturation. Another possibility is that ß-oxidation might have been inhibited, blocking the fatty acid from entering the mitochondria. As a result, the aged oocyte might have failed to utilize both enzymes Cpt1 and Acsbg1, resulting in the accumulation of both enzymes in the aged oocyte's mitochondria.

There is emerging evidence for a relationship between insulin resistance and beta-oxidation (215). Studies have suggested that insulin inhibits fatty acid oxidation by blocking lipolysis in adipose tissue and stimulates lipogenesis (216). This is likely to be occurred in aged oocytes as well. Supporting this hypothesis is the high expression of Solute carrier family 27 (Slc27a2) in aged oocytes (FC= 3.58), which is a long chain fatty acid. Clinical studies suggest that this gene is potentially a key marker for insulin resistant syndrome. This may

be apparent, as Slc27a2 was indeed involved in activation of diabetes mellitus signalling pathways in aged oocytes (Table 6.5). In addition, CCAAT/enhancer binding (Cebpb), a transcription factor involved in the diabetes mellitus pathway (Table 6.5), was reported to be a key regulator in fat metabolism. Cebpb was found to control total body fat content by governing fatty acid mitochondrial β-oxidation in *C. elegans*. as Cebpb deficiency has resulted in reduced lipid accumulations (217). Moreover, accumulation of Cebpb has induced fatty acid synthesis and triggered insulin resistance pathways. One study has shown that overexpression of Cebpb has induced Mapk10 and Protein Kinase C beta (Pkcb) in adipose mouse tissue (218). Both of these genes were altered in the aged oocyte in this study. Further evidence that enhanced activity of Pkcb is linked with activation of insulin resistance pathway. Free fatty acid-induced insulin resistance is associated with activation of protein kinase C and alterations in the insulin signalling cascade. Therefore, these mechanisms must be well coordinated. They are altered in oocyte ageing and consequently disturb oocyte biological processes.

Collectively, it is evident that the dysregulation of the diabetes mellitus pathway potentially contributes to the mechanism of impaired lipid metabolism. Whether increased fatty acid beta-oxidation prevents cytoplasmic lipid accumulation and decreases insulin resistance is still controversial. What could be ultimately concluded from both assumptions is that impaired lipid synthesis in aged oocytes is linked with insufficient ATP levels required for subsequent normal development.

sc-RNA-seq revealed that age-related changes in mitochondria function primarily contribute to alterations in oocyte energy. A balanced lipid haemostasis and ATP level are crucial to

meet the demands for successful maturation, fertilisation and subsequent embryo development. Data showed that three pathways, mitochondrial L-carnitine, fatty acid activation, and Linolenate Biosynthesis (all of which are involved in fatty acid metabolism) were dysregulated in the GV of the aged oocyte, which was not seen in the 9-week-old oocyte. Thereby, mitochondrial dysfunction is mainly contributed to oocyte senescence. These age-associated deteriorations play a main role in the depletion of the intracellular metabolic process.

Three canonical pathways, signalling by Rho family GTPases, sertoli cell-sertoli cell gab junction and integrin linked pathway were altered (Table 6.6 and 6.7). All of these primarily function in cytoskeleton organization, microtubule and spindle formation, as well as in oocyte communication with the microenvironment. There is a high expression of the cytoskeleton molecules such as Acta2, myl9, Vim, spindle formation and Sept9, and a decreased level of cadherin, which related to cellular adhesion and gap junctions, control ion and molecule flow through adjacent cells. This revealed the abnormal global alteration in the cytoskeleton of the aged oocyte compared with the young adult GV. Interestingly, most of the genes involved in these pathways were expressed in similar patterns in pre-pubertal oocytes (discussed in Chapter 5) known as having poor developmental competence. This cytoskeleton alteration may consequence in chromosomal abnormalities, inhibition of the extrusion of first polar body, increased cellular fragmentation. Based on the fact that ATP is crucial for driving meiosis, it may be indicated that mitochondria have not generated sufficient energy required for completing functions properly.

The Gp6 pathway involves: collagen alpha 1(Colla1), collagen type alpha 2 (Colla2) and collagen typeIII alpha1 (Cola3a1). Collagen I and Collagen III were found to be the most

expressed ECM molecules in mouse ovary (219). The three collagen transcripts had the highest fold change in the differential genes expressed in aged oocyte (Table 6.8). Changes in Collagen level are also linked with tissue fibrosis. For example: collagen accumulates in heart with age, leading to cardiovascular dysfunction (220) . Therefore, the elevated level of collagen primarily reflects the ovary's advanced age. It has been documented that ovaries form older mice are more rigid compared to those from younger mice (221) Also, a recent observation confirmed that fibrosis increases in ovarian stroma with advanced age (222). Thus, the ovarian microenvironment may promotes the accumulation of collagen in the one-year-old oocyte, thus, compromising oocyte quality.

Changes in the melatonin degradation pathway was one interesting finding that correlated with the nitric acid and reactive oxygen species (ROS) pathway. Myeloperoxidase (MPO) is a peroxidase enzyme that is encoded by the MPO gene, derived from leukocytes (223). Myeloperoxidase is released from activated leukocytes at inflammatory sites, and uses hydrogen peroxide to generate a wide array of ROS (224). Reactive oxygen and nitrogen species are a family of antimicrobial molecules derived from nitric oxide and superoxide anions generated during normal metabolic reactions. However, if these molecules and their metabolites are not adequately neutralized, they may pose a serious threat to cellular viability. Therefore, a balanced level between ROS and antioxidants are detrimental for the developing oocyte health and to maintain a normal ovarian physiology (225).

Myeloperoxidase (MPO) was reported as a key marker for oxidative stress in ovarian cancer and premature ovarian failure (226). High levels of ROS deteriorate oocyte quality by inducing apoptosis (227). therefore the accumulation of MPO promotes oxidative damage

in aged oocytes. More interestingly, one study has shown that the MPO serum level was found to be elevated in PCOS patients, in particular those who suffer from insulin resistance (228). A further observation supporting the association between MPO and insulin resistance is that mice deficient in MPO exhibit an improved insulin resistance, elevating energy consumption (229). Taken together, our sc-RNA-seq revealed a novel mechanism of MPO in age-related deterioration in aged oocytes. The implication of MPO in the dysregulation of nitric oxide and ROS pathways emphasises the contribution of MPO in modulating oocyte functions. ndeed, oxidative stress is one of the primary factors that have a direct negative impact on oocyte quality. Oxidative stress is believed to be a consequence of mitochondrial dysfunction (230).

To further understand the main biological functions represented in the GV of aged oocyte, transcripts were grouped according to their molecular functions. The characterization of the networks shown in Figure 6.6 and 6.7 is an agreement with the dysregulation of canonical pathways. For instance: RAR activation pathways, Sertoli cell to Sertoli cell pathway, signalling by Rho GTPase family, ILK pathway are implicated in growth and development functions. The second and third network were cell death and organismal injuries and abnormalities included genes involved in apoptosis, inflammation that related with mitochondrial dysfunction. Finally, lipid metabolism is potentially the most altered metabolite in aged oocyte.

Based on these data, this study proposes that mitochondrial dysfunction appears to be the main factor and the first to be altered at the earliest stage of oocyte senescence. Despite the fact that the oocyte is compromised on multiple levels, ATP supply is reported to be altered

early in 8 months oocyte (51).. Therefore, it could be assumed that energy dramatically declines, which is then followed by a cascade of disrupted intracellular functions. Most biological functions are dependent on mitochondrial activity to obtain the necessary energy for accomplishing and achieving oocyte maturation events, including cell cycles and organelle reorganization (16). Therefore, mitochondrial dysfunction triggers the dramatic dysregulation of several pathways, thereby leading to declined developmental competence for an aged oocyte. Further experiments are necessary to confirm this speculation.

6.5. Conclusion

Oocyte ageing can be summarized as progressive failure in the maintenance of oocyte haemostasis, leading to impaired developmental competence. Single cell transcriptomics have enabled the investigation of novel pathways that are involved in oocyte ageing. Mitochondrial dysfunction, lipid metabolism and retinoic acid metabolism play a major role in age-related deterioration in oocytes.

Elucidation of aberrant transcriptome profile in aged oocytes could lead the development of potent, non-invasive markers that can be used for selecting the oocytes with the highest competence. Furthermore, preventing lipid accumulation or balance retinoic acid could be potential methods to restore the normal levels of these metabolites in aged oocytes, hence reversing the ageing process and ultimately enhancing the outcomes of assisted reproduction for older women.

Chapter 7
Transcriptome Analysis
of the
Cumulus Cells Derived from Single Oocyte

7.1. Introduction

Oocyte development is coordinated by bidirectional interaction between the oocyte and the surrounding cumulus cells (CCs). The CCs are in physical contact with the oocyte; together they form the cumulus-oocyte complex (COC) (231,232). Cumulus cells are a key player in transmitting signals to and from the oocyte via cytoplasmic projection and the gap junction at the oocyte surface (233). The signalling between the oocytes and their CCs is vital for producing a viable and competent oocyte (234).

The oocyte is the central regulator of follicle growth, regulating a broad range of follicular CCs' differentiation, proliferation and luteinisation. In turn, CCs are responsible for the external follicle's metabolism of the glucose and pyruvate used for energy production in the oocyte. Intimate communication between the oocyte and its CCs is especially crucial in terms of the nuclear and cytoplasmic maturation of the oocyte. Hence, both oocytes and CCs grow in a synchronised manner (57,235).

Since oocytes and CCs share the same environment, it is believed that CCs are able to disclose relevant information concerning an oocyte. Numerous studies have proposed using CCs as a non-invasive approach of predicting an oocyte's quality and developmental potential. Previous studies have demonstrated that CC apoptotic rates are linked with abnormal oocyte morphology, incomplete oocyte maturation, and impaired fertilisation (67,236,237). Intriguingly, CCs associated with chromosomally abnormal oocytes were found to exhibit a characteristic gene pattern. This observation indicates that a suboptimal follicular microenvironment potentially contributes to meiotic chromosomal abnormalities in oocytes. Thereby, aneuploidy oocytes are correlated with abnormal gene expression in

CCs (69). These studies collectively suggest that changes in the microenvironment mediated by CCs are linked with reduced oocyte quality. Both intrinsic (oocyte) and/or extrinsic (follicular) factors may be involved in the oocyte's decline in quality (55,238). However, it is not clear whether companion abnormal CCs induce the oocyte abnormality or whether poor oocyte lead to the apoptotic CCs.

Although the developmental potential of a pre-pubertal oocyte and an adult oocyte clearly differ, the role of CCs associated with pre-pubertal oocytes is poorly understood. Since CCs support pre-pubertal oocyte maturation *in vitro*, it is doubtable whether incomplete pre-pubertal oocyte competence is brought about by the companion CCs. What may support this hypothesis is a study demonstrating that follicular development occurs via two waves of folliculogenesis in the mouse ovary. The first primordial follicles begin to appear before puberty and are then gradually replaced by the second follicle wave, which supports adult fertility (239).

In addition, an aged follicular microenvironment could impact oocytes by retaining a characteristic transcriptional footprint in the surrounding CCs. A better understanding of the transcriptomic mechanism that governs the CCs associated with oocytes of different ages may result in new insights regarding CCs' contribution to the production of competent oocytes. No high-throughput study based on gene expression profiling has been conducted on pre-pubertal CCs, and little is known about age-related dysfunction in CCs. This chapter sheds light on the key genes implicated in CC development.

The aim of this chapter is to use RNA sequencing to analyse the transcriptomic differences among CCs at 3 weeks (pre-pubertal), 9 weeks (mature), and 1 year (aged) isolated from single oocytes to identify age-related changes potentially reflecting diminished oocyte development potential.

7.2. Methods

24 individual COCs were recovered from three different ages of female C57BL/6 mice: 3 weeks old (pre-pubertal oocyte), 9 weeks old (adult oocyte) and one-year-old (aged oocyte). Mice were stimulated with PMSG 48 hours prior to the collection of ovaries. Cumulus cells were mechanically separated from each individual COC in which the oocyte is at GV stage, washed in PBS/BSA and kept in lysate buffer at -80°C. Ployadenylated RNA was isolated, amplified and processed for Illumina sequencing using Smart-seq2 approach (149) (preparation details in section 2.4 and 2.5).

The CC sequencing data were processed using FastQC format. Duplicates and adaptors were removed. However, generated reads from CC sample were re-mapped twice in order to remove the duplicates sequences. Then, only uniquely mapped reads were retained using STAR software. Differential gene expression was carried out using edgeR, the data were then enriched for pathways analysis using IPA, web based software, and Metacore Thompson (detailed in section 2.9)

7.3. Results

The analysis of CC transcriptomics is divided into four sections:

- Sequencing quality
- Principle component analysis (PCA)
- Differential gene expression
- Functional analysis

7.3.1. Sequencing Quality

7.3.1.1 Mapping Rate

Illumina sequencing resulted in 24 million reads per CC (n=24; 8 CCs of each age [3 weeks, 9 weeks and 1 year]). In Figure 7.1, the three columns represent the mapping rate for each CCs sample. Of the approximately 24 million reads generated from each CC, an average of 18 million reads were mapped, and about 13.5 million were mapped uniquely. Roughly 54.5% of the reads were uniquely mapped to the mouse genome. The mapping rate differed among the samples (Figure 7.1), with certain samples featuring a lower mapping rate than others.

7.3.1.2. Cumulus cell-specific gene expression

To evaluate the validity of the generated sequencing data, a number of CC-specific genes were assessed to ensure high-quality sequencing. A boxplot was generated to demonstrate the average gene expression in each of the 8 samples from each of the 3 age groups. The genes in each group were sorted based on the level of expression from highest to lowest. These genes were selected from previous mice CCs studies showing that these genes found to be higher in cumulus cells than other follicular cells such as mural granulosa cells. Because this is the first sc-RNA-seq for mouse CCs, it was fundamental to relay on previous finding (Table 7.1, Figure 7.1)

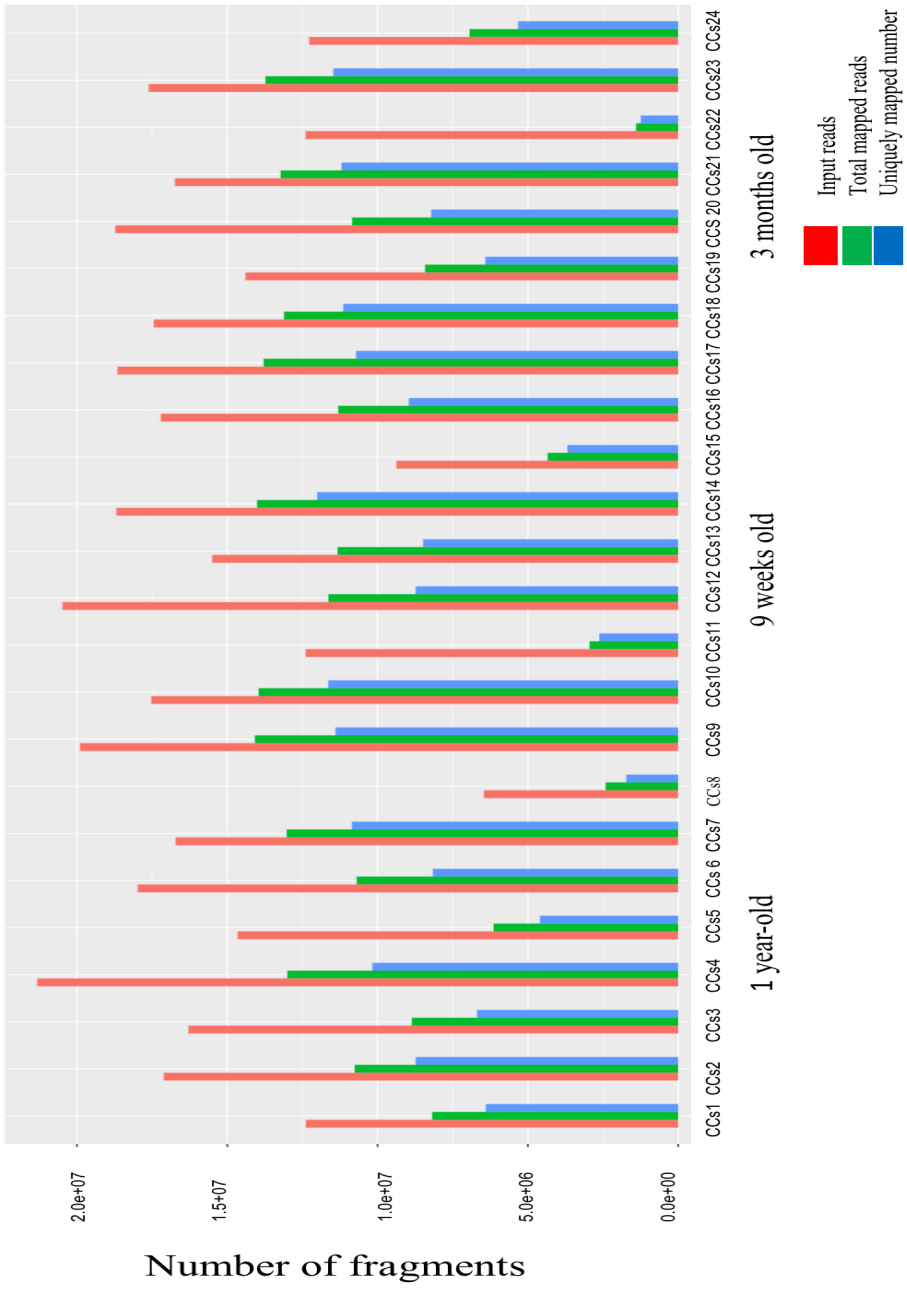


Figure 7.1 Mapping rate graph for CC samples. Each cumulus cells derived from a single oocyte (n=8) at 3 weeks, 9 weeks, and 1 year of age generated approximately 24 million reads. Of these CCs, roughly 18 million reads were mapped, and about 14 million reads were mapped uniquely.

Table 7.1: List of CCs specific genes used as positive control genes for assessing the quality of CCs sequencing, adopted from (58)

Gene Name	Gene Symbol
gremlin 1	Grem1
core-binding factor, runt domain, alpha subunit 2, translocated to, 3 (human)	Cbfa2t3
melanocortin 2 receptor accessory protein	Mrap
tubulin, beta 2B	Tubb2b
calcium channel, voltage-dependent, P/Q type, alpha 1A subunit	Cacna1a
actin filament associated protein 1-like 2	Afap1l2
mannan-binding lectin serine peptidase 1	Masp1
Vimentin	Vim
plexin B1	Plxnb1
camello-like 3	Cml3
RNA binding protein, fox-1 homolog (C. elegans) 3	Rbfox3
protein phosphatase 1, regulatory (inhibitor) subunit 3G	Ppp1r3g
cytochrome P450, family 19, subfamily a, polypeptide 1	Cyp19a1
inhibin beta-A	Inhba
CCAAT/enhancer binding protein (C/EBP), delta	Cebpd
slit homolog 3 (Drosophila)	Slit3
tissue inhibitor of metalloproteinase 2	Timp2
maternally expressed 3	Meg3
sterile alpha motif domain containing 4	Samd4
inhibin beta-B	Inhbb
inhibitor of DNA binding 2	Id2
procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha II polypeptide	P4ha2
GTP cyclohydrolase 1	Gch1
protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 3	Ppfia3
inhibitor of DNA binding 1	Id1
stearoyl-Coenzyme A desaturase 1	Scd1
autoimmune regulator (autoimmune polyendocrinopathy candidiasis ectodermal dystrophy)	Aire
fatty acid binding protein 3, muscle and heart	Fabp3

vascular endothelial growth factor A	Vegfa
activation-induced cytidine deaminase	Aicda
melanoma inhibitory activity 1	Mia1
disabled homolog 1 (Drosophila)	Dab1
high mobility group nucleosomal binding domain 3	Hmgn3
serum deprivation response	Sdpr
low density lipoprotein receptor-related protein 8, apolipoprotein e receptor	Lrp8
like and metalloprotease (reprolysin type) with thrombospondin type 1 motif, 5 (aggrecanase-2)	Adamts5
RNA imprinted and accumulated in nucleus	Rian
glutathione S-transferase, alpha 4	Gsta4
frizzled homolog 2 (Drosophila)	Fzd2
cyclin-dependent kinase inhibitor 1A (P21)	Cdkn1a
growth arrest and DNA-damage-inducible 45 gamma	Gadd45g
neuromedin B	Nmb
v-ral simian leukemia viral oncogene homolog B (ras related)	Ralb
F-box and leucine-rich repeat protein 15	Fbxl15
dopamine receptor 4	Drd4
thymosin, beta 4, X chromosome	Tmsb4x
solute carrier family 2 (facilitated glucose transporter), member 1	Slc2a1
Phosphofructokinase, platelet	Pfkp
serine (or cysteine) peptidase inhibitor, clade A, member 5	Serpina5
GSG1-like	Gsg1l

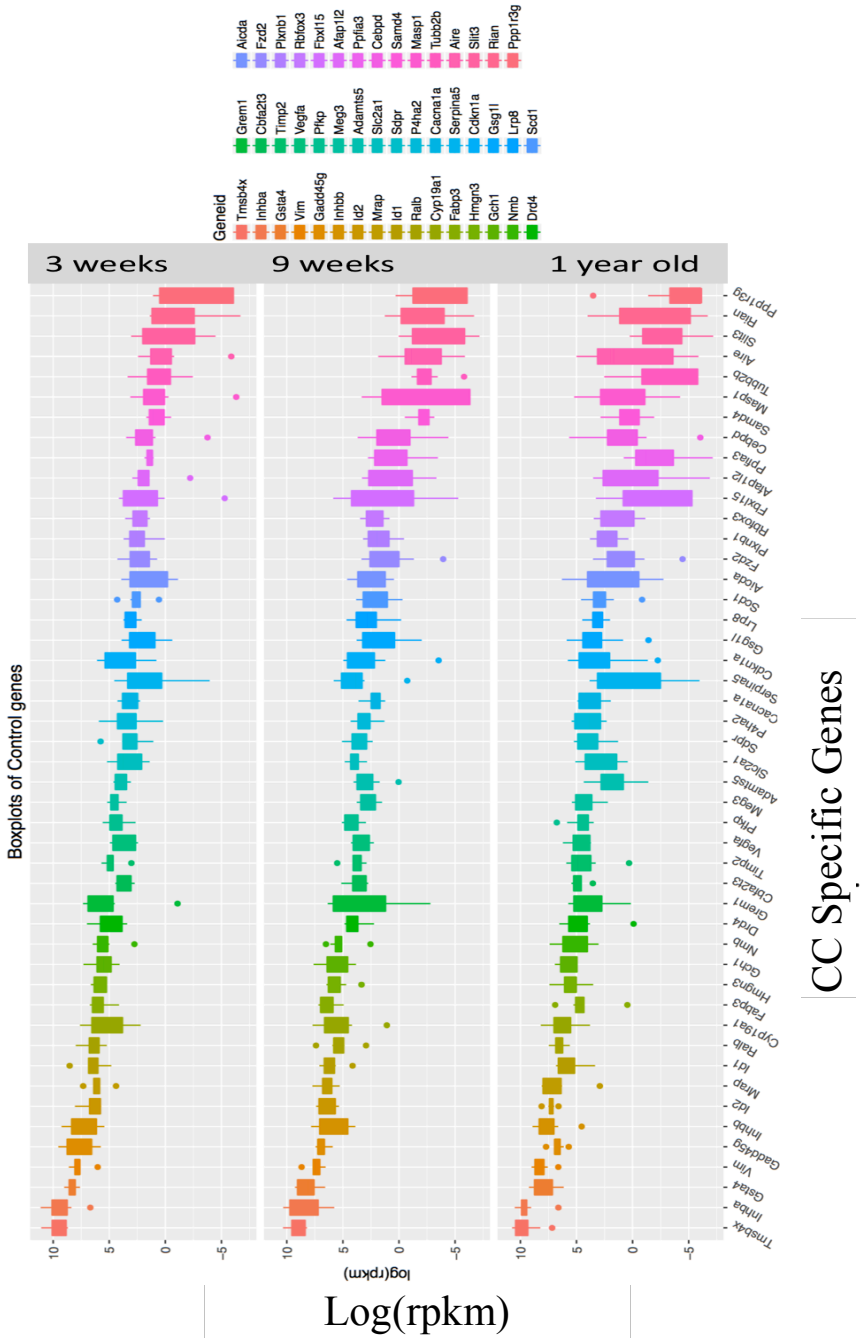


Figure 7.2 Boxplot demonstrating the average expression of the CC-specific genes in the sequencing data. Boxes correspond to the first and third quartiles; whiskers extend to 1.5 times the interquartile, and the dots indicate outliers. The genes in each group are sorted from the highest to lowest level of expression.

7.3.1.3. Number of expressed genes (quality assessment)

The study stipulated that a gene was retained if it expresses a threshold of 1 count-per-million (cpm) for any given sample from each age group. The number of expressed genes averaged 12,000 at 3 weeks, 11,600 at 9 weeks, and 11,300 at 1 year. Identifying the number of genes expressed in each sample was vital for ensuring that the sequencing data were reliable. The 3-week-old CCs demonstrated a smaller black box, indicating the consistent value of the expressed genes of the CC samples whereas variable values of the global expressed genes of the CCs samples in both 9-weeks and 1-year old (Figure 7.3).

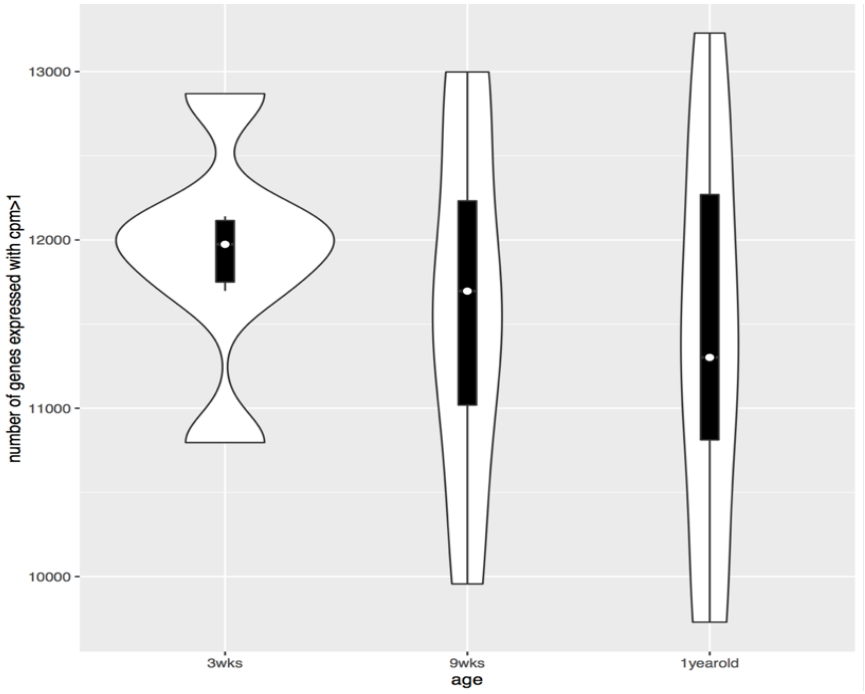


Figure 7.3 A Villon plot of the number of genes expressed in the CCs from each age group. A threshold of 1 cpm was used to deem a gene expressed for any given sample. The larger the black box is, the more variable the number of expressed genes in the associated with the most consistent gene expression figures.

7.3.2. Principal component analysis

To investigate gene expression pattern variations among the CCs within a given age group, the sample data were plotted along the first and second PCAs to identify age-related differences (Figure 7.4). The generated sample clusters enabled an evaluation of similarities and differences regarding global transcriptome profiles between the age groups. The clusters did so by revealing how the principal components (each dot represents a sample in Figure 4) of all three ages were correlated with one another.

Surprisingly, CCs have not clustered clearly. Principal component analysis revealed that the 1-year-old oocytes featured a distinct transcription pattern that distinguished them from the oocytes belonging to the other two age groups. Thus, transcriptomic profiling seemed to likewise differentiate the 1-year-old CCs from those CCs from the 9-week and 3-week old groups.

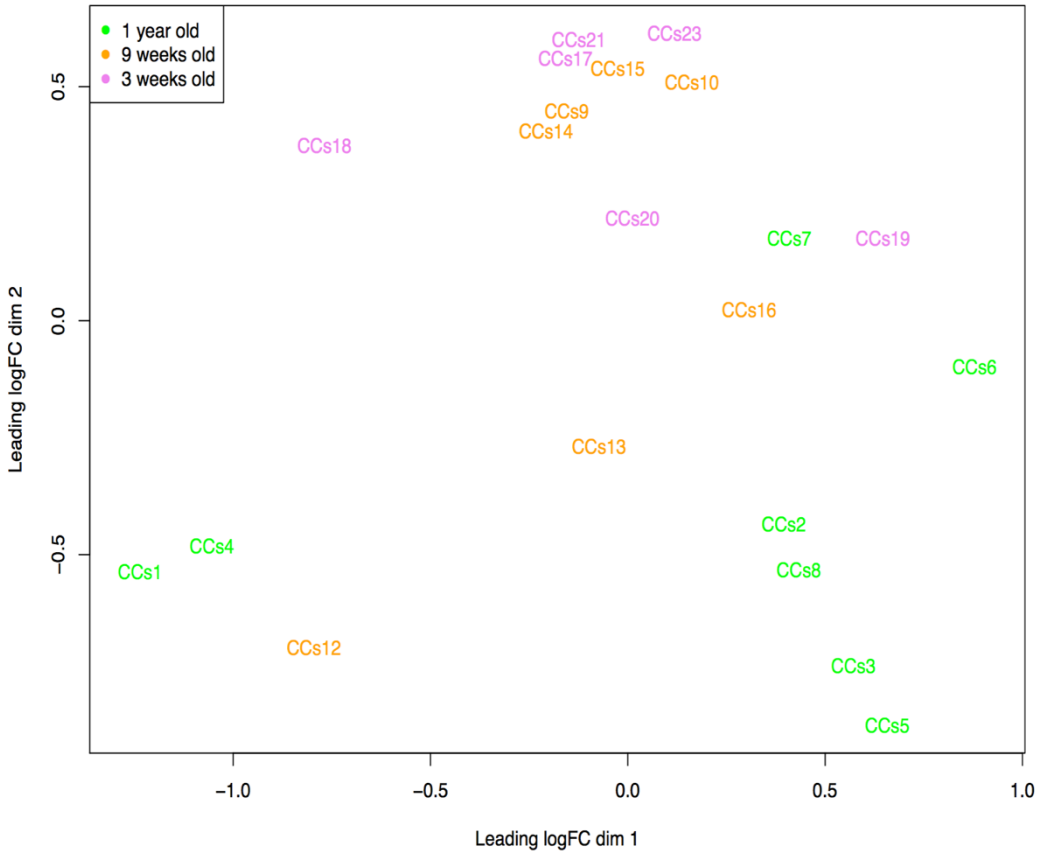


Figure 7.4 Principle component analysis (PCA) plot based on the 500 most variable genes; each point represents a sample deriving from CCs stripped from the individual mouse oocytes. The data reveal how the samples from the three age groups are related to each other. Most of the 1-year-old mouse samples separated well, whereas the 3-week-old samples did not separate from the 9-week-old samples as clearly.

7.3.2.1. Pre-pubertal versus adult cumulus cells (3 weeks versus 9 weeks)

At 3 weeks, the CCs had not distinctly clustered from those at 9 weeks (Figure 7.5), indicating that all components of the two groups (the 3-week and 9-weeks oocytes) possessed comparable global gene expression patterns. Six pre-pubertal CCs sample were plotted into the PCA where two samples from the pre-pubertal CC group, marked by low sequencing quality, were excluded from this PCA. This is because the suspicion that they might have adversely impacted the sample clustering during the PCA. However, even after plotting the high-quality CC samples, they failed to separate from each other, suggesting that the 3-week-old and 9-week-old CCs featured comparable transcription patterns.

7.3.2.2. Aged versus young-adult cumulus cells (1 year versus 9 weeks)

Principal component analysis plotting revealed that the oocytes from the 1-year-old mice (green squares in Figure 7.6) were highly correlated with each other, as they clustered separately from the oocytes taken from the 9-week-old mice (orange circle in Figure 7.6). Moreover, PCA demonstrated that the transcription patterns of the aged CCs were influenced by advancing age.

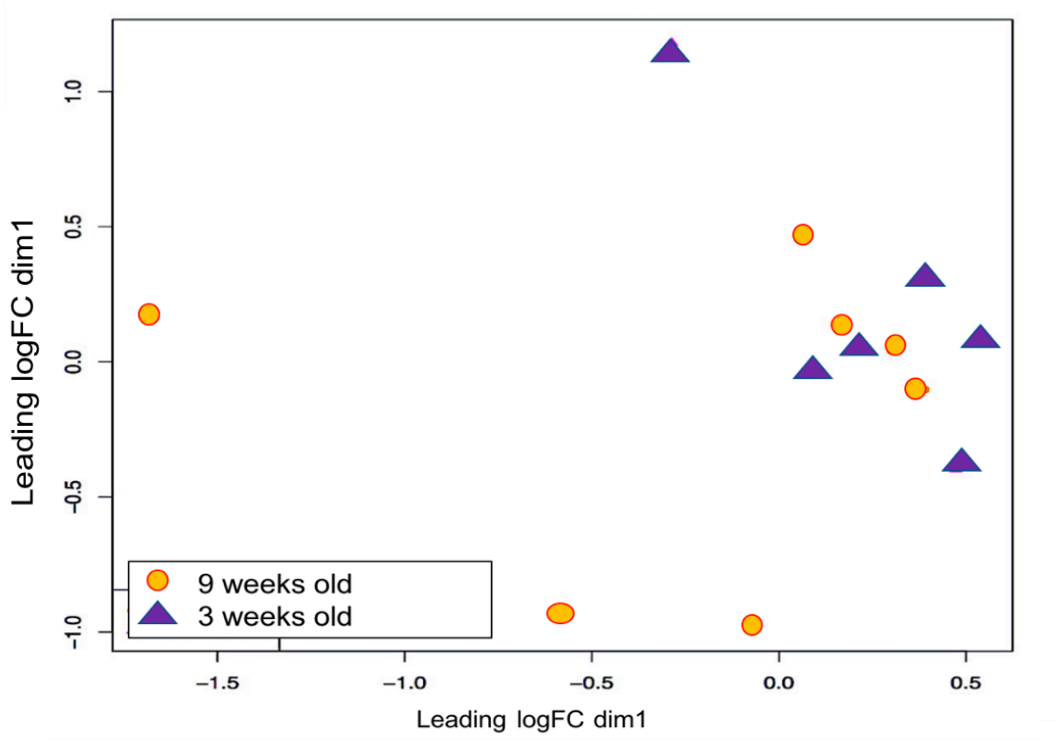


Figure 7.5 Principle component analysis (PCA) revealed that the CCs belonging to the pre-pubertal mice (3 weeks) and the 9-week-old mice did not have highly pronounced differences.

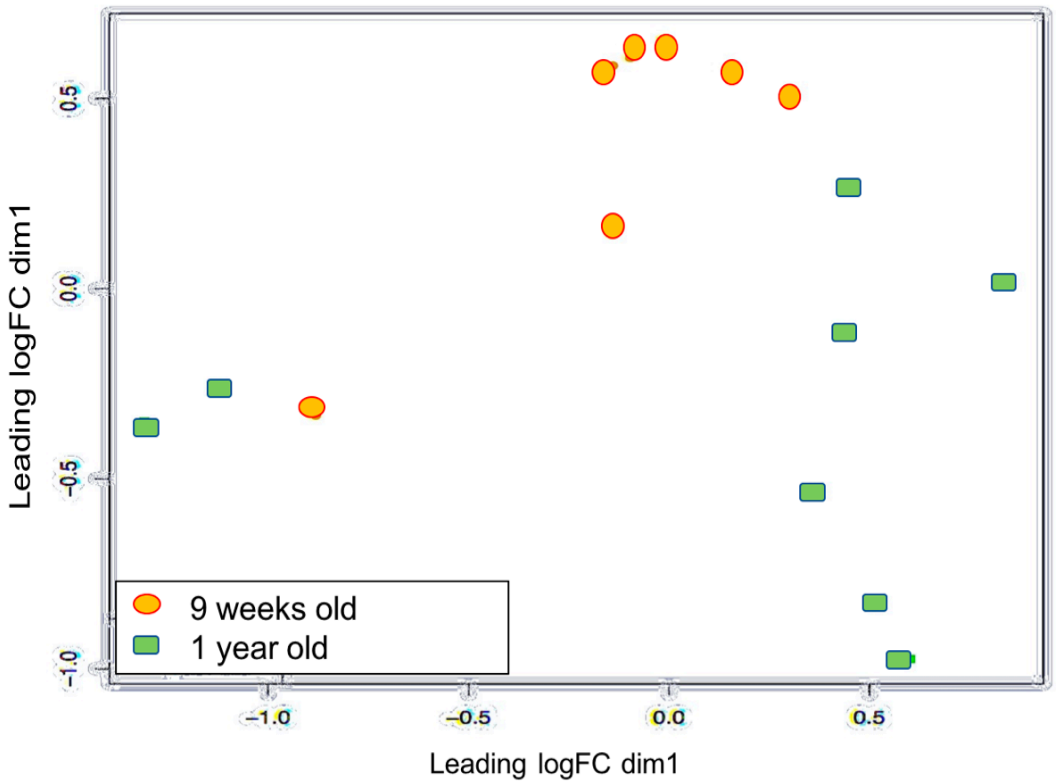


Figure 7.6 Principle component analysis (PCA) of oocytes at 1 year (aged) versus 9 weeks (young adult). Oocytes from the aged mouse cluster were fairly separately from those of the young adult mice.

7.3.3. Differential gene expression

7.3.3.1 Pre-pubertal versus young-adult cumulus cells

Surprisingly, no genes were identified as differentially expressed in the 3-week-old CCs as compared to the 9-week-old CCs. This finding suggests possible similarities between pre-pubertal and young-adult CCs in terms of transcriptional activity.

7.3.3.2 1-year-old versus 9-week-old cumulus cells

The scatterplot comparing the 1-year-old CCs and the 9-week-old CCS depicts the differentially expressed genes in red (Figure 7.7; adjusted *P* value <0.05). The RNA-seq detected a total of 20 genes that were differentially expressed in the 1-year-old oocyte versus the 9-week-old oocyte. The difference in gene expression indicates changes in the CC transcriptome profiles of the oocytes of various ages, demonstrating an alteration caused by age. Interestingly, all of the differential genes had a fold change as low as negative 10.

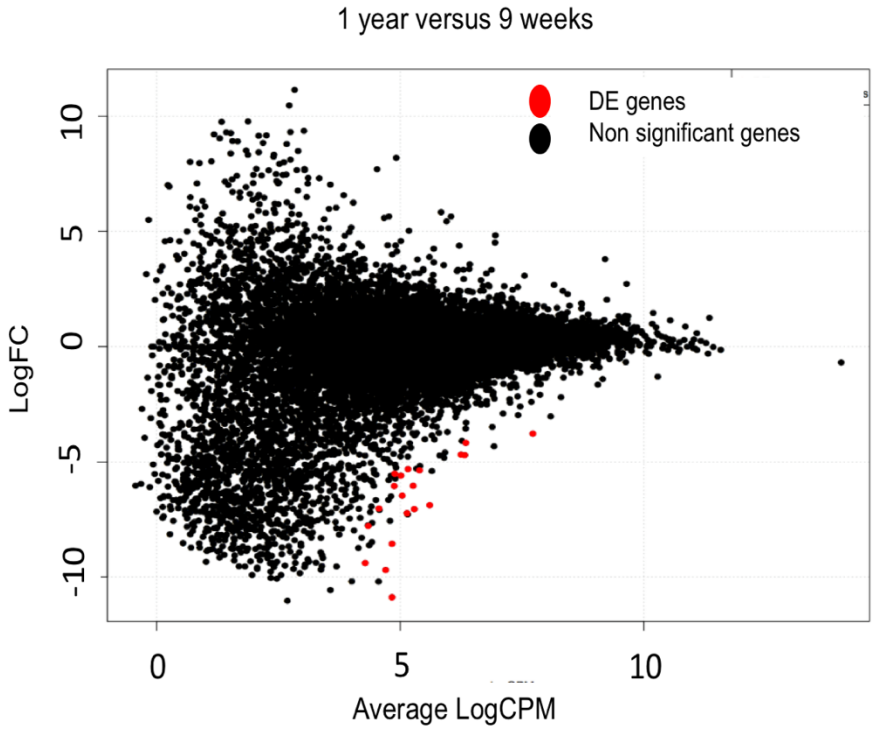


Figure 7.7 Scatterplot of average counts (cpm) versus log2-fold changes for genes generated by DEseq. Genes were retained if 1 count-per-million (cpm for any given sample.

7.3.4. Functional network analysis

To assess the transcriptomic differences in a comprehensive manner, the False Discovery Rate (FDR) cutoff was increased to 0.1, and consequently, the number of genes included in the analysis also increased. Using this approach, 57 genes were differentially expressed in aged CCs, 54 genes were downregulated in the aged CCs, and only 3 genes were upregulated. A list of DE genes is provided in Table 7.2

Table 7.2: List of the expressed differential genes in 1-year CCs relative to 9-weeks CCs (FDR<0.1)

Gene Name	Geneid	logFC	FDR
NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 4-like 2	Ndufa4l2	8.191785489	0.089647315
ATP-binding cassette transporter G1 (Abcbg1)	Abcb1b	4.513904167	0.089647315
Farnesyl pyrophosphate	Fdps	2.27934726	0.089647315
Replication rermination factor 2 domain containing 1,	Rtfdc1	-2.259085327	0.092066339
Branched chain ketoacid dehydrogenase kinase	Bckdk	-2.430085427	0.089647315
Pre-MRNA processing factor 18	Prpf18	-3.447487829	0.089647315
RAS Like family 10 member A	Rasl10a	-3.484978104	0.080224553
Coenzyme Q6 monooxygenase	Coq6	-3.705487574	0.089647315
Connective tissue growth factor	Ctgf	-3.776659609	0.039274005
Autophagy-related protein 13	Atg13	-4.093071275	0.080224553
Tumor protein D52	Tpd52l1	-4.173442836	0.038648039
Putative splicing factor, arginine/serine-rich 14	Sugp2	-4.188139125	0.089647315
Kazal Type Serine Peptidase Inhibitor Domain 1	Kazald1	-4.684360739	0.027560214
Kelch Like Family Member 31	Klhl31	-4.706772123	0.027560214
Ectonucleoside triphosphate diphosphohydrolase-1	Entpd1	-4.727967481	0.089647315
Ras-Related Protein Rab-43	Rab43	-5.023752809	0.089647315
Vacuolar protein sorting-associated protein 8 homolog	Vps8	-5.312694674	0.039274005
G patch domain containing 4	Gpatch4	-5.367285394	0.007996236
zinc finger protein, multitype 2 protein	Zfpm2	-5.490073991	0.089647315
Coiled-Coil Domain Containing 157	Ccdc157	-5.50881712	0.039274005
Ethylmalonic encephalopathy 1 protein	Ethe1	-5.726452391	0.089647315
TRNA-YW Synthesizing Protein 3 Homolog	Tyw3	-5.769272417	0.089647315
Membrane Bound O- Acyltransferase Domain Containing 1	Mboat1	-5.792604659	0.079526906
Glycerophosphodiester Phosphodiesterase Domain-Containing Protein 2	Gdpd2	-5.820379117	0.089647315

Gene Name	Geneid	logFC	FDR
Interleukin 17A	Il17ra	-6.040409767	0.038648039
Elastin Microfibril Interfacer 2	Emilin2	-6.050495185	0.039274005
Homeobox protein TGIF2	Tgif2	-6.168923687	0.089647315
Interleukin 15 Receptor Subunit Alpha	Il15ra	-6.343953371	0.089647315
Transmembrane Protease, Serine 13	Tmprss13	-6.46907042	0.038648039
Alanine And Arginine Rich Domain Containing Protein	Aard	-6.650722631	0.089647315
	1810014B01Rik	-6.656760661	0.079526906
Spondin-1	Spon1	-6.881817396	0.027560214
Pyruvate Dehydrogenase Kinase 4	Pdk4	-7.025715078	0.038648039
Tumor necrosis factor receptor superfamily member 19L	Relt	-7.038369037	0.089647315
Osteomodulin	Omd	-7.050367035	0.000802651
	E030030I06Rik	-7.083605055	0.076591473
Lymphocyte Antigen 96	Ly96	-7.18224366	0.089647315
Protien Kinase D1	Prkd1	-7.222064898	0.038648039
Tetraspanin-4	Tspan4	-7.279420435	0.066336277
Kruppel-like factor 8	Klf8	-7.429391148	0.089647315
Ataxin-1-like	Atxn1l	-7.456732636	0.057454785
Tubulin Folding Cofactor E	Tbce	-7.757538966	0.089647315
Loss Of Heterozygosity 12 Chromosomal Region 1	Loh12cr1	-7.780678509	0.038648039
Asteroid Homolog 1	Aste1	-7.979489861	0.089647315
Chemokine (C-X-C motif) ligand 1	Cxcl1	-8.483264428	0.089647315
Collagen Type IX Alpha 2 Chain	Col9a2	-8.552240981	0.038648039
Calcium Voltage- Gated Channel Subunit Alpha1 D	Cacna1d	-8.738633482	0.089647315
	4632415K11Rik	-9.22329134	0.081213016
TBC/LysM associated domain containing 1	2010111I01Rik	-9.390542088	0.034921709
Ras-association domain family protein 6	Rassf6	-9.52212348	0.089647315

Gene Name	Geneid	logFC	FDR
Cyclin D3	Ccnd3	-9.541985649	0.089647315
Ankyrin Repeat And SOCS Box Containing 4	Asb4	-9.691720932	0.038648039
Platelet Derived Growth Factor Receptor Like	Pdgfrl	-10.03738935	0.09825491
Zinc finger protein 956	Zfp956	-10.18321396	0.089647315
Ydroxyacylglutathione Hydrolase Like	Haghl	-10.19330783	0.079526906
Retrotransposon Gag like 8C (Rlt8)	Cxx1c	-10.87482248	0.006075601

The biological networks of the differentially expressed genes were generated using IPA software. Table 7.3 illustrates the networks potentially involved in the CCs' senescence and the genes involved in each network.

1. First Network: Endocrine system, development and function, organ Morphology, tissue morphology (Figure 7.8).
2. Second Network: Cell death and survival, DNA replication, recombination and repair, Organismal injury and abnormalities (Figure 7.9).
3. Third Network: Cancer, Organismal injury and abnormalities, Haematological disease

Table 7.3: Top functional networks involved in the aged CCs (1-year-old) in comparison to young adult CCs (9-weeks-old), genes involved in each network, generated by IPA.

	Network name	Score	Focus Molecule	Genes
1	Endocrine system, development and function, organ Morphology, tissue morphology	33	15	Abcb1b, Bckdk, Canaid, Ccnd3, Col9a2, Ctgt, Cxcl2, Entpd1, Fdps, Ili5ra, Ili7ra, Entpd1,Ly96, Prkd1, Spon1, Zfpm2
2	Cell death and survival, DNA replication, recombination and repair, Organismal injury and abnormalities.	30	14	Asb4, Atg13, Atxnil, Emilin2, Ethe1, Klf8, Ndufa4l2, Pdgfrl, Rab43, Rassf6, Relt, Spon1, Sugp2, Tyw3
3	Cancer, Organismal injury and abnormalities, Hematological disease.	18	9	Cacnaid, Ccdc157, Kazald1, Klhl31, Pdk4, Rasl10a, Tbce, Tmprss13, Vp58

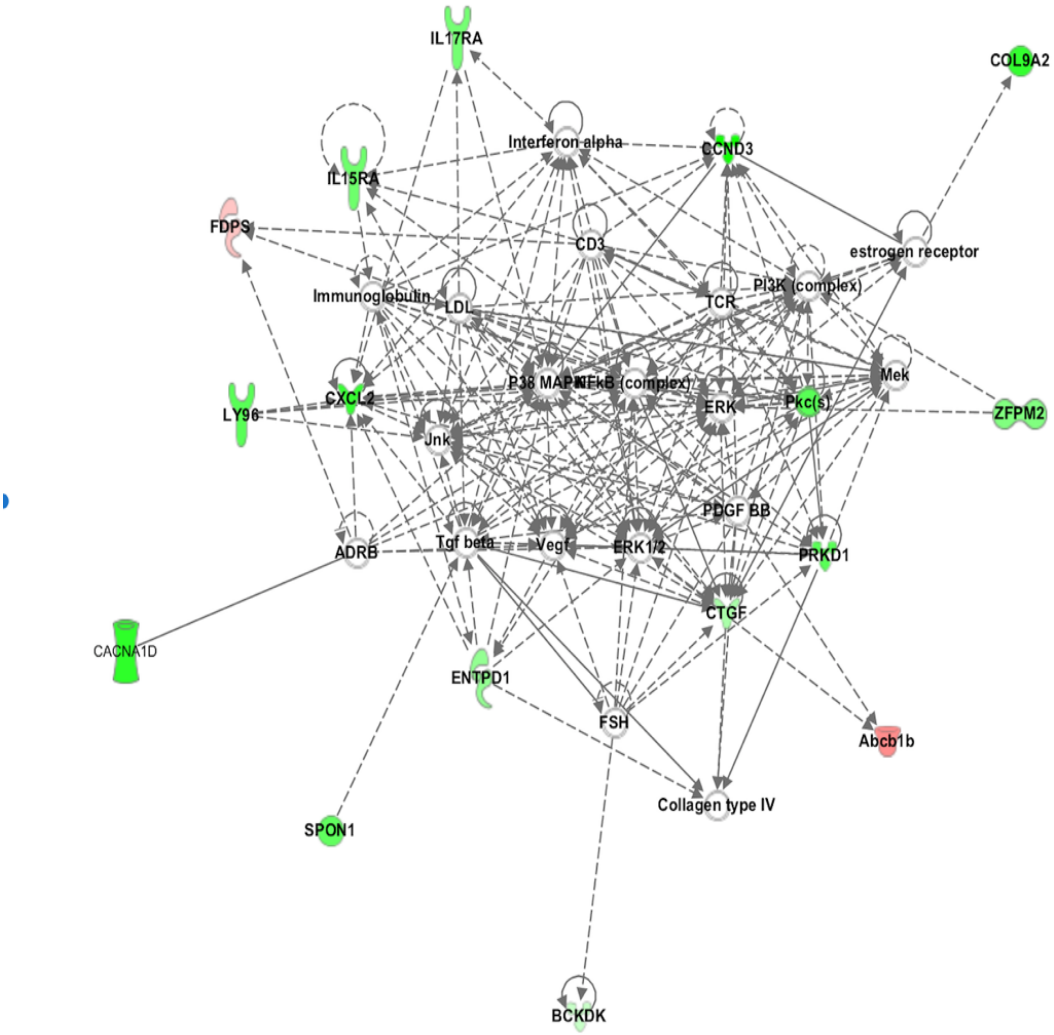


Figure 7.8 The top IPA-generated network enriched in the differentially expressed CC genes was that denoted as 'endocrine system and cell development' (enrichment score: 33). The network includes Fdps and Abcbg1 (upregulated in aged CCs) in central positions, and genes encoding proteins involved in the extracellular matrix and adhesion (a) and with proliferation and differentiation, are in the same location.

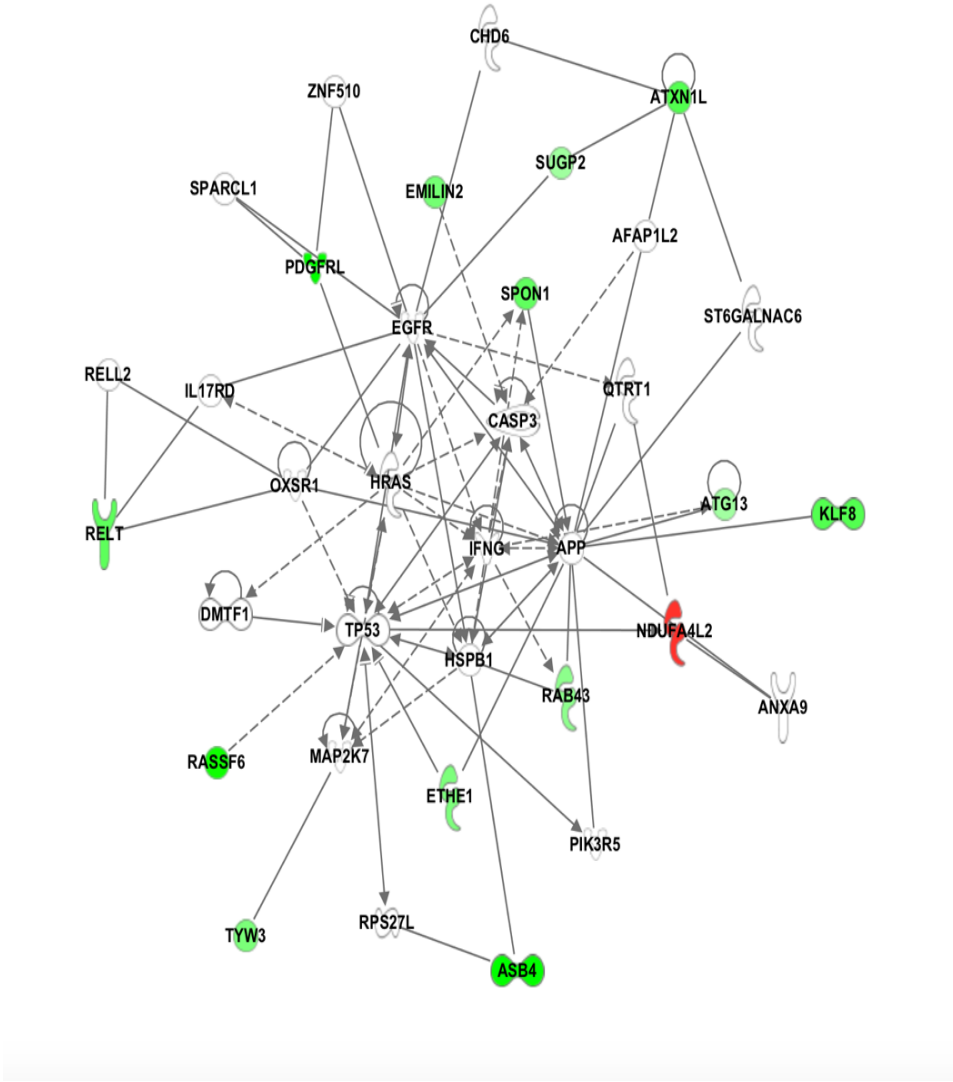


Figure 7.9 The second network affected in the aged CCs was cell death and survival, organismal injuries, and abnormalities.

7.4. Discussion

Oocyte maturation is complex event which is tightly regulated by the coordinated interaction between the developing oocyte and its companion CCs. While the oocyte is directly connected to the surrounding CCs via a network of gap junctions facilitating communication and the exchange of low molecular weight molecules, it is also influenced by the greater follicular environment, which facilitates both metabolic and oocyte meiosis regulation. The oocyte is arrested at the prophase I at by arresting factors transmitted from granulosa cells to the oocyte. Furthermore, the resumption of meiosis is mediated by transmitting a signal via the same mechanism.

This chapter aimed to define the transcriptomic characteristics of the CCs in each age group: pre-pubertal oocyte (3 weeks), young adult (9 weeks) and aged (1 year). Transcriptome profiling analysis was therefore conducted to evaluate the CCs' effect on the production of viable oocytes at each age. Thus, this investigation was intended to reveal whether intrinsic oocyte characteristics or extrinsic factors mediated by the surrounding CCs are responsible for impaired oocyte quality.

The strength of this study is that both oocytes and CCs were sampled from the same COC. It was highly important to ensure that both cell types faced similar experimental conditions and that they were at the exact same developmental stage. Since all samples were subjected to the same conditions, the differences observed across groups should reflect actual variability among them.

The variations in CC sequencing data illustrated in Figure 7.1 are partially attributed to the low number of recovered cells in some samples. This has resulted in low amount of starting RNA. The CCs were intended to be isolated from each of the oocytes and then processed very quickly to avoid RNA degradation. There was no ideal way to achieve perfectly equal CCs numbers, taking into accounting that both the CCs and oocyte had to be from the same complex and processed separately due to the technical challenges related to the sample processing time; additional time could have possibly altered gene expression.

As pre-pubertal oocytes are known to be less competent than adult oocytes, the transcriptomic profile of the pre-pubertal oocyte was investigated earlier in this study and found to be distinct from that of young-adult mice (9 weeks). The differences are potentially associated with the incomplete developmental competence of the oocyte at the pre-pubertal stage. It was therefore crucial to determine whether CCs participated in the impaired acquisition of oocyte competence at pre-puberty. It is unclear whether the CCs associated with pre-pubertal oocyte likewise demonstrate transcriptional activity that is different from that of CCs derived from young-adult mice. In short, whether the CCs and the oocytes are similar in this regard requires further research.

Tracing follicular development revealed two waves of folliculogenesis in mice: the first wave generates primordial follicles that subsequently begin to grow pre-pubertally, with the second, cortical population, supporting adult fertility (239). For that reason, Zhang *et al.* (2014) claimed that the reason that pre-pubertal oocytes are less competent is that when oocytes are collected from pre-pubertal mice, they likely derive from the first wave of primordial follicles, which may not possess the same genetic features of adult primordial

follicles. Supporting this hypothesis is a study that labelled somatic cells using autoradiographs of the embryonic gonad. Hrishfield *et al*. observed two populations of primordial follicles in post-natal rat ovaries, each with different developmental dynamics (108). More recently, Zheng *et al*. employed another tracing technique to demonstrate the existence of two waves of follicles: The first primordial follicle were activated until 45–50 days and then gradually replaced by the adult primordial follicles (109).

The RNA sequencing conducted as part of this study was able to answer this assertion at the molecular level. The CC samples from both 3-week-old and 9-week-old mice were plotted via PCA (Figure 7.4) to determine whether they exhibited similar gene expression patterns. Interestingly, PCA did not reveal any marked differences in the global gene expression of the CC samples from the two age groups. The CC samples had not been separated (Figure 7.5), demonstrating that the CCs associated with pre-pubertal and young-adult oocytes display similar transcriptional activity. Confirming this finding, when a differential expression comparison was conducted, no differential gene expression was detected between the 3-week-old CCs and the 9-week-old CCs.

Despite the 3-week-old and 9-week-old oocytes differing in terms of gene expression, no similar difference was found between the 3-week-old CCs and the 9-week-old CCs. While this finding might suggest that the 3-week-old oocyte transcriptome is independent of CCs, further evidence is required. The pre-pubertal mice were stimulated with gonadotropin (PMSG), which influences the maturation of both the oocyte and CC, which might indicate that CCs respond better to hormonal stimulation. Although with the notion of synchronised maturation between oocyte and CCs, the CCs seemingly underwent a molecular maturation

process. An *in vitro* maturation study has reported that CC proliferation was promoted in response to FSH *in vitro*, even when the CCs were separated from the oocyte. Cumulus cells further acquired the ability to undergo differentiation and to express transcripts related to expansion (57).

Thus, the similarities between the CCs of both ages could be explained by the affect of PMSG used to stimulate the mice ovaries promoting CC development during the transition from antral to pre-ovulatory follicle. In light of this explanation, a study has investigated the CC transcriptome in relation to the stages of oocyte nuclear maturation: GV, MI, and MII. That analysis found that CC transcriptomic profiles were similar during the GV and MI stages, only becoming differentially expressed in the MII oocyte (72).

On the other hand, in this study, RNA-seq has successfully identified transcriptomic changes between CCs derived from 1-year-old oocytes and 9-week-old CCs. A total of 57 genes were found to be differentially expressed in aged CCs, mainly involved in cholesterol biosynthesis, angiogenesis, cell apoptosis. Interestingly, of these 57 genes, only three genes were upregulated, which are mainly associated with mitochondrial function.

Aged human CCs have been investigated to compare the transcriptome profiles of young patients less that 30-years old and aged CCs from patients older than 37 years old. Aged CCs were found to be enriched with genes related to angiogenesis suggesting that CCs ageing is mainly associated with hypoxia(76).

The network analysis conducted via IPA revealed that endocrine function and organ development is the most biological functions affected (Table 7.3). Cumulus cells play a

central role in transmitting endocrine signals to the oocyte, and most of the genes with high folds were involved in this network.

The two key genes, farnesyl pyrophosphate (Fdps) and ATP-binding cassette transporter G1 (Abcbg1), both involved in steroid metabolism, have increased expression in aged CCs (Figure 7.7). The first of these, Fdps, is a key intermediate in cholesterol and sterol biosynthesis, a substrate for protein farnesylation and geranylgeranylation, and a ligand or agonist for certain hormone receptors and growth receptors. It has previously been identified as specifically expressed in CCs other than follicular cells (240). The elevated concentration of Fdbs indicates the increased synthesis of cholesterol in aged CCs, an outcome not normally reported in young CCs. Su *et al* (2007) has documented that increased Fdps expression is correlated with cholesterol-deficient oocytes, suggesting that cholesterol biosynthesis is promoted in CCs as a compensatory mechanism (234). That finding might suggest a similar mechanism in aged oocytes since the oocyte is reliant on CCs to provide it with cholesterol, and oocytes stimulate this activity in CCs via Bone morphogenetic protein 15 (Bmp15) and Growth differentiation factor-9 (Gdf9). Cholesterol is essential in that it supports successful cell cleavage in the developing embryo. Thus, cholesterol synthesis in CCs changes for the purpose of reducing the deleterious impact of age on lipid metabolism in aged oocytes. This alteration in cholesterol composition alters the microenvironment surrounding the oocyte and hence contributes to the deterioration of the quality of the aged oocyte.

The second key gene, Abcbg, was significantly highly expressed (FC= 4; Table 7.2) in the aged CCs. A member of the ABC transporter superfamily, Abcbg1 mediates the transport

of cholesterol, phospholipids, and other lipophilic molecules across cellular membranes and removes excess cellular cholesterol from cells. Elevated levels of Abcbg1 may potentially enhance cholesterol synthesis in CCs. The dysregulation of the cholesterol balance in the aged CCs may have altered their steroidogenesis and lipid profile.

On the other hand, the loss of Abcbg1 in mice leads to a severe developmental defect of the placenta, leading to intrauterine growth retardation and neonatal death (241). It is therefore believed that Abcbg1 is a key regulator for maintaining a normal level of cholesterol in tissue, as Abcbg1 dysregulation is accompanied by unbalanced steroidogenesis (241).

Thus, both Fdbs and Abcbg1 seem to increase the cholesterol level in aged CCs. The high expression of Fdbs and Abcbg1 in aged CCs may compensate for aged oocytes, which are deficient in ATP, as cholesterol is one of the main sources for energy production in follicles. Cholesterol and it derivatives are essential components of cell membranes and are precursors of steroid hormone synthesis. Aged CCs are characterised by an altered cholesterol composition, which weakens their functioning.

On the other hand, transcripts related to angiogenesis were markedly decreased in the aged CCs. Among those transcripts was connective tissue growth factor (Ctgf), which has diverse cellular functions (Table 7.2). Its primary role is in normal follicular development, and particularly granulosa cell differentiation and proliferation (242). It was also reported to contribute to oestrogen formation in rat granulosa cells (243). Therefore, the very low level of Ctgf in the aged CCs might have induced apoptosis.

A knockout Ctgf in mouse ovary has significantly resulted in granulosa cell apoptosis, which leads to less pre-antral and antral follicles and more atretic pre-antral and antral follicles than seen in control ovaries (244). Increased CC apoptosis was also reported in oocytes from aged mice (245). Since Ctgf is an important paracrine factor involved in the continuous remodelling of the extracellular matrix supporting granulosa cell development, it is also considered an angiogenic factor responsible for providing oxygen and nutrients in later stages of follicle development (246). It could therefore be indicating that low Ctgf levels initiate hypoxia in aged CCs. In agreement with this finding, a recent CCs transcriptome study revealed that the link between Ctgf and hypoxia is a major hallmark of CC senescence (77).

Moreover, cytokines and pro-inflammatory molecules, which are important for CCs expansion and ovulation, also markedly decreased in the aged CCs. These included receptors of interleukins 5 and 7 (l17ra, Il15a) and chemokine (C-X-C motif) ligand 1. Cumulus cells expansion is based on the synthesis of a large quantity of extracellular matrix occurring in the enlargement of COCs. It takes place in the follicle shortly before ovulation in vivo, as well as during meiotic maturation *in vitro*. Interleukins 7 is believed to be involved in steroidogenesis, the ovulation process, and oocyte maturation (247). Substantially high II7r expression levels in CCs are known to be a marker for CCs expansion and oocyte maturation.(248) Therefore, it is evident that downregulating cytokines has an adverse impact on CC expansion, which is essential for oocyte maturation, fertilisation, and embryonic development (65). Evidently, decreased cytokine activity alters the normal CC expansion process.

Furthermore, of those genes involved in the endocrine system and organ development network, protein kinase C (Pkc), which has a key role in the regulation of gap junctions, was decreased seven fold in the aged CCs. Protein kinase C plays a role in phosphorylating connexin 43, which control intracellular communication by opening and closing gap junction channels and facilitating ion and growth factor transport within ovarian follicles (249). Since normal Pkc levels are required for connexin 43 activity, the diminished level of Pkc in aged CCs may reduce the functionality of the gap junctions, ultimately affecting oocyte development.

The only gene with very high expression was involved in the second network (cell death and survival, organism injuries, and abnormalities; Table7.3) was NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 4-like 2 (Ndufa4), A subunit of NADH dehydrogenase (ubiquinone), Ndufa4 is located in the mitochondrial inner membrane and is the largest of the five complexes of the electron transport chain. Steuerwald *et al*. have demonstrated that Ndufa4l2 levels are 1.94 times higher in aged human oocytes than in young human oocytes (53). The present study has found that this gene's expression was eightfold higher in the aged CCs than in the young CCs. It is believed to act as an oxidative stress mediator and is reportedly highly expressed in placental tissue in pre-eclampsia in pregnant women (250). Moreover, Ndufa4l2 is expressed in response to hypoxia. The Ndufa4l2 activation mechanism under hypoxic conditions is as follows: Cells activate an anaerobic switch that favours glycolysis and that attenuates mitochondrial activity. Thus, Ndufa4l2 helps to reduce mitochondrial oxygen consumption and activity, which in turn initiates a shift from mitochondrial respiration to anaerobic glycolysis. Since the glycolytic process is alternatively used for sourcing ATP (251), a CC transcriptome study reported the enrichment

of transcripts involved in hypoxia-related stress, including Ndufa4l2, in aged human CCs (252).

Interestingly, a study has revealed that treating renal cells with Ndufa4l2 induces changes in TCA cycle metabolism, suggesting that mitochondrial bioenergetics are improved when Ndufa4l2 levels are reduced (253). When examined alongside the altered lipid metabolism in the 1-year-old oocytes discussed in Chapter 6, this finding seemingly suggests that the altered metabolism of aged CCs is linked to the aged oocyte's altered lipid profile and energy production.

Since Ndufa4L2 was the only gene with very high expression in the aged CCs, CCs hypoxia is potentially one of the main characteristics of aged CCs contributing to poor oocyte quality. The tight regulation of mitochondrial function is an essential metabolic adaptation to fluctuations in oxygen availability.

The detailed CC transcriptome revealed that the CCs' biological functioning was compromised and implicated in oocyte senescence, suggesting a link between oocyte quality and CC ageing since oocyte development requires a tightly controlled endocrine, paracrine, and autocrine interplay among the oocyte itself and the CCs. The poor quality of the aged oocyte is likely influenced by the altered CC activity, as CCs with compromised functioning impair the endocrine signals required for successful oocyte maturation and further embryo development.

This study has demonstrated the dysregulation of key factors in cholesterol biosynthesis; in aged CCs, CC expansion and gap junctions were associated with impaired CC functioning.

Moreover, transcripts related to hypoxia were confirmed to be a key characteristic of aged CCs. Given that both oocytes and CCs were retrieved from the same COCs, this study reveals the molecular mechanism underlying oocyte competence, confirming that an altered microenvironment is a major contributor to the poor quality of the aged oocyte.

7.5. Conclusion

Cumulus cells play an essential role in the maturation and competence acquisition of the developing oocyte. The fact that these cells are closely associated with the oocyte, share the same microenvironment, and can be easily collected during IVF procedures makes them attractive targets for basic research and the development of clinically relevant assays. Analyses of follicular cells are likely to reveal important information concerning the viability and genetic constitution of their associated oocyte; such studies will likely also shed light on normal follicular processes and the impact of disorders or medical interventions.

Chapter 8

OpenArray Real-Time PCR Validation

8.1. Introduction

Quantitative, real-time polymerase chain reaction (RT-qPCR) has been the gold standard method in quantitation transcript abundance since its advent (254). Real-time PCR provides the opportunity to quantify mRNA in small samples, such as those from oocytes and pre-implantation embryos at the picogram level (255).

Real-time PCR has better accuracy, a wider dynamic range and higher sensitivity than other transcriptomic methods, such as microarrays, do. Therefore, RT-PCR is known to be a superior, alternative method for the calibration and validation of microarrays (256). Most studies have used RT-PCR to validate only a small number of genes; screening a list of genes using 96 or 384 wells is somewhat undesirable due to its time-consuming nature and the need for a large quantity of material and intensive labour (257). The development of OpenArray qPCR offers the ability to investigate a panel of genes in a nanolitre volume of sample. Moreover, OpenArray qPCR has the potential to provide gene multiplexing, allowing several thousand PCRs to be performed in parallel on one platform (258–260).

The most important element that must be accounted for when performing qPCR on immature oocytes is that the RNA content in immature oocytes (germinal vesicle [GV] stage) are susceptible to transcriptional changes during the maturation process (261,262). This is because oocyte maturation (the transition from GV to meiosis [MII]) undergoes selective degradation of the transcripts (197,263). Thus, genes nominated for use as reference genes could be a part of the regulation of target genes. As result, this increases the concern about the accuracy of gene expression measurement .

Previous transcriptomic studies relied on qPCR for validating the gene expression generated

by microarrays or RNA sequencing (RNA-seq) studies. Although oocyte and cumulus cells CCs studies have found relative agreement between the qPCR and microarray results, it cannot be ascertained whether both methods exhibit a similar concordance (72,130). Because the genes chosen for validation are always standard genes, not novel ones, qPCR becomes a routine procedure, without a real consideration of using a second method to confirm the data it generates. This chapter begins by explaining the steps of the optimisation preparation approach for single oocytes and the surrounding CCs for the use of OpenArray qPCR. This is achieved by conducting four pilot studies before the experiment is performed.

The overall aim of this chapter is to employ an OpenArray RT-PCR for validating the RNA-seq results of the following groups:

- Pre-pubertal oocytes relative to young adult oocytes (3-weeks-old vs. 9-weeks-old);
- Aged oocytes relative to young adult oocytes (1-year-old vs. 9-weeks-old); and
- Aged CCs relative to young CCs (1-year-old vs. 9 weeks old).

To achieve the main aim, this chapter is structured to fulfil the requirements described below.

A. Section one: Optimising a preparation protocol

The optimisation of a preparation protocol underwent many experiments to develop the best approach for single oocytes. The main requirement was that the approach would enable investigation of a panel of genes in an individual oocyte and its CCs. Three sequential pilot experiments were conducted to develop a reliable preparation approach, ensuring the validity of each step of the approach. The method of each pilot experiment is explained

below, followed by its result. Because the protocol was being developed, each small study relied on results from its predecessor until the final preparation protocol was determined. The development of this protocol included the following:

- Pilot 1: Assessment of RNA recovered from a single oocyte and its CCs;
- Pilot 2: RT-PCR of the cumulus–oocyte complex (COC); and
- Pilot 3: RT-PCR of a single oocyte and CCs, including a preamplification step.

B. Section two: OpenArray real-time PCR (Section 8.3).

The actual oocyte and CC samples were run according to the optimised protocol developed in section one. However, an experimental OpenArray qPCR plate was run using four samples prior to the actual experiment (as explained in section 8.3.2).

8.2. Optimising a preparation protocol for RNA recovery from a single oocyte and its CCs.

8.2.1. Pilot 1: Assessment of RNA recovered from a single oocyte and its CCs

The first fundamental step when optimising an RNA extraction protocol is to ensure that the RNA recovery from a single cell is sufficient for the requirements of the gene expression experiment. Existing methods are unable to measure the RNA recovered from a single oocyte as the RNA.

8.2.1.1. Aim

A single-gene RT-qPCR experiment was run to ensure that the RNA was intact and evaluate its compatibility with downstream reverse transcription and subsequent PCR using two genes, as follows: Glyceraldehyde 3-phosphate dehydrogenase (Gapdh), which is known to be a reference gene in many gene expression studies (264,265), was used as a housekeeping gene, while Cytochrome P450 family 11 subfamily A member 1(Cyp11a1) was used as the gene of interest; Cyp11a1 was one of the differentially expressed genes in both the oocyte and CC samples generated in RNA-seq data. Therefore, it was selected to be run in parallel with Gapdh.

8.2.1.2. RNA recovery

The TaqMan Gene Expression Cell-to-Ct kit (Cat. No. 4399002, Ambion by Life Technologies, UK) was used for the RNA extraction. A single COC was retrieved from a mouse ovary, washed in bovine serum albumin (BSA)/phosphate-buffered saline (PBS; 1 mg/ml) and separated into the oocyte and CCs by mechanical pipetting in and out. Both the oocyte and CCs were immediately immersed in 50 µlX lysis solution (1:100, DNase: lysis buffer) and mixed five times up and down to ensure complete mixing with the lysis buffer. The lysates were then incubated for 5 minutes at room temperature (25°C). After incubation, 5 µlX Stop solution was added to each lysate tube and mixed five times.

8.2.1.3. Reverse transcription

Samples were reverse transcribed to cDNA. To ensure recovery of the maximum RNA material from a single oocyte, the protocol provided by the manufacturer's instructions was modified. However, it was not clear which volume would provide the best results, and thus, samples were divided into two groups, where each sample group had a different volume of lysates to determine the best volume for sufficient RNA amount from a single oocyte.

In the first group, each sample contained 12.5 µl of lysate sample added to 25 µl of 2X reverse transcription (RT) buffer, 2.5 µlX Gapdh TaqMan® Gene Expression Assays (Cat. no. 4453320, ThermoFisher, UK) and 10 of µl distilled water to make up 50 µl of reaction volume. Samples in the second group contained 22.5 µl of lysate sample added to 25 µl and 2.5 µl of enzyme (no water was added). Two similar groups were prepared in parallel using Cyp11a1 TaqMan® Gene Expression Assays (Cat no., 4453320, ThermoFisher) instead of Gapdh.

Table 8.1: Reverse transcription reagents

Component	**First group**	**Second group**
2X Reverse Transcription (RT) Buffer	25 µl	25 µl
20X RT enzyme mix	2.5 µl	2.5 µl
Nuclease-free water	10 µl	-
cDNA sample	12.5 µl	22.5 µl
Final volume RT Master Mix	50 µl	50 µl

8.2.1.4. Reverse transcription programme

Reactions were run in the thermocycler as follows: Samples were held at 37°C for 60 minutes, followed by incubation for 5 minutes at 95°C, and then held at 4 °C. Each sample was run in three technical replicates. A negative control, containing all the RT reaction components except the RT enzymes, was also run.

8.2.1.5. Real-time PCR

Following cDNA synthesis, the cDNA samples were prepared for qPCR reactions. A total of 50 µl of PCR reaction was prepared, as indicated in Table 8.2.

Table 8.2: Reagents for PCR mix reaction

Component	**Volume**
TaqMan Gene Expression Master Mix	25 µl
TaqMan Gene Expression Assay (20X)	2.5 µl
Nuclease-free water	12.5 µl
cDNA sample	12.5 µl

8.2.1.6. Pilot 1 results

Unfortunately, the amplification of Gapdh and Cyp11a1 presented in Figure 8.1 showed an unsuccessful detection of the gene expression in both the oocyte and CC samples (the amplification curves were comparable to the minus RT samples), demonstrating the failure of the PCR amplification. Although five additional amplification cycles were added (5 cycles added to 40 PCR cycles), the samples could not be amplified. Amplification curves are created when the fluorescent signal from the cDNA sample is plotted against the cycle number. Therefore, an amplification curve represents the accumulation of the product over the duration of the RT- PCR experiment. The two negative controls, namely minus RT and non-template control, were both run with oocyte and CC samples to ensure that the PCR reaction was contamination free. The amplification failure was the result of two possible scenarios; either the template or the primers contributed to the failure of this amplification. To eliminate the primer as the reason for amplification failure, the experiment was repeated using other assays. In these tests, the amplification continued to fail.

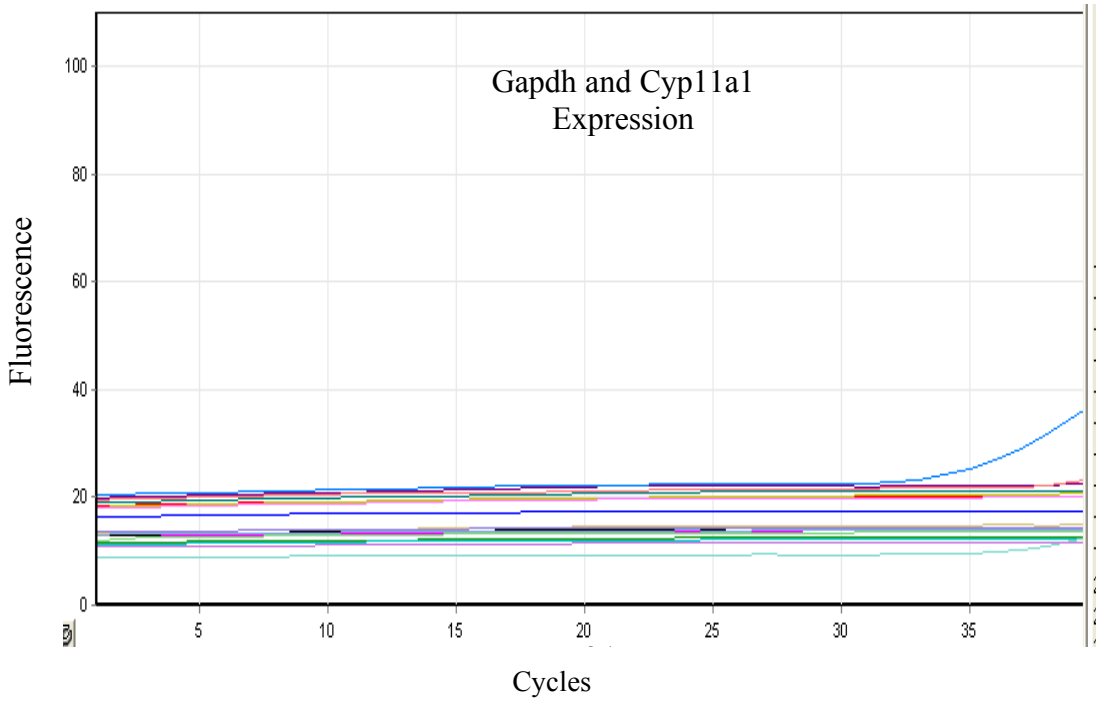


Figure 8.4 First pilot: Gapdh and Cyp11a1 expressions in single oocytes and CCs.

Both genes have not been detected due to failed amplification of all samples: single oocyte and cumulus cells samples (samples were run in duplicate). Amplification plots are created when the fluorescent signal from each sample is plotted against cycle number; therefore, the amplification plots represent the accumulation of fluorescent product. Relative fluorescence is given on the y axis vs. cycle number on the x axis.

8.2.2. Pilot 2: Real-time PCR of COC

8.2.2.1. Aim

Due to the amplification failure of the single oocytes and CCs in pilot 1, certain factors needed to be addressed to determine the reasons for this failure. A low starting quantity of RNA material was suspected to be a primary reason for the amplification failure. Therefore, pilot 2 was designed to determine whether using a large volume of RNA would improve the amplification.

8.2.2.2. Method

To obtain a large amount of RNA, COCs were collected for the analysis rather than single cells. Each five COCs were placed in a 50-ul lysis buffer tube. The samples were prepared reverse transcribed using a similar method to that explained in section 8.2.1.

8.2.2.3. Pilot 2 results

Gapdh and Cyp11a1 expressions were detected by real-time qPCR, as shown in Figure 8.2. The threshold of the real-time PCR reaction was the level of signal reflecting a significant increase over the calculated baseline signal (Ct value, the default was set up). The main purpose of this pilot experiment was only to verify whether the detection of Gapdh expression and Cyp11a1 level was achievable using this preparation approach. The results confirmed that limited RNA material was the main reason for the RT-qPCR failure in pilot 1. Therefore, RNA must be concentrated when isolating it from a single oocyte and the

surrounding CCs. Thus, pilot 3 was designed to address this issue, as described in the next section.

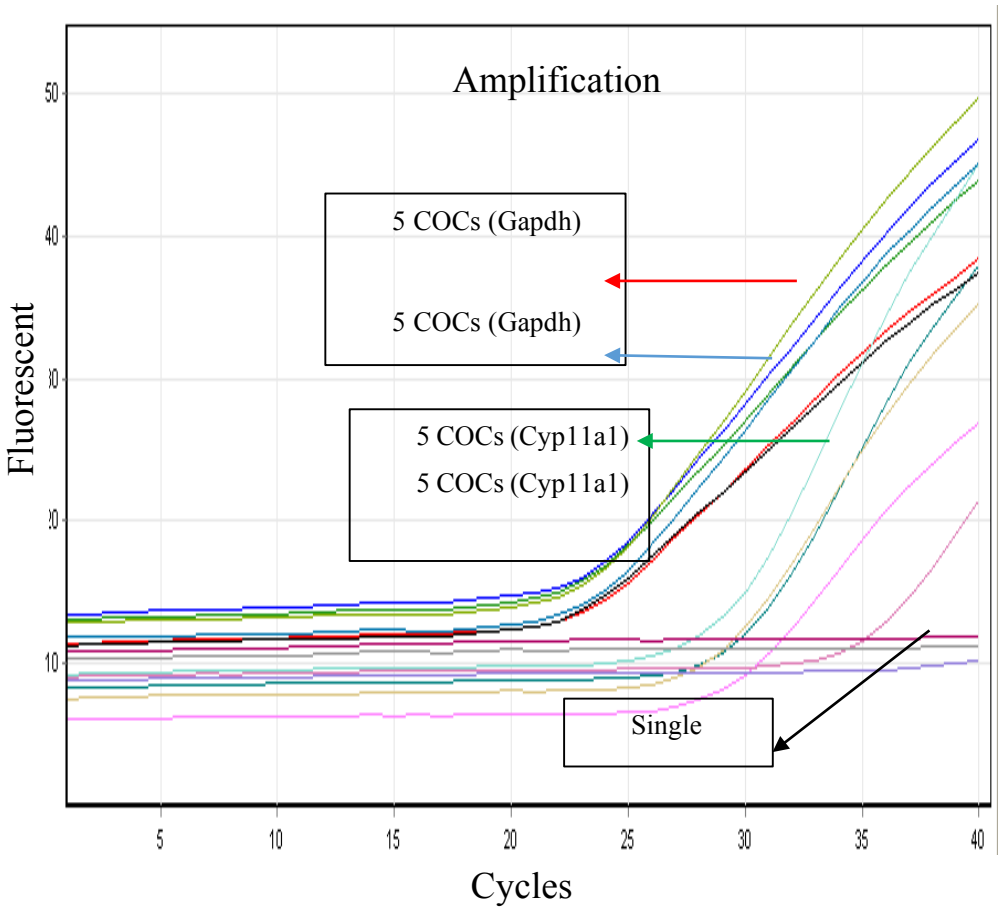


Figure 8.5 Pilot 2: Gapdh and Cyp11a1 expression detected by RT-PCR in mouse COC samples.

Each amplification curve represents one sample (each set of two identical curves demonstrates the duplicates). Two negative controls are included (straight curves): The first is the negative control without RT enzyme, and the second is the no template control; they were run together with the samples. The negative control showed no amplification.

8.2.3. Pilot 3: Gapdh and Cyp11a1 expression in a single oocyte and its CCs

8.2.3.1. Aim

Based on the findings of pilots 1 and 2, it was concluded that the PCR product could not be detected in a single oocyte (pilot 1) due to dilution of the RNA recovered from this oocyte and its CCs. This means that the volumes used for preparing the lysates were unsuitable for the RNA quantity derived from small cells. This assertion was confirmed when a concentrated sample of RNA was used in a sample of five COCs in pilot 2, as this enabled the detection of both genes of interest. To increase the quantity of starting material to allow for the detection of genes, it was necessary to increase the cDNA yield generated from a single oocyte and its CCs. The aim of pilot 3 was to determine whether Gapdh and Cyp11a1 could be detected in a single oocyte and its CCs using small reagent volumes.

8.2.3.2. Method

The reagents used in the previous two pilots were scaled down to a quarter of their volume. In addition, a preamplification step was implemented in the preparation protocol just prior to the qPCR. Figure 8.3 demonstrates the workflow for pilot 3. Each single oocyte was washed in BSA/PBS (1 mg/ml) and immediately immersed in 10 µl of lysis solution (DNAase: lysis buffer, 1:100 volume) and mixed five times up and down. The sample was then incubated for 5 minutes at room temperature (19–25°C). After incubation, 1 µlX of

stop solution was added to the lysate and mixed five times. Using this method, the oocyte's or CCs' RNA should have been concentrated in small volumes. Then, the samples were reverse transcribed in a similar method as described in section 8.2.1.4. Next, the samples were prepared for preamplification as described below.

8.2.3.3. Preamplification step

Equal volumes of each 20X TaqMan Gene expression assay (Gapdh and Cyp11a1) were pooled together and diluted with Tris base, acetic acid and EDTA (TAE) buffer, reaching a final concentration of 0.2X for each assay. For instance, 1 µl of Gapdh and 1 µl of Cyp11a1 were added to the TAE buffer. Then, approximately 12.5 µl of pooled assay mix (0.2X, each assay) was added to 25 µl of TaqMan preamplification master mix (2X; Cat. no. 4488593, ThermoFisher, UK), and 12 µl of nuclease-free water was added to constitute a total final volume of 50 µl. The sample mixtures were centrifuged. A preamplification reaction was run in a thermocycler, programmed as follows: samples were held for 10 minutes at 95°C, followed by 14 cycles of 95°C for 15 seconds and 60°C for 4 minutes (Table 8.3).

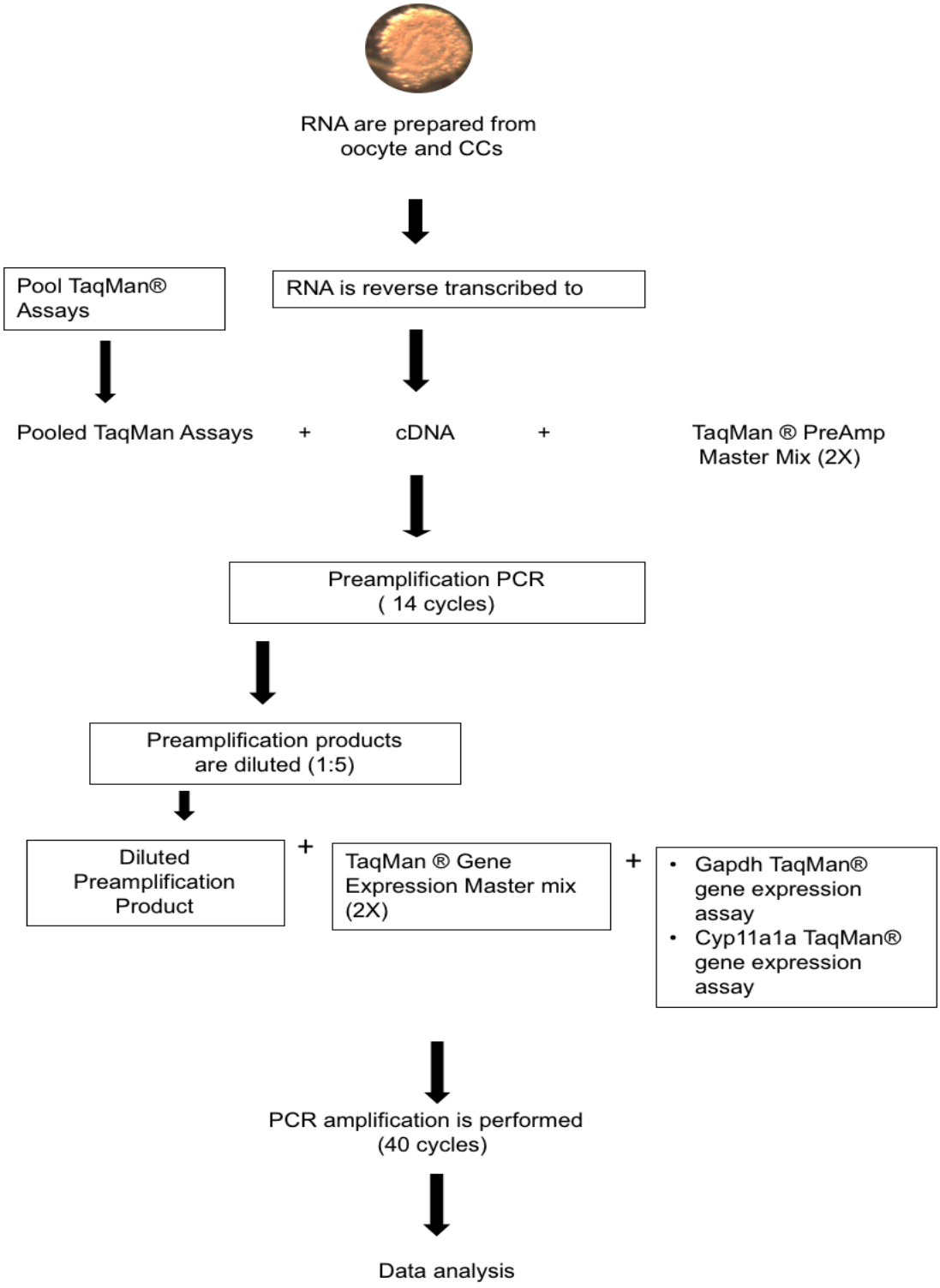


Figure 8.6 Workflow of the RT-qPCR, including the sample preparation, reverse transcription, preamplification and qPCR reaction.

Table 8.3: Reagents used for the preamplification mixture

Reagent	Each reaction
TaqMan Gene Expression Assay (20)	2.5 µl
Preamplified cDNA products (diluted to 1:5)	12.5 µl
TaqMan Gene Expression Master Mix (2X)	25 µl
Nuclease-free water	10 µl
Final volume RT Master Mix	50 µl

8.2.3.4. Polymerase chain reaction

The preamplification product was diluted to a ratio of 1:5 with the TAE buffer. The PCR reactions were prepared in a similar manner to a previous experiment (Table 8.2). The samples were run in RT- PCR instrument according to the following programme: samples were held at 50°C for 2 minutes and 95°C for 10 minutes, followed by 40 cycles consisting of 15 seconds at 95°C and 4 minutes at 60°C. The experiments were repeated twice to confirm the results.

8.2.3.5. Pilot 3 results

Successful amplification of the single oocyte and CCs is indicated in Figure 8.4. The Gapdh and Cyp11a1 expression were detected in all samples, as follows: (1) single oocyte 1, (2) single oocyte 2, (3) companion CCs 1, (4) companion CCs 2, (5) positive control–COC, (6) negative control (red line).

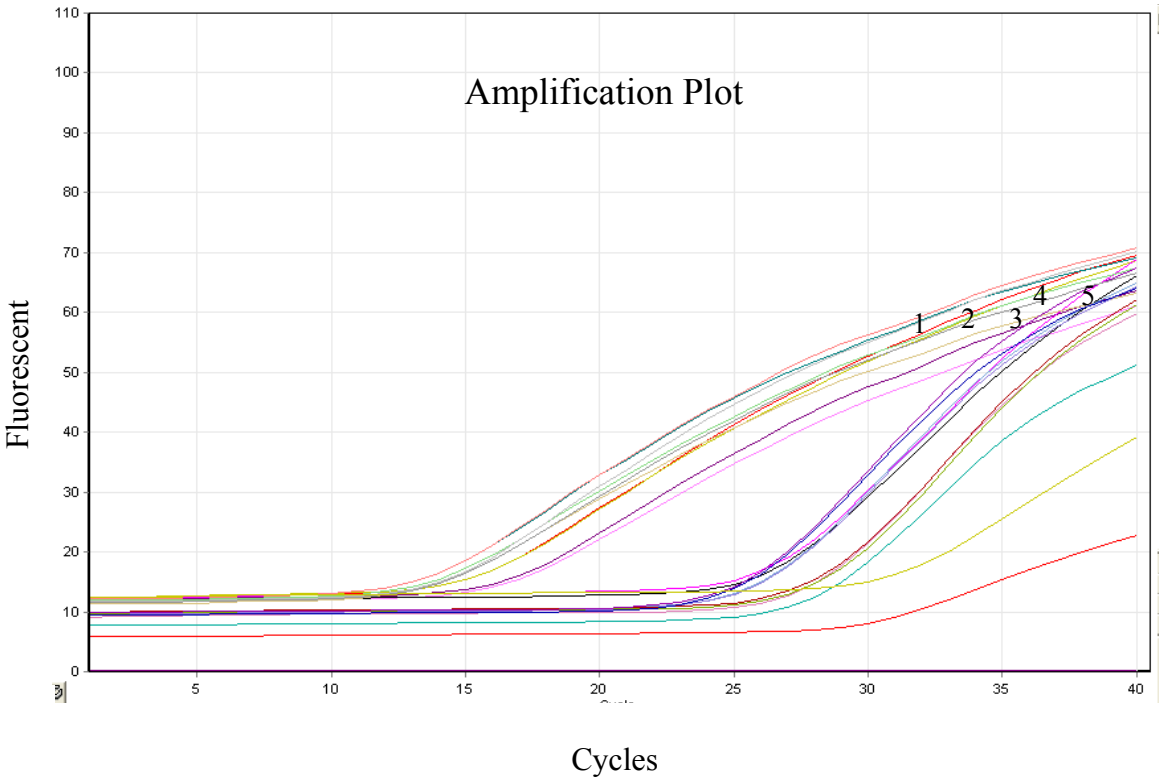


Figure 8.7 Pilot 3: Successful RT-PCR showing the Gapdh and Cyp11a1 transcripts' expression in the samples: (1) oocyte 1, (2) oocyte 2, (3) CCs 1, (4) CCs 2, (5) positive control–COC, (6) negative control (red line).

8.3. OpenArray RT-PCR

Oocyte and CCs samples were collected based on the optimised preparation method determined in pilot 3, as described below.

8.3.1. Assay panel selection

8.3.1.1. Technical criteria

Each TaqMan® gene expression assay has a unique assay ID by which the assay placement can be assigned. The selected assay IDs were those with probes spanning the exon–exon junction of the associated genes that do not detect genomic DNA (this is suffixed by _m1, which is mouse, *Mus musculus*, abbreviated 'Mm') and have an amplicon size of 50–100 bp, as the shorter amplicon size allows better amplification efficiency. It is worth mentioning that TaqMan assays consist of a pair of unlabelled PCR primers and a TaqMan probe with a dye label (FAM). The selected 224 assays and their corresponding IDs (e.g., assay ID for mouse Gapdh: Mm00484668_m1) were organised in OpenArray Excel format (Figure 8.7) and ordered from ThermoFisher manufactory. OpenArray qPCR plates (Cat. no. 4471128) were configured with the 224 assays selected from the sequencing data and set according to the selection criteria described below.

8.3.1.2. Gene selection criteria (selected genes from the RNA-seq results)

Each list of differential genes expressed in each age group was sorted based on the highest fold changes in both upregulated and downregulated genes with a set fold change of > 2 or

< –2. Genes were sorted in ascending order based on the lowest false discovery rate (FDR). Although the initial plan was to select equal numbers of genes from each group, some assays did not meet the technical criteria stated in the guidelines from ThermoFisher manufactory mentioned in section 8.3.1.1.

The gene panel consisted of differentially expressed genes from the following groups:

- **Group A:** 3-week-old vs. 9-week-old oocytes:
    - Upregulated genes;
    - Downregulated genes;
- **Group B:** 1-year-old vs. 9-week-old oocytes:
    - Upregulated genes;
    - Downregulated genes;
- **Group C:** 1-year-old vs. 9-week-old CCs;
- **Group D**: Oocyte-specific genes; and
- **Group E:** CC-specific genes.

These genes were organised as described below.

A. Oocyte-expressed genes (130 assays)

There were 130 gene assays imported from the oocytes' RNA sequencing data. Of these, there were 21 upregulated genes in pre-pubertal oocytes, 20 upregulated genes in aged oocytes and 39 genes overlapping for both ages. Regarding the downregulated genes, 11 downregulated genes overlapped for both ages, 21 genes in pre-pubertal oocytes and 18 genes only in aged oocytes.

B. Cumulus cell–expressed genes (75 assays)

Seventy-five genes were selected from the CCs' RNA-seq data. Twenty genes, which were the only differentially expressed genes between aged and young adult CCs, were selected for validation, whereas the remaining genes were selected from the genes expressed in 9-week-old CCs.

C. Oocyte-specific genes

Thirteen genes were chosen from the oocyte-specific genes. These were used to test the sequencing data quality (Section 5.3.1.2).

D. Cumulus cell–specific genes

Thirteen genes were chosen from the CC-specific genes, which were previously used to evaluate the CCs sequencing data (Section 7.3.1.2).

E. Housekeeping genes

Five genes were selected from the literature that are known to be expressed at constant levels. These were as follows: Gapdh, which is a housekeeping gene for both oocytes and CCs; hypoxanthine guanine phosphoribosyltransferase 1 (Hprt1), a housekeeping gene for oocytes; ß-actin (Actb); ribosomal protein L19 (Rpl19); and ß-tubulin (Tubb), which is a microtubule component.

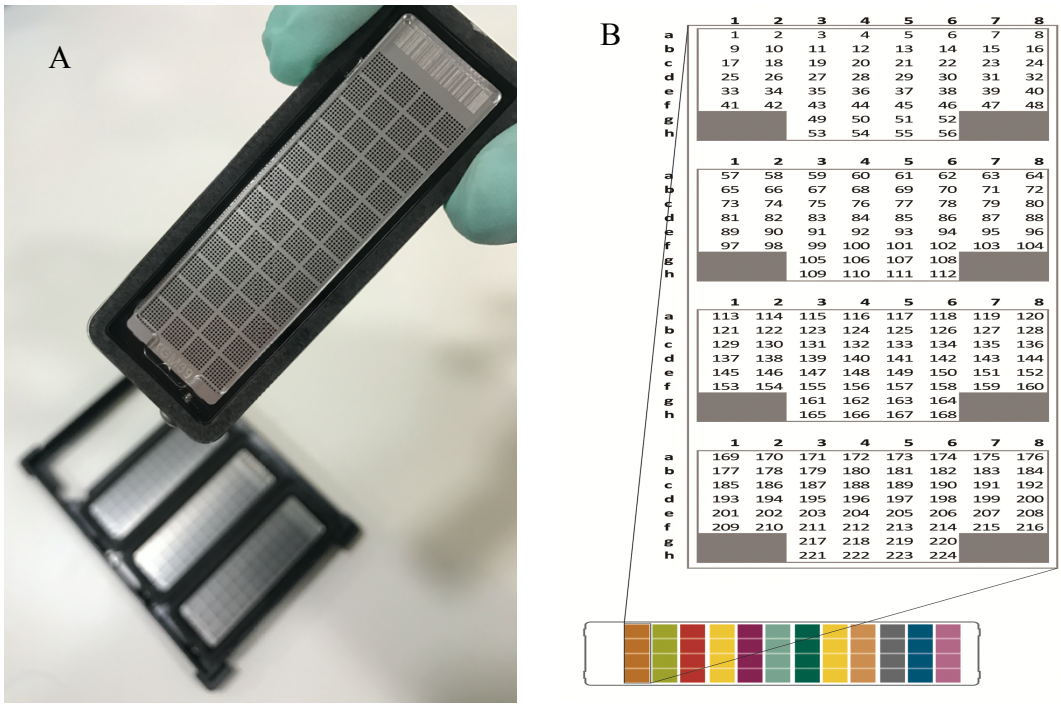


Figure 8.8 A) Photograph of a stainless-steel plate of OpenArray qPCR, the size of a microscope slide, taken during the experiment. Wells are grouped in 48 arrays of 64 holes (33 nl reaction volume), in which the primer pairs are immobilised. The array loaded with primers is stored in an evacuated Mylar™ bag at –20°C. All four plates can be run at once to reduce experiment variation. B) Schematic illustration of the holes showing 224 assays organised in four subarrays (squares).

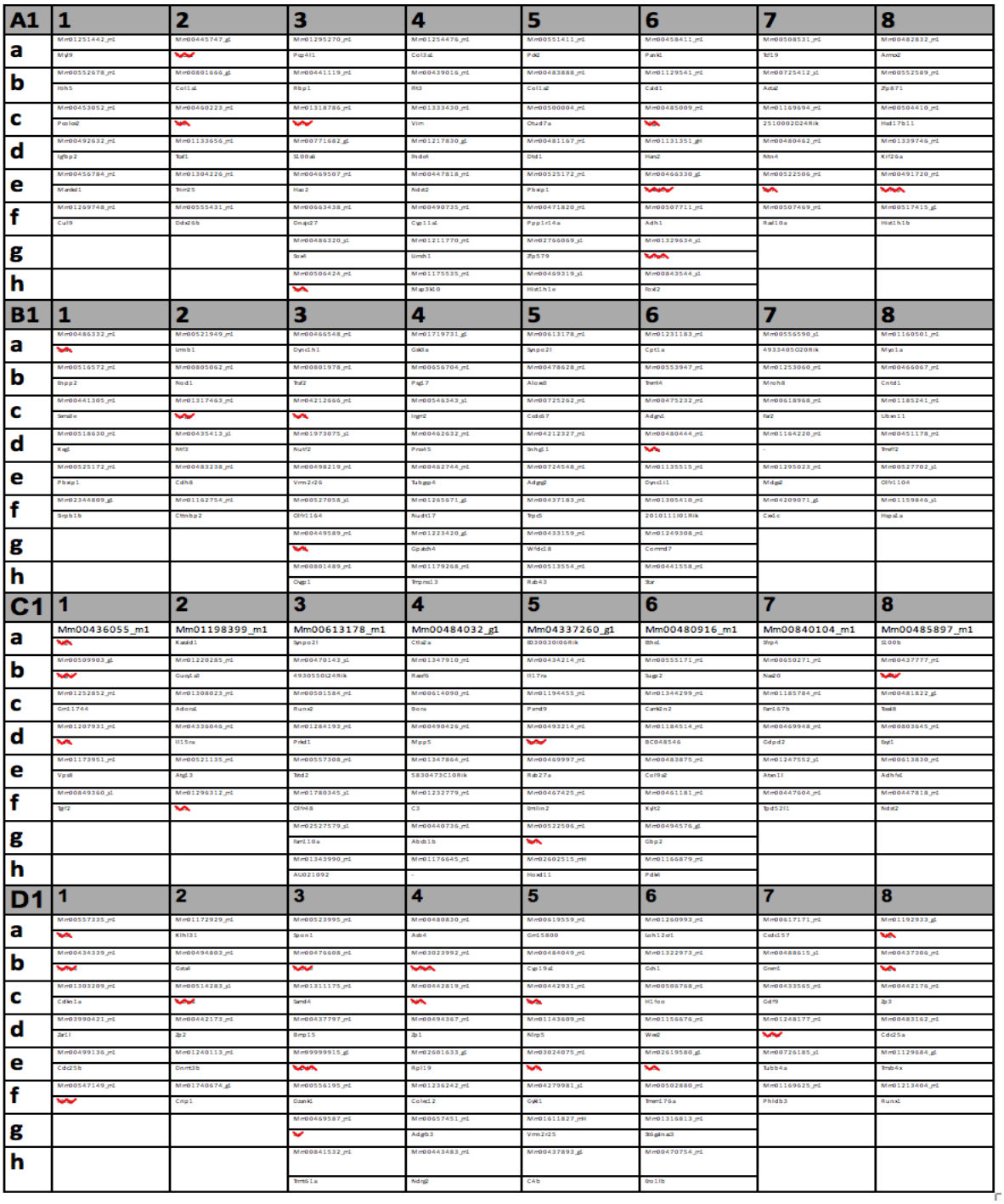


Figure 8.6 Illustration of the first column in an OpenArray plate, which is divided into four sub-arrays (A, B, C, D) horizontally and eight small columns vertically. Each sample was dispensed in one column containing the primers of 224 assays (ThermoFisher, UK).

8.3.2. Experimental design

Six biological replicates were collected from each age group and each cell type, that is, six single oocytes, six 9-week-old oocytes and six 1-year-old oocytes, as well as six CC samples from each group. In total, there were 18 oocytes and 18 CCs in the total number of biological samples. Regarding the technical replicates, each cDNA sample was run in three replicates. The total number of oocyte samples, including the technical replicates, was 54 samples, which was the same for the CC samples. In total, 108 samples were run in nine OpenArray plates.

8.3.2.1. Mice

Three female 3-week-old and three female 9-week-old C57BL/6 mice were purchased from BMS (South Park, University of Oxford). Five female 1-year-old C57BL/6 mice were purchased from Envigo, Netherlands. The mice were sacrificed by cervical dislocation and RNAase zap was used to clean the mice to ensure an RNA-free environment. The mice were dissected, and their ovaries were retrieved and transferred into tubes containing BSA/PBS.

8.3.2.2. Sample collection

Six single oocytes and their six surrounding CCs were collected from the 3-week-old, 9-week-old and 1-year-old mice and kept separate. In total, 21 single oocyte samples (six single oocytes with their six companion CCs from each age group, plus extra samples) were

selected for use in the validation experiment. Extra samples were collected to allow for replacements if needed.

8.3.2.3. Sample preparation

RNA was prepared using the TaqMan Gene Expression Cell-to-CT kit (Cat. no. 439902, Ampion, Life Technologies) as described in pilot 3 (section 8.2.3). Sample collection was performed on one mouse per day to avoid late preparations and lengthy collection procedures, which have adverse effects on RNA integrity. Samples were kept frozen at –20ºC until all collections were complete and ready for the RT reaction.

8.3.3. Preamplification

To prepare the preamplification mix, approximately 4.2 µl of the preamplification assay's pooled preamplification master mix (Cat. no. 4462159, ThermoFisher) was added to 8.3 µl of preamplification master mix (Cat. no. 44262159, ThermoFisher). Then, 4.2 µl of cDNA sample was added to the mixture, resulting in a final volume of 16.7 µl. The samples were then run in a thermocycler as follows: after 10 minutes of incubation at 95°C, preamplification was carried out for 14 cycles of amplification at 95°C for 15 seconds and 60°C for 4 minutes (Table 8.4).

Table 8.4: Reagents for pre-amplification reaction for the OpenArray qPCR

Component	**PCRs in each reaction**
Preamplification master mix	8.3 µl
Preamplification assays pool (224 assays)	4.2 µl
cDNA sample	4.2 µl
Total	16.7µl

8.3.4. Real-time PCR preparation

Preamplified cDNA was added to the TaqMan™ Real-Time *Master* mix (Table 8.5).

Table 8.5: Reagents used for OpenArray qPCR

Component	**Each reaction**
TaqMan™ Real-Time *Master* mix	16.7 µl
Preamplified cDNA	5.2 µl
Total	21.9 µl

8.3.5. Real-time PCR workflow

The protocol was set definitively in pilot 3, as reflected in the workflow shown in Figure 8.4. After the preamplification step, the preamplified samples were diluted before running them for real-time qPCR. However, this dilution was suitable for single-tube qPCR, as

conducted in pilot 3. As for the OpenArray plate, the volumes used for preparing the PCR reaction (Table 8.5) were different from those of the single-tube qPCR (Table 8.3); because the OpenArray plate is designed for nanolitre volumes, it was unclear how much dilution of cDNA was required. Due to this uncertainty, one of the OpenArray plates was sacrificed in an experimental run in which a series of diluted preamplified cDNA samples were examined before conducting the actual validation experiment.

8.3.6. Preliminary experiment

8.3.6.1. Aim

Although pilot 3 was successful in showing Gapdh and Cyp11a1 transcripts detected in single oocytes, it was crucial to translate the preparation approach into small volumes for the OpenArray qPCR. The aim of this experiment was to determine the amount of dilution required for the preamplification.

8.3.6.2. Method

Samples were prepared as indicated in section 8.2.2.4. Then, the preamplification step was performed as described in section 8.3.3. To determine how much dilution was needed for the RT-PCR reaction, three dilutions were applied, as follows: 1:5, 1:10 and 1:15. The PCR product was diluted with nuclease-free distilled water.

8.3.6.3. Results

The results showed that no significant difference in the Ct values was generated by the three dilutions. This suggests that the preamplification product may be used for real-time preparation without prior dilution.

8.4. OpenArray RT-PCR experiment

Measuring the RNA quantity from a single oocyte assigned for use in the subsequent OpenArray qPCR experiment was found to be infeasible

Each oocyte sample (cDNA sample) was run in three technical replicates. The quantities required to run triplicates were calculated and prepared together to avoid the pipetting steps. In new PCR-RNA-free tubes, 16.7 µl of real-time master mix was dispensed, and then 5.2 µl of preamplified cDNA was added. Thus, each cDNA of either a single oocyte or CC sample had three tubes of real-time reaction. The sample mixtures were then vortexed and spun down.

8.4.1. Sample loading

Each OpenArray chip contain 48 subarrays composing 12 wells sufficient for 12 samples. Therefore, each oocyte sample was prepared and run in three technical triplicates; CC samples were prepared similarly. All 12 samples were loaded onto one chip, as the OpenArray chip is divided horizontally into four subarrays (A, B, C and D). From each

sample, 5 µl was dispensed into A, B, C and D, resulting in 20 µl from each sample. After loading, the chips were covered, sealed, filled with immersion oil and centrifuged for 30 seconds at high speed.

8.4.2. Real-time PCR instrument

The customised 224-assay OpenArray real-time PCR plates were run in the Quaintstudio 12k Flex system. Four plates were run at one time. A standard three-step thermal cycler protocol with an annealing temperature of 55°C was followed by a product melt–curve analysis consisting of an increase from 70°C to 94°C, with image collection at every 0.25°C.

8.5. Data analysis using the comparative Ct method

The threshold cycle (Ct) values generated from the RT-PCR were analysed using the delta delta Ct which is a comparative Ct method, developed by Livak and Schmittgen (2001), and defined by the following formula: $2^{\Delta\Delta CT}$. The equation is $2^{\Delta CT\ (Ct\ calibrator\ -Ct\ sample)}$, where 2 = reaction efficiency of 100%, assuming that the amplification efficiency of both target and reference are equal (266). For example, primers with a reaction efficiency of 95% will be $1.95^{-\Delta CT\ (CTsample-CTcalibrator)}$. The relative expression levels for each housekeeping gene were calculated in each dataset for each age group, the 3-week-old, 9-week-old and 1-year-old oocytes. Then, the delta $C_{T\ (}\Delta Ct)$ was calculated by subtracting the Ct value of the target gene (224 genes in each sample) from the housekeeping genes of the same age group. Using this method, the Ct values were normalised to the endogenous control in each run. Then, the average of the delta Ct of six samples were calculated in each age group. Using the 9-week-

old oocytes as a reference, the ΔΔCT values were calculated by relating the delta Ct of the 3-week-old oocytes to the delta of the control (9-week-old oocytes). This was used to transform the C_T values into normalised relative expression values by relating the C_T values of the samples to those of the control/calibrator for each gene (266).

8.6. OpenArray real-time qPCR result

8.6.1. Housekeeping genes

To evaluate the reference (housekeeping) genes' stability for the real-time qPCR oocytes, the variance in the Ct values of the candidate housekeeping genes (Gapdh, Rpl19, Hprt1, Actb and Tubb) were calculated. The results showed that Gapdh was the most stable gene for the 9-week-old age group (n = 6), which was the control oocyte group (Figure 8.7). Gapdh also had the most constant Ct value among other housekeeping genes for the 3-week-old age group (Figure 8.7B), although with different Ct values for the 3-week-old, 9-week-old and 1-year-old age groups (Figure 8.10B).

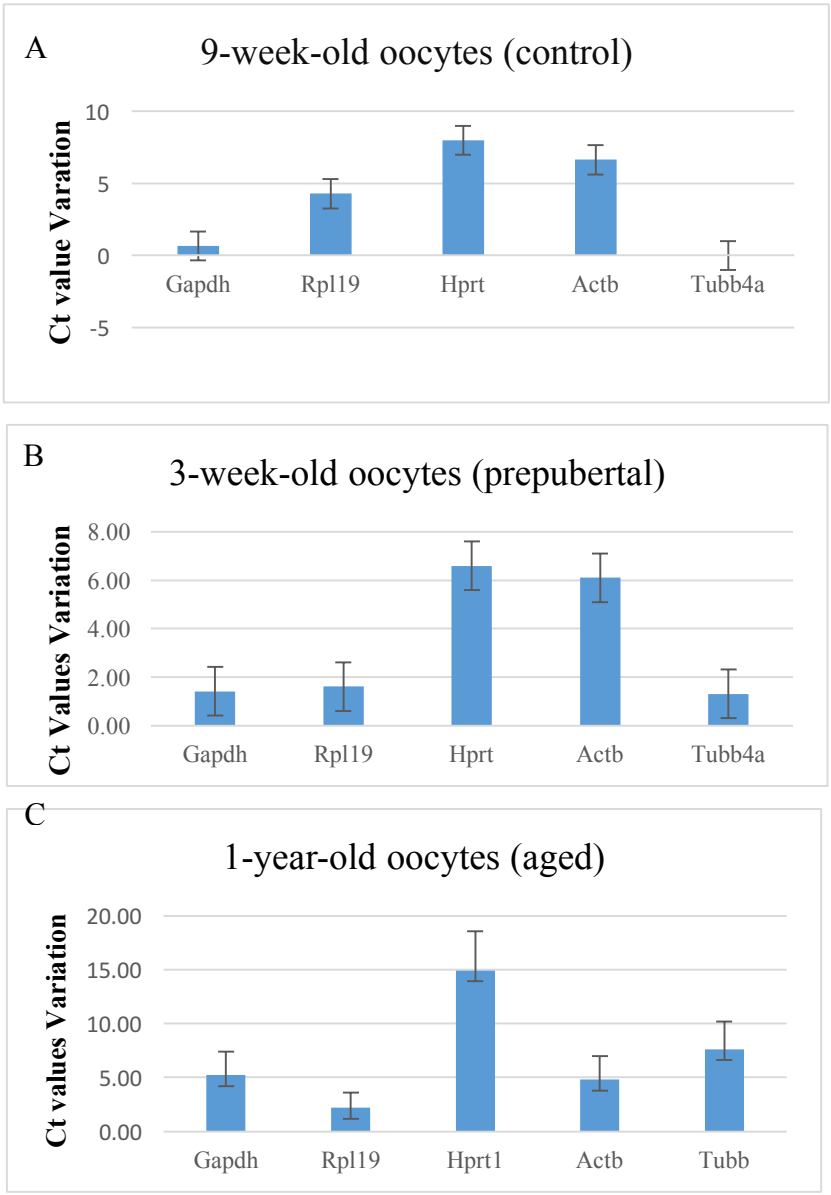
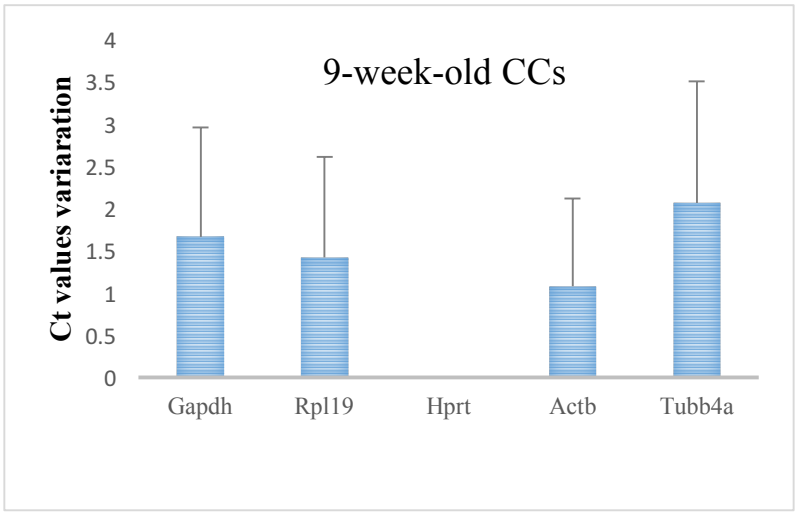


Figure 8.7 Ct value analysis of five candidate housekeeping genes (Gapdh, Rpl19, Hprt, Actb and Tubb).

The mean of Ct values was calculated from three technical replicates. Then, variance was analysed across six biological replicates.

A) The variation of five housekeeping genes in 3-week-old oocytes (n = 6); B) Ct value variation in 9-week-old oocytes (n=6); C) Ct variation in 1-year-old oocytes (n = 6).



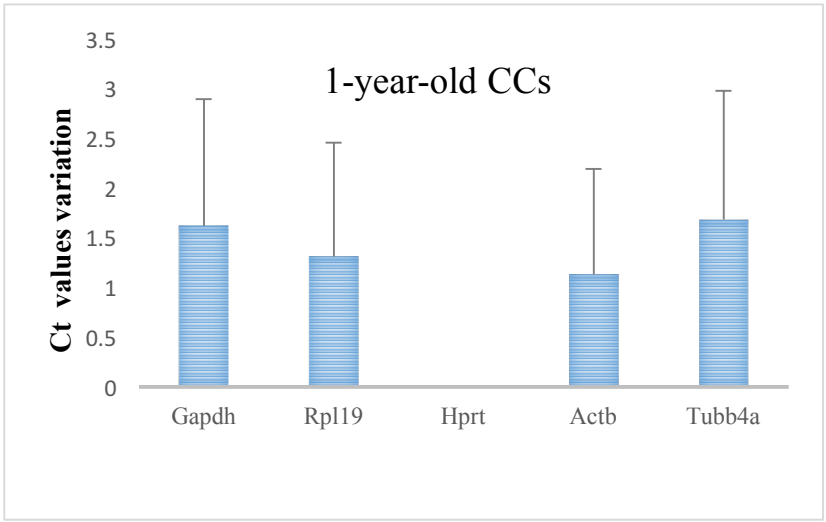


Figure 8.8 Variation analysis of Ct values of five housekeeping genes (Gapdh, Rpl19, Hprt1, Actb and Tubb) in CCs samples (n = 6).

A) Actb has the most constant Ct value for 9-week-old CCs whereas Hprt1 is not expressed in most CC samples; B) Actb shows the least variance in Ct value, while Hprt1 is not expressed in 1-year-old CCs (n = 6).

8.6.2. Three-week-old oocytes

Approximately 73 assays were suspected to be detected by OpenArray qPCR in 3-week-old oocytes. Of these, 13 are known to be oocyte-specific genes and 5 are known as housekeeping genes. Fifty-five genes were differentially expressed in the 3-week-old oocyte data generated by RNA-seq. Of those, oocyte-specific genes were also detected in the 3-week-old oocytes by the OpenArray real-time qPCR (Figure 8.9). The relative expression levels of the expressed genes were calculated using the formula: $2^{\Delta\Delta CT}$ and relating to the 9-week-old oocytes, which were used as a calibrator (control group).

About 30 of the 55 genes were expressed in the 3-week-old oocytes by qPCR. All of these were upregulated (Figure 8.10). OpenArray qPCR was not able to detect the downregulated genes. Gene expressions generated by qPCR were then compared with their corresponding values from sc-RNA-seq (Figure 8.11), as the OpenArray qPCR plates contain all the selected genes from the RNA-seq data, including those that are CC specific. CC-specific genes were not expressed in the 3-week-old oocytes, which confirms the qPCR reaction specificity.

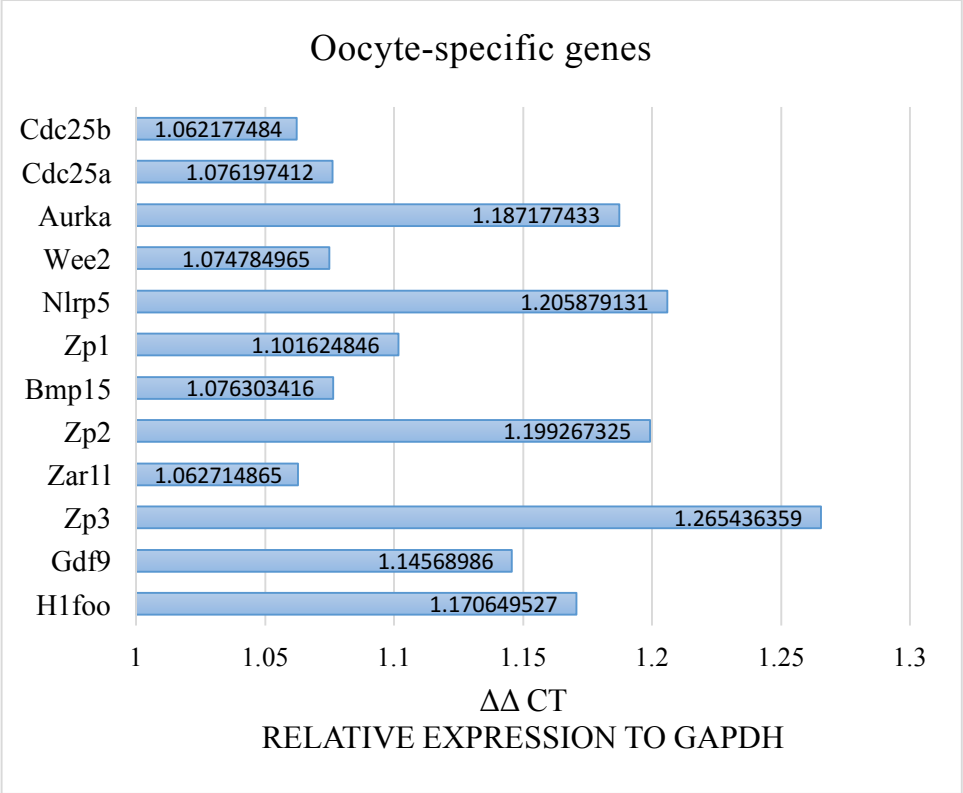


Figure 8.9 Oocyte-specific genes were detected in 3-week-old oocytes (n = 6), generated by OpenArray qPCR. The relative expression levels were calculated using the formula: $2^{\Delta\Delta CT}$.

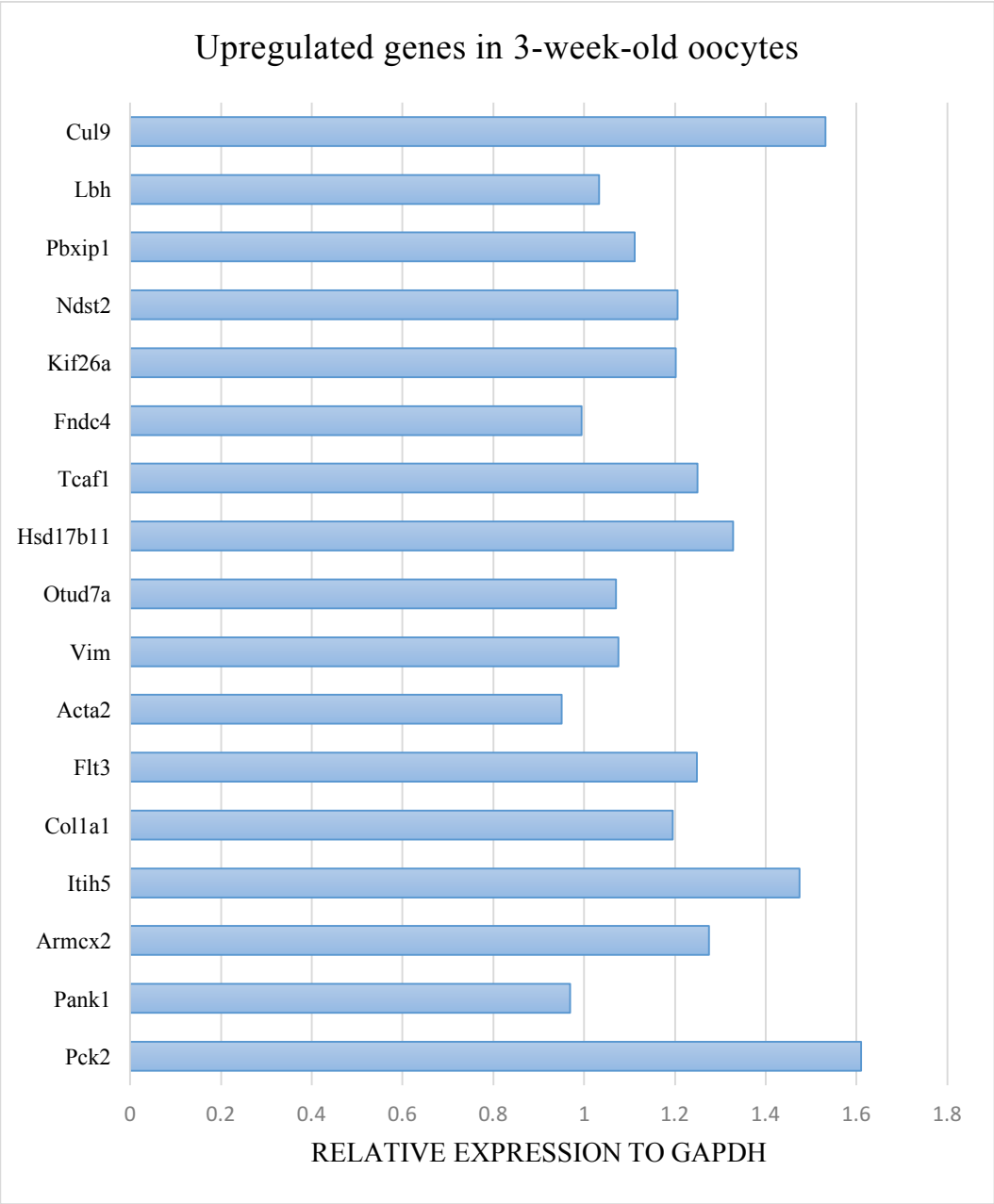


Figure 8.10 Relative expression of the upregulated genes in 3-week-old oocytes determined by OpenArray qPCR. The relative expression levels of the genes were calculating using 2ΔΔ Ct method.

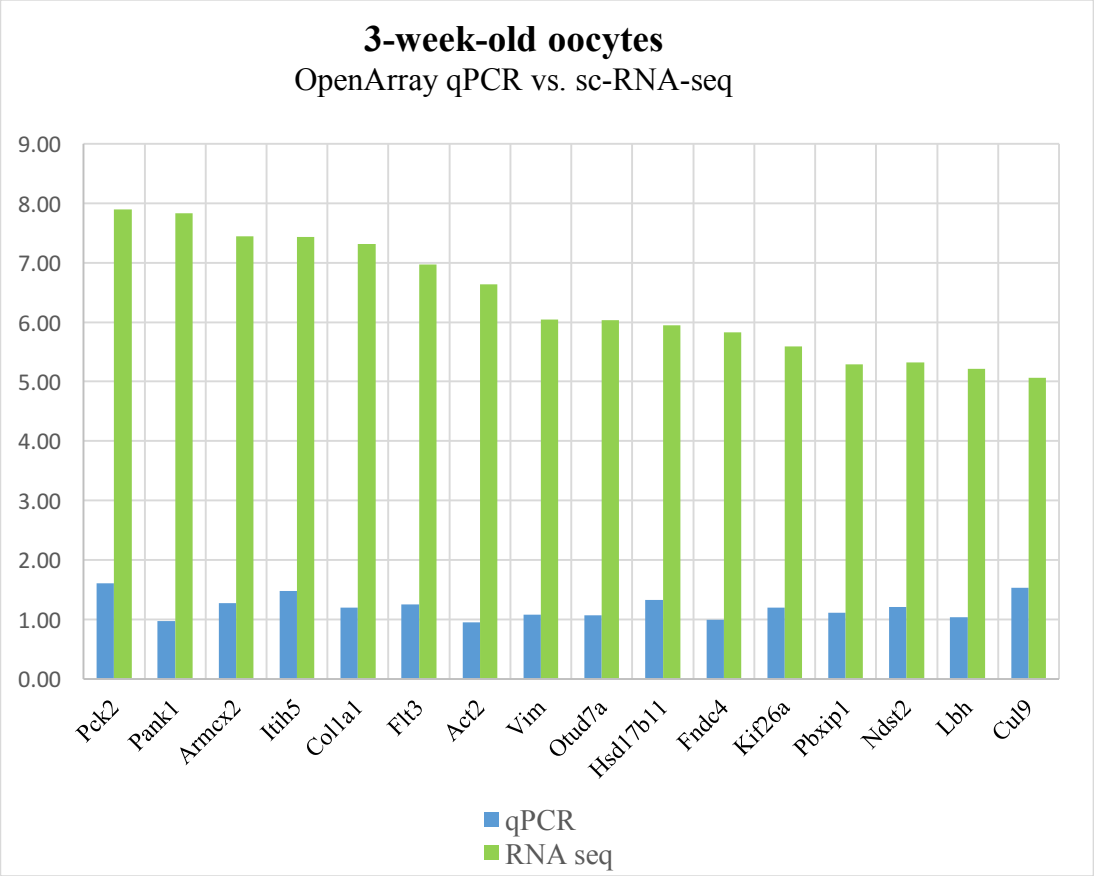


Figure 8.11 Comparison of sc-RNA-seq (green) and OpenArray qPCR (blue) data. The expressed genes were detected by both techniques. RNA-seq generated accurate measurement of the gene abundance.

8.6.3. One-year-old oocytes

Seventy-one genes were expressed in the 1-year-old oocytes determined through RNA-seq. These were validated by OpenArray qPCR, where 44 of the 71 genes were detected. Of those, the 13 oocyte-specific genes expressed in RNA-seq data were also examined in real-time qPCR, and they were all expressed (Figure 8.12). As determined by OpenArray qPCR, approximately 24 genes agreed with the selected genes from the 1-year-old oocytes' sequencing data. The list of genes was compared with the fold change of differentially expressed genes generated from RNA-seq in Figure 8.13, whereas downregulated genes from the RNA-seq were not expressed by OpenArray qPCR.

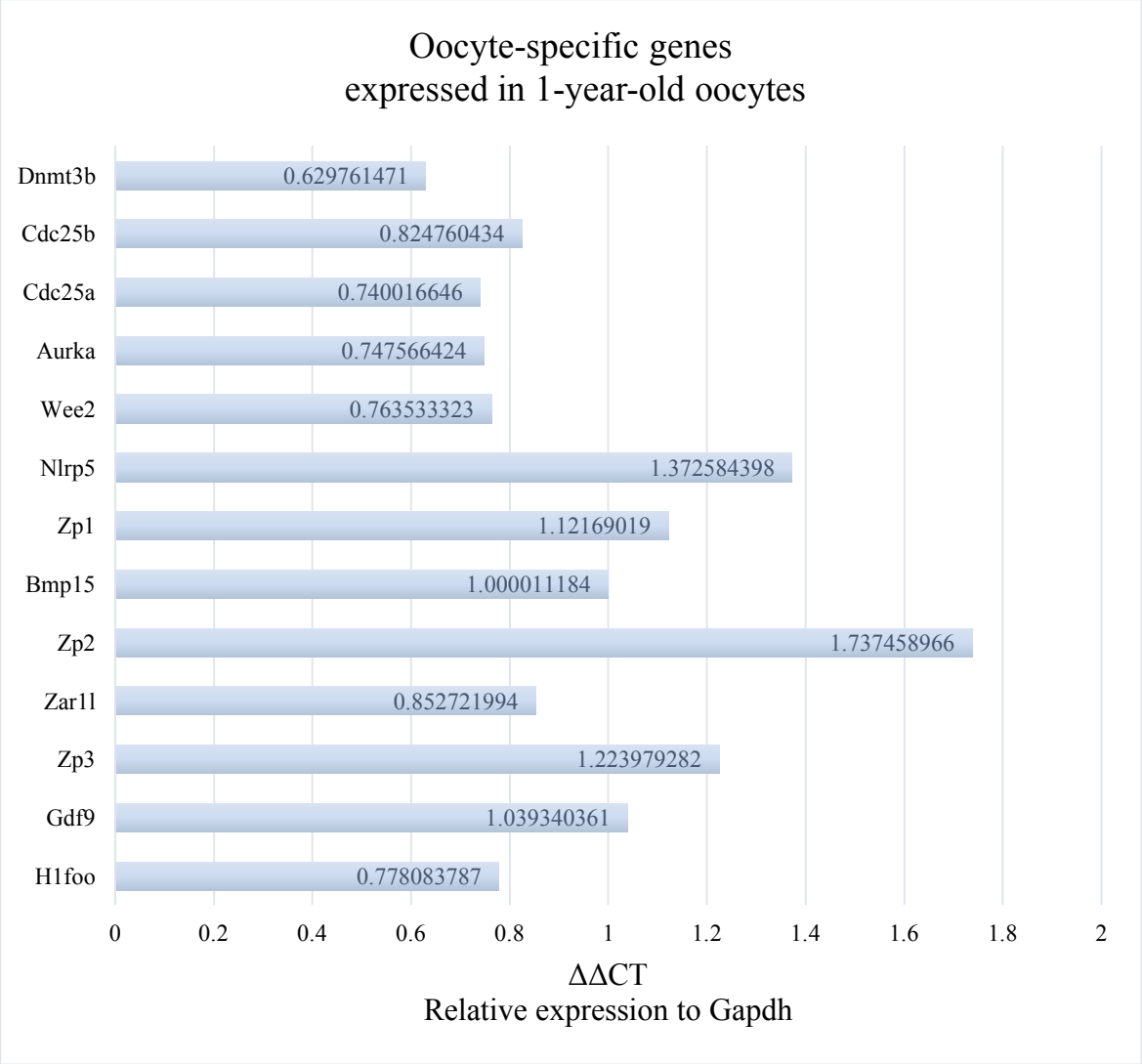


Figure 8.12 Oocyte-specific genes were detected in 1-year-old oocytes (n = 6). The gene expression is relative to that seen in 9-week-old oocytes, generated by OpenArray qPCR. The relative expression levels of the expressed genes were calculated using the formula: $2^{\Delta\Delta CT}$.

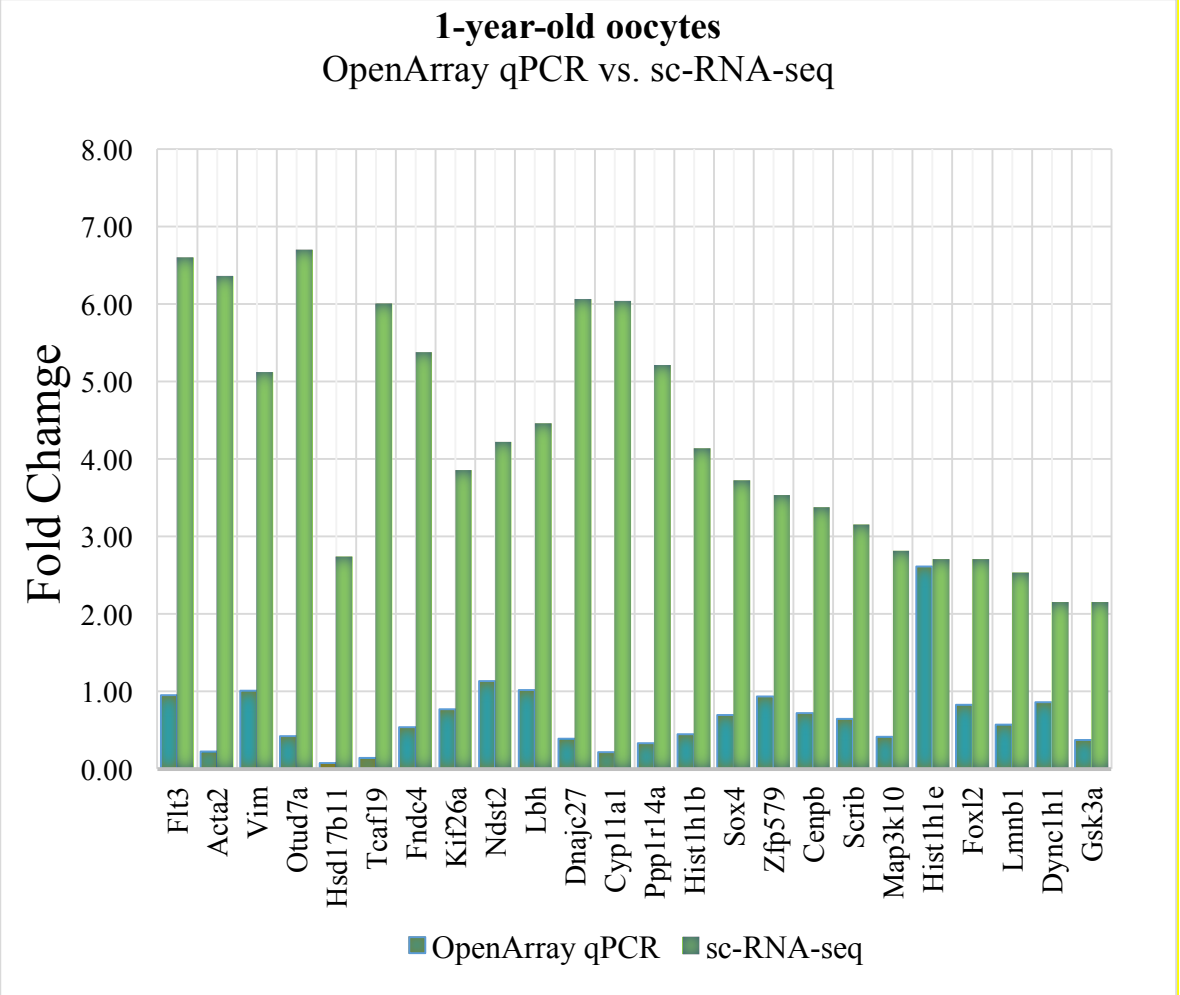


Figure 8.13 Comparison between the expression of upregulated genes generated by sc-RNA-seq and OpenArray qPCR.

8.6.4. One-year-old CCs versus 9-week-old CCs

As a part of validating the RNA-seq data for CCs, it was crucial to assess the RT-PCR performance in CCs using the positive control genes known to be expressed in CCs. The OpenArray qPCR revealed the expression of CC-specific genes (Figure 8.14) that were previously generated by RNA-seq (Chapter 7, Figure 7.2). These data confirmed the successful RT-PCR reaction in the CC samples. Then, to validate the 20 downregulated genes, which were the only differentially expressed genes in aged CCs generated by RNA-seq, they were examined by qPCR. Surprisingly, none of these downregulated expressed genes were detected in OpenArray qPCR except the Ctgf gene, which has the least negative fold in RNA-seq data of aged CCs. The remaining CCs genes that were configured in OpenArray were designed for the comparison between the 3- and 9-week-old CCs. However, since no difference in gene expression was found by the RNA-seq, which is far more sensitive then qPCR, these genes were not assessed.

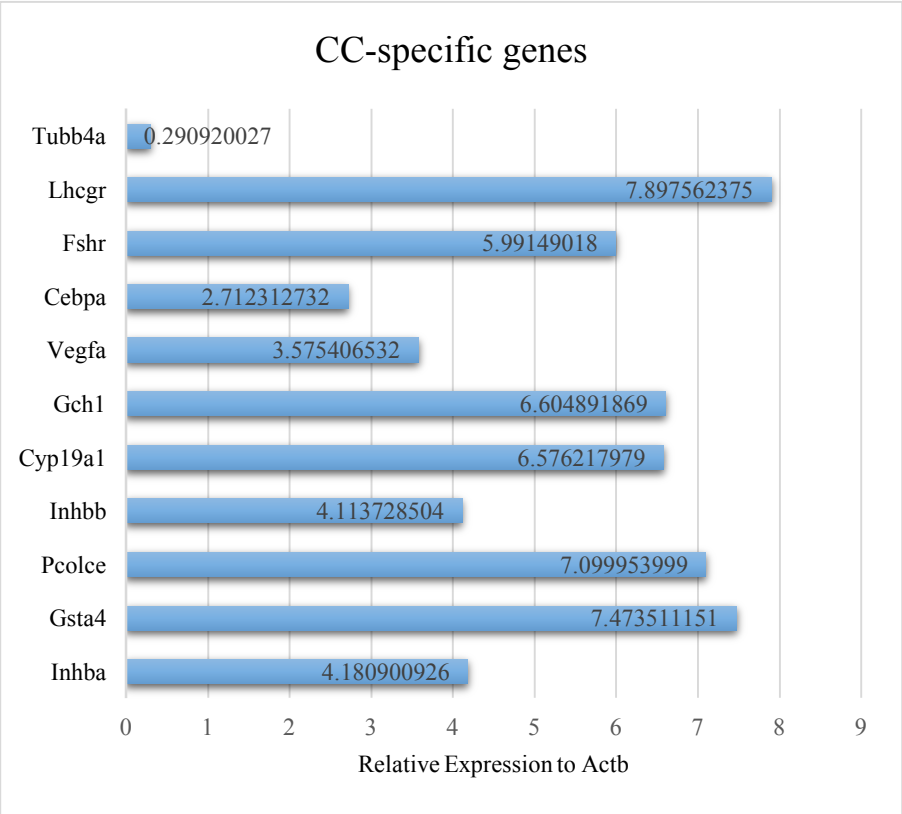


Figure 8.14 Cumulus cells-specific genes were detected in aged CCs using OpenArray qPCR. These genes are known to be expressed by murine CCs (58). They were also detected by RNA-seq in aged CCs. Gene expression was calculated in relation to Actb.

Table 8.6: RNA-seq data showing 20 downregulated expressed genes. OpenArray qPCR was not able to detect genes with very low expression except for Ctgf, which had the least negative fold in aged CCs.

Geneid	**LogFC**
Cxx1c	–10.87482248
Asb4	–9.691720932
2010111I01Rik	–9.390542088
Col9a2	–8.552240981
Loh12cr1	–7.780678509
Prkd1	–7.222064898
Omd	–7.050367035
Pdk4	–7.025715078
Spon1	–6.881817396
Tmprss13	–6.46907042
Emilin2	–6.050495185
Il17ra	–6.040409767
Gm15800	–5.597793897
Ccdc157	–5.50881712
Gpatch4	–5.367285394
Vps8	–5.312694674
Klhl31	–4.706772123
Kazald1	–4.684360739
Tpd52l1	–4.173442836
Ctgf	–3.776659609

8.7. Discussion

Real time- PCR is known to be a gold standard technology for accurately quantifying transcript abundance in biological samples (252). It is a powerful molecular technique that enables gene expression studies to be performed on small samples, including those from oocytes (255,267). OpenArray real-time PCR is a hybrid approach that possesses the accuracy and sensitivity of the qPCR approach coupled with the ability to generate high-throughput data similar to the hybridisation-based microarrays (256,257). High-throughput RT-PCR was recently developed to analyse larger numbers of genes and samples using small reaction volumes. The microlitre volumes used in 384-well microplates were scaled down to nanolitres in the OpenArray qPCR platform, providing the advantage of eliminating the need to use large quantities of biological material (258).

The aim of this chapter was to use an OpenArray RT-PCR for validating the RNA-seq data of single oocytes and their CCs in detailed validation analyses, rather than using pooled oocytes, which could mislead the interpretation. Traditionally, the first step in gene expression experiments is obtaining pure, high-quality RNA. Thus, RNA isolation is the key element in accurate gene expression measurement. To apply single-cell gene expression profiling using OpenArray RT-PCR, optimising a good preparation strategy is fundamental for ensuring the RNA input is sufficient for the OpenArray RT-PCR requirements.

Cell-to-Ct Gene Expression from ThermoFisher (Applied Biosystems, UK) is a two-step preparation designed for RNA input from 10–10^5 cells. However, because the oocyte is a large single cell (unlike somatic cells), and at the same time, has limited RNA, it was not clear if it would be compatible with the OpenArray plate. Because one plate is equal to eight

384-well plates, the amount of RNA (input of cDNA) must span the expression levels of the target genes. Therefore, emphasis was placed on reducing the preparation steps and as much hands-on manipulation as possible to avoid losing RNA. The cell-to-Ct gene expression extraction kit was modified for use for RNA recovery from single oocytes and their CCs, as it has the advantage of not requiring RNA purification, a step normally conducted after PCR amplification in other RNA extraction methods. It eliminates the need for any washing steps, which certainly lead to RNA losses (268). DNase was added to the lysis buffer to eliminate the susceptibility of presence of genomic DNA in RNA samples and ensure they were not contaminated by DNA, as it is virtually impossible to completely remove genomic DNA from RNA preparations.

The first pilot was performed to assess the RNA input in single oocytes and compatibility with downstream RT and subsequent PCR. The study was based on the use of Gapdh and Cyp11a1 as experimental genes in a single oocyte and its CCs. To ensure that samples were genomic DNA free, minus reverse transcriptase control ('no amplification control', NAC) was run in parallel on all samples tested in the RT-PCR experiments. Typically, minus RT is a mock reverse transcription that has all the RT-PCR reagents except the reverse transcriptase and another non-template negative control that contains all the PCR content expect the template. using this method, each experiment was confirmed to be genomic DNA free. However, the data showed failed amplification from all samples. The gene of interest was not detected, nor was the housekeeping gene identified. Many reasons for this were suspected, including the incompatibility of the RNA extraction kit with a single oocyte.

The second pilot was designed to figure out if using a high RNA volume would allow the detection of Gapdh and Cyp11a1 expression. Therefore, COCs were not separated, but instead, transferred immediately to the lysis buffer and processed. Pilot 2 showed successful RT-PCR amplification that enabled the detection of Gapdh and Cyp11a 1 in COC samples. This result confirmed that the amount of RNA input recovered from a single oocyte in the first pilot study was not compatible with the reagent given by the protocol. Thus, it was considered concentrating the cell suspension may result in a better amplification. Based on this, in pilot 3, the reagent volumes were scaled down to be compatible with the oocyte size. Moreover, a preamplification step was implemented in the preparation protocol to increase the cDNA yield. cDNA was preamplified by 14 PCR cycles, after which, the preamplified cDNA underwent a RT-PCR reaction that successfully enhanced the cDNA yield from a single oocyte. A positive control from pilot 2 was run in parallel with the single oocytes and their CC samples.

A proper normalisation strategy must be applied to obtain an accurate measurement of gene expression. Real-time PCR experiments should include stable endogenous controls to correct the sample-to-sample variation when calculating the qRT-PCR efficiency, as well as reducing error in sample quantification (269,270). A measured difference of transcript abundance between two samples may result from both true biological and technical variations. Technical variations must be removed or minimised. In the study, three technical replicates of each single oocyte and CC sample were conducted using similar PCR components, and extremely similar Ct values were produced in all three triplicates, which were then aggregated into a single measurement (Ct average). These results confirmed the high reproducibility of the OpenArray RT-PCR. Conducting technical replicates enhances

the precision of the data and allows for a clearer picture of the true biological variation related to the transcriptional activity rather than technical artefacts (271,272).

The most common method of data normalisation is using housekeeping genes, which presumably remain constantly expressed and should not be affected by the experimental treatment or condition (259). Several transcriptome studies have utilised constitutively expressed housekeeping genes, such as Gapdh in human and mouse oocytes (139). Other studies use Hprt1 and β-actin (Actb) to measure the variation in oocyte or embryo mRNA abundance (273–275). Housekeeping genes are often used as reference genes, as they are present in all nucleated cell types and required for fundamental cell survival (259,262). Despite this, one study revealed that the expression of the traditional housekeeping genes, such as β-actin and Gapdh, are significantly regulated in various cell and tissue, making them unreliable. Therefore, five genes – Gapdh, Rpl19, Hprt1, Actb and Tubb – were chosen as potential gene references for the oocytes and CCs, as the RT-PCR experiments in most ovarian transcriptome studies have used these genes. Initially, it was planned to use more genes to conduct a *geometric averaging* of multiple selected housekeeping genes to achieve accurate normalisation. However, the OpenArray RT-PCR data revealed that the reference genes' stability varied in a manner that was not expected. Although oocyte housekeeping genes were presumed to be constantly expressed in all developmental stages of the oocyte, Hprt1 appeared to be the most variable housekeeping gene among the oocytes in each age group (Figure 8.7), whereas the Rpl19 expression level was nearly consistent in most of the samples, making it a reliable control. The fluctuation of the reference genes in the oocytes could indicate that these genes contributed in part to the transcriptomic changes during

maturation, as the current study revealed that oocytes undergo dynamic transcriptional changes associated with age, as described in Chapters 5 and 6.

While the housekeeping genes were found to be relatively variable in both prepubertal and aged oocytes, they were more constant in the CC samples (Figure 8.8). The present data are consistent with study that found dramatic variations in mRNA levels during oocyte development, which in turn, made the normalisation of oocyte and preimplantation stage embryo data more challenging (262). Such normalisation cannot be based on the identical principles to those used for somatic cells. In fact, oocytes are thought to be different from somatic cells, in that mRNA levels fluctuate during oogenesis and early development, and thus, normalisation strategies are prone to mistakes (276).

Based on the findings, only one gene of each cell type was selected to be used in normalising the gene expression data. Gapdh and Actb were selected for the oocyte and CCs samples, respectively. Some papers have suggested using two genes to generate more accurate results (270,271); however, when this strategy was applied to the present data, no differences were found.

To assess the performance of OpenArray qPCR reactions, genes known to be oocyte-specific, which were detected by the sc-RNA seq quality (chapter 5, Figure 5.4), were also implemented in the OpenArray RT-PCR plate. These oocyte-specific genes were successfully detected in the 3-week-old, 9-week-old and 1-year-old oocytes. The oocyte-specific genes were initially selected from the literature (130). These genes were found to

be upregulated in 3-week-old and 1-year-old in comparison with 9-week-old oocytes, in agreement with the RNA sequencing of single oocytes generated in this study.

Real-time PCR revealed about 40 genes in the 1-year-old oocytes; however, the genes detected by qPCR were compared with those that met the criteria (fold change > 2, FDR < 0.05). Of these, 24 upregulated genes were consistent with the sequencing data, while downregulated genes (from sc-RNA-seq data) were not detected. From a technical point of view, as most of downregulated genes selected from RNA seq data had very low expression, it seems that low levels of gene expression are not easily measured. Perhaps downregulated genes with different ranges of negative fold expression (from RNA-seq data) should be chosen to be examined by OpenArray RT-PCR.

The present study revealed that RT-PCR has limited accuracy when it comes to measuring transcripts expressed in very high or low abundance. This was seen in both 3-week-old and 1-year-old oocytes. In addition, it seems that RT-PCR has a limited dynamics range compared with RNA-seq. This was evidenced by the CCs' qPCR data, as 20 genes selected from RNA-seq data were not validated, since they were not detected by OpenArray RT-PCR. The only gene that was found was Ctgf, which had the least negative fold (Table 8.6). CC-specific data were detected in aged CC samples and showed a high level of expression (Figure 8.9).

The ideal design of the validation experiment should use the same samples employed for the main technique, and the RNA from a biological sample must be divided and run in both sc-RNA-seq and RT-PCR. This is crucial when it comes to comparing the gene expression levels generated by the two different approaches. However, applying this method was

infeasible due to the extremely limited RNA material isolated from a single oocyte. Another alternative method used in other publications is using an exact amount of RNA input in both transcriptomic methods (RNA-seq and RT-PCR) (261,268); however, measuring the RNA quantity from a single cell is still infeasible with the existing methods. Therefore, the problem faced here is that each experiment has a specific RNA extraction protocol and procedure representing different levels of potential and ability.

One possible reason for the inconsistent gene expression is the priming strategy used in sc-RNA-seq compared with that used in qPCR. In sc-RNA-seq, the Smart seq2 approach is based on using the oligo-dT primers, which bind to the poly(A) RNA, whereas RT-PCR uses random hexamers that randomly bind to the RNA. In turn, the RNA regions are not represented similarly by both priming strategies (277). A previous study found that the expression levels of reference genes, such as TUBB and Succinate Dehydrogenase Complex Flavoprotein Subunit A (SDHA), were affected by using different priming strategies during RNA preparation (276).

One of the main differences between RNA-seq and qPCR is that level of downregulated genes expression could not be detected with qPCR. However, this was expected to some extent, as the genes chosen from the RNA-seq data exhibited very low expression. This could be the result of two potential scenarios. The first possibility is that it is extremely likely that the expression of these genes is outside the dynamic range of the chosen designed assay in RT-PCR. The other possibility is that the cDNA input was not sufficient for all assays in one subarray; that is, when the cDNA sample was dispensed (5 µl), the cDNA was distributed to all the assays (64 assays) in one subarray. Multiplexing many assays in one

well, it appears that the target primer competes with others to obtain the cDNA. Thus, moderately expressed genes had the most cDNA, whereas low and highly expressed genes could not obtain enough cDNA to present a clear result.

Overall, the data generated are considerably biased toward moderately expressed genes. In general, RNA-seq exhibited much better performance than qPCR did. The data confirmed that the coverage of 30 million reads was powerful in revealing novel genes that cannot be detected in RT-PCR.

Different variables inherent to the RT-qPCR workflow must be controlled to minimise technical variation. This includes the amount and quality of starting material, enzymatic efficiencies and overall transcriptional activity. During the configuring of the qPCR, some of the primers were inventoried for first time use, as they had not been tested before. TaqMan assays were designed using a proprietary algorithm developed by ThermoFisher Scientific. The primers were a set of oligonucleotides, and the algorithm evaluated a set of optimal assays, considering criteria like the melting temperature and nucleotide composition of primer-pair combinations. The algorithm selected the assays with the highest specificity based on nucleic acid sequence comparisons of assay primers and probes with genomic sequences from other closely related species. From the sequence comparison, the assay with the highest mismatch score with other organisms was chosen, thereby minimising the possibility of the generation of false-positive results during testing.

Several studies have compared microarray results with qPCR data, and they have found an excellent correlation between the results from the two assays. As for RNA-seq and RT-PCR, the discrepancies existing in this qPCR study are most likely due to bias in the RT-PCR

experiment rather than the RNA-seq data. sc-RNA-seq appears to be the most reliable technique, since normalisation to a reference gene is avoided. In addition, RNA-seq does not require predesigned probes; rather, it can detect and quantify entire transcripts, including unknown transcripts. Thus, it avoids the major technical issues of RT-PCR and microarrays.

Most RNA-seq publications run qPCR for a few genes (77,130,278–280). In these RNA seq studies, the genes from RNA-seq data selected for validation by RT-PCR were standard genes, for which all the above technical challenges are avoided, whereas novel genes were not validated. Thus, the RT-PCR results do not necessarily resemble the RNA-seq data in these previous findings. Furthermore, from a technical perspective, single-tube RT-PCR was used, which relied on using the whole sample for one gene whereas the present study was designed to investigate 224 genes in one sample.

The most important factor that must be addressed for improving –RT-PCR performance is determining how much cDNA input OpenArray RT-PCR requires to detect the transcripts with high or low expression. This study demonstrated that high throughput qPCR requires more cDNA to achieve the level of accuracy expected from this technology. Another suggestion is organising the assay in the OpenArray plate in a manner that allows the high-expression genes to be placed in different subarrays (in validation experiments) rather than organising them ascendingly according to the highest fold expression level, as in the design of the OpenArray plate in this study. In addition, when choosing the TaqMan assays, it is fundamental that the dynamic range of an assay is tested in advance and compared with the reference genes' dynamic range.

Overall, the results showed high reproducibility across the technical replicates and acceptable data quality but moderate agreement with the expression measurements of the RNA-seq data. High-throughput qPCR revealed the expression of more than 30 genes for each prepubertal and aged oocyte at the single-cell level. However, OpenArray RT-PCR may require extensive optimisation and troubleshooting to obtain reliable data. If the technical issues are addressed, OpenArray could have the potential to be used as a rapid molecular tool for multiple clinical applications, including assessing embryo and oocyte quality.

8.8. Conclusion

OpenArray qPCR is a promising technique for analysing gene expression in single oocytes and CCs, provided that the technical issues are addressed. Although RT-PCR appears to inherit some technical limitations of the hybridisation-based microarrays in terms of designing probes and primer characteristics, it has the advantage that minute quantities of material can be used – a crucial factor in single-cell gene expression profiling studies. Real-time PCR is still superior to microarrays and RNA-seq in terms of cost efficiency, simplicity of the protocols and quick turnout technique; improving cDNA amplification would provide a first step toward the goal of adapting OpenArray RT-PCR for clinical screening of potential markers for oocyte competence.

Chapter 9

General Discussion

The advent of single-cell RNA sequencing (sc-RNA-seq) has principally overcome the obstacles presented by previous transcriptomic analysis techniques. It solves the problem of a shortage of material required for the analysis and avoids the technical artefacts associated with hybridisation-based technologies (138). This study presented the successful application of sc-RNA-seq on individual oocytes and their surrounding cumulus cells (CCs). The overall aim was to investigate the changes in the transcriptome of a single oocyte and its companion CCs in relation to age. sc-RNA-seq has provided a unique view of alterations concerning the transcriptional activity in oocytes and CCs at the pre-pubertal, prime and advanced ages of reproductive life.

At the start of this project, no publications had addressed single-oocyte RNA transcriptomics. It was therefore necessary to optimise a preparation protocol for the sc-RNA-seq library. The Smart-seq2 approach was optimised to enable the generation of high-quality RNA from single oocytes and their CCs. Despite the technical challenges faced during the development of the preparation protocol, Smart-seq2 proved to be an efficient, reliable approach, allowing us to achieve high sequencing depth and high library complexity from a single oocyte. This was proven by the detection of 12 000–13 000 expressed genes in an individual oocyte and about 1642 differentially expressed genes in an aged oocyte; this high number of differentially expressed genes was not achieved in previous transcriptomic research, where 500 pooled oocyte samples were used and only 530 significantly differential genes were detected (52,53,113).

The present sc-RNA-seq study was distinct in that it revealed profound differences between pre-pubertal and young adult oocytes. About 2912 genes were differentially expressed in

pre-pubertal oocytes relative to 9-week-old oocytes. Three-week-old oocyte transcriptomics data revealed that several pathways associated with cytoplasmic functions continued to be expressed in an immature manner. The dysregulation of the Rho family GTPases signalling pathway, together with the integrin-linked kinase (ILK) signalling pathway, which both compose the cytoskeleton remodelling, indicated the immature characteristics of the cytoplasm. The changes in the levels of genes responsible for regulating the cytoskeleton dynamics (e.g. Mapk10, Vim, Sep9) indeed affected the spindle and microtubule formation, and consequently, the competence in achieving proper meiotic division.

Previous findings showed that *in vitro* maturation for MI (IVM) oocytes closely resembles that of *in vivo* MII oocytes in terms of cellular pathways related to nuclear maturity, but this is not evident in pathways associated with cytoplasmic maturity (281). The current study revealed that retinoic acid metabolism (PXR/RXR pathway) had a novel role in the cytoplasmic maturation of the pre-pubertal oocyte, and evidence was accumulated that RA and its derivatives have a major impact on diverse cellular functions, including regulating meiotic resumption, controlling lipid metabolism and steroidogenesis (282–284). Thus, this significant role of RA should be considered for implementation in the culture media composition for the IVM of pre-pubertal oocytes.

The pre-pubertal oocyte data demonstrated that the transcriptional activity is still undergoing global changes at this age, and it seemed that cytoplasmic maturation was not achieved. The increased accumulation of several metabolites in 3-week-old oocytes demonstrated that the cytoplasmic function remains underdeveloped. Successful fertilisation depends highly on the inherent qualities of the oocytes, and thus, it relies on oocyte maturation. The developmental competence seems to be gradual and sequential, which explains why the pre-

pubertal oocyte can be fertilised but has limited ability to develop into a healthy embryo. The molecular mechanisms that underlie competence acquisition were demonstrated, confirming that cytoplasmic competence, like meiotic competence, correlates closely with approaching puberty.

The present study provides baseline data for the pre-pubertal transcriptome, which has not been discussed before in the literature. Despite the increased clinical demand for using pre-pubertal oocytes for pre-pubertal cancer survivors who lost their fertility after cancer treatment, the transcriptomes of the pre-pubertal oocytes are barely discussed, and extremely limited research has been done, even in mammalians. Therefore, the differentially expressed genes and enriched pathways provide clues for the optimisation of IVM techniques, especially for infertile, pre-pubertal patients for whom IVM is the only available option so far.

In contrast to other methods, sc-RNA-seq provides a clear picture of the upregulated and downregulated genes in aged oocytes in relation to oocytes at a young adult age (9 weeks). The present data identified new regulatory genes that govern the ageing process, confirming that changes are obscured when oocytes are pooled.

The unbalanced lipid metabolism evidently has a major influence on the oocyte ageing process. The dysregulation of genes involved in ß oxidation, such as Cpt1a and Scd1, indicate the deterioration of mitochondrial function, which has adverse consequences in energy production. An interesting finding from 2016 demonstrated that several classes of phospholipids are affected in aged oocytes (285). The study used CellMask Deep Red staining to compare mouse oocytes aged 45–50 weeks with some as young as 4 weeks old; the results showed partial disruption of the oocyte zona and suggested that the integrity of

the plasma membrane was intact in aged oocytes. Because lipid is the structural backbone in oocytes, as well as their source of energy, its alteration is associated with a decreased quality of mouse oocytes (285).

Interestingly, when H2O2 was used as an exogenous agent to induce the oxidative stress in a group of 4-week-old oocytes as an attempt to partially mimic the oxidative stress originating from the mitochondria in aged oocytes, by comparing the lipid profiles among the young oocytes, aged oocytes and oocytes treated with H_2O_2, the aged and H_2O_2-treated oocytes were found to have abnormal lipid profiles. This finding shows an evident association between oxidative stress and lipid dysregulation (285). While lipid dysregulation seems to be a major hallmark of the ageing process in oocyte senescence, oxidative stress has also been labelled as a major characteristic of aged oocytes; however, it is difficult to ascertain which changes in mitochondria induce the cascade of oocyte deterioration. Elevated reactive oxygen species (ROS) production is linked with the increased lipid oxidation observed in several age-related diseases (286). Both lipid dysregulation and oxidative stress were noticed in both oocytes and aged CCs. Oocyte ageing involves the dysregulation of several interconnected biological processes.

It was surprising to observe that the atypical expression of some pathways was common in both pre-pubertal and aged oocytes (Figure 9.1). The examples include the retinoic acid metabolism (PXR/RXR pathway); many genes in this pathway are dysregulated in both prepubertal and aged oocytes. The patterns of gene expression were different in some differential genes, but this generated a new hypothesis that atypical transcriptional activity of the retinoic acid pathway has a strong link with poor oocyte quality, regardless of age. Retinoic acid seems to be a key regulator in oocyte development that should be given more

attention. Lipid metabolism and steroidogenesis have been found to be controlled by RA; therefore, the RA level must be tightly regulated to improve the oocyte quality (287). Nevertheless, this assertion requires further investigation.

Overall, the involvement of the four pathways in the dysregulation of the transcriptional activity in both ages revealed a strong association of these pathways with decreased oocyte competence, regardless of the oocyte's age. The comparison of gene expression between oocytes with poor competence, such as prepubertal and aged oocytes, and those with full developmental competence, indicated that the malfunctioning of these certain pathways may govern the transcriptome and have a major influence in compromising oocyte maturation. It would be of great benefit if these genes were to be researched extensively to used them as markers for oocyte quality.

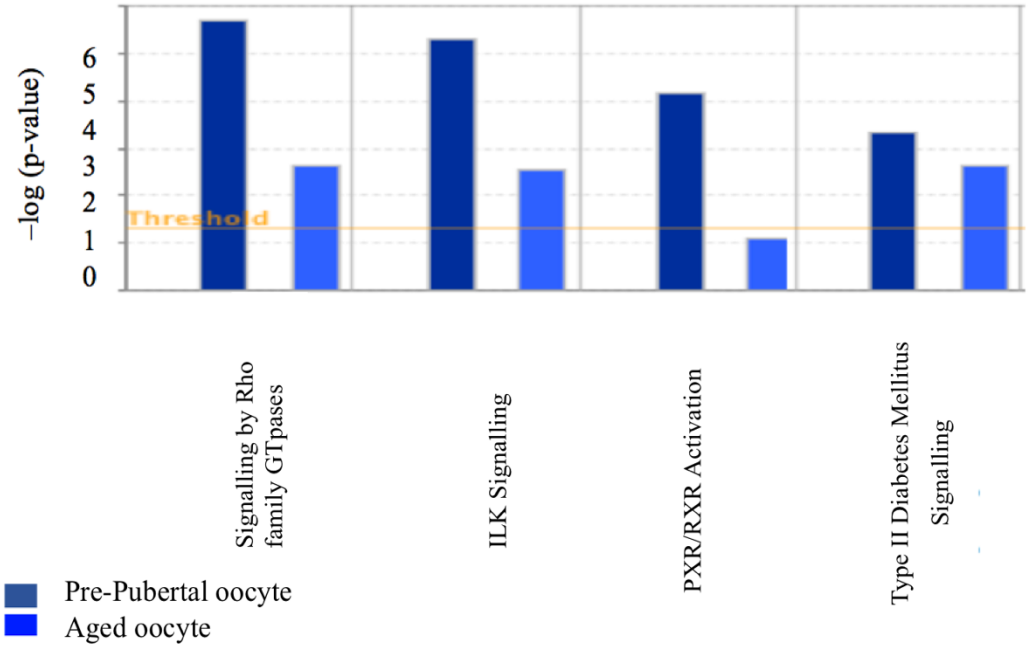


Figure 9.1: Common pathways between the pre-pubertal and aged oocytes relative to young adult oocytes.

The dysregulation of four pathways, namely GTPases by the Rho signalling family, lLk pathway, PXR/RXR activation pathway, and type II diabetes mellitus, indicate certain molecular mechanisms underlying the poor quality of oocytes.

Evidently, genes fluctuate in response to the oocyte's needs. In addition, changes in the ovarian microenvironmental conditions could induce changes in the genes' expression in oocytes. Cumulus cells seem to provide the oocyte with factors that rescue the deteriorated functions in aged oocytes. The changes in aged CCs are presumably an adaptation mechanism; thus, unbalanced CC factors come into concordance with the altered metabolic requirements of the aged oocyte. The current study showed that changes in lipid metabolism were found in both aged oocytes and CCs. Because many lipid classes have common structural features, and they are often regulated by the same enzymes in class-specific manners, a high degree of co-regulation is presumed in the lipid profiles of oocytes and their CCs. It is evident that the surrounding CCs play a role in supporting oocyte lipid homeostasis, as a large part of the beta oxidation process occurs in CCs (288,289). The fatty acid metabolism in CCs was recently analysed via a microarray study demonstrating that a significant enrichment of lipid metabolism-related genes was detected in CCs from *in vitro* mature oocytes, suggesting that it is one of the most important functions of CCs during oocyte maturation (188,290,291).

Taking the results together, the altered levels of lipid metabolism-related genes in aged CCs at 1-year-old shown in the present study illustrated that the CCs' lipid was clearly is important for keeping a normal oocyte development. Moreover, the findings highlighted how alteration in CCs lipids affects the early stage of embryo development at an advanced age.

It was mentioned above that retrieving the oocyte and CCs from the same complex represents a strength of this study, as most publications have contradictory findings; for instance, an increase in apoptotic CCs was correlated with poor-quality oocytes in one study,

whereas a later study showed that poor oocytes have increased CC proliferation (67,68). Therefore, conducting the analysis of both cells taken from the same follicle rules out this suspicion; more importantly, it correlated the relationship between the oocyte and CCs in an aged ovary.

The oocyte is largely responsible for regulating events in fertilisation and embryonic development through the maternal genome, which drives the oocyte growth until embryonic activation (13,292). Therefore, most aberrations are a consequence of inherent defects in the oocyte (47). The main concern is that oocytes with diminished developmental competence have non-visualised abnormalities, as the event compromising cytoplasmic maturation cannot be identified. Atypical cytoplasmic maturation may cause failure at any step of a fertilisation event, without any morphological indicator, such as the extrusion of the first polar body in the oocyte or pronucleus formation, which are indicators of nuclear – but not cytoplasmic – maturation (155,293).

Developing a molecular-based assay could represent a new, reliable way to assess the oocyte's competence, as the judgment of oocyte quality is still infeasible in clinical practice, and this is mainly due to the lack of knowledge of what oocyte competence entails (294). This further supports the necessity of performing deeper investigation of maturation regulation before pursuing an explanation for fertilisation failure.

The OpenArray qPCR demonstrated the capacity to measure a panel of more than a hundred genes in one sample. However, comparing the results of gene expression generated by sc-RNA-seq and OpenArray qPCR indicated that sc-RNA-seq exhibits better performance than OpenArray qPCR does in terms of being more powerful in the quantitation of gene expression. Unfortunately, the highly expressed genes detected by the RNA-seq experiment

were not detected in OpenArray qPCR; similarly, qPCR was unable to detect the genes with low expression. This is because sc-RNA-seq relies on direct access to the sequencing of mRNA rather than needing a designed probe like qPCR does. Accordingly, sc-RNA-seq allows a larger dynamic range of gene expression compared with qPCR.

Improving ART treatment will mainly rely on improving oocyte health by understanding the oocyte's biological process. There are multiple ways of improving the quality if the causes of any problems are defined. Culture media can be modified and developed in a way that meets the requirements to remediate poor oocyte quality.

Overall, this thesis has added a database of the transcriptomic differences in pre-pubertal, adult and aged oocytes, as well as their companion CCs, proposing new insights into the influence of age on the oocyte and surrounding CCs. sc-RNA-seq shows great promise for revealing deeper insights into ovarian function. It offers deep coverage and throughput, making single-oocyte transcriptomics applicable to studies of deep differences at the molecular level.

References

1. Macklon NS, Fauser BCJM. Aspects of Ovarian Follicle Development throughout Life. *Horm Res*. 1999;52(4): 161–70.
2. Pedersen T, Peters H. Proposal for a classification of oocytes and follicles in the mouse ovary. *J Reprod Fertil.* 1986;17(3): 555–7.
3. Peters H. The development of the mouse ovary from birth to maturity. *Acta Endocrinol.* 1969;62(1): 98–116.
4. Macklon NS, Fauser BC. Aspects of ovarian follicle development throughout life. *Horm Res* 1999;52(4):161–70.
5. Albertini DF, Combelles CM, Benecchi E, Carabatsos MJ. Cellular basis for paracrine regulation of ovarian follicle development. *Reproduction.* 2001; 121(5): 647–53.
6. Himelstein-Braw R, Byskov AG, Peters H, Faber M. Follicular atresia in the infant human ovary. *J Reprod Fertil*. 1976; 46(1):55–9.
7. Peters H, Byskov AG, Himelstein-Braw R, Faber M. Follicular Growth: The Basic Event in The Mouse and Human Overy. *J Reprof Fert.* 1975;45(3): 559-566.
8. BORUM K. Oogenesis in the mouse. A study of the meiotic prophase. *Exp Cell Res*. 1961;24:495–507.
9. Li R, Albertini DF. The road to maturation: somatic cell interaction and self-organization of the mammalian oocyte. *Nat Rev Mol Cell Biol* . 2013;14(3):141–52.
10. Sánchez F, Smitz J. Molecular control of oogenesis. Biochim Biophys Acta. *Mol Basis Dis*. 2012;1822 (12):1896–912.
11. Tanghe S, Van Soom A, Nauwynck H, Coryn M, de Kruif A. Minireview: Functions of the cumulus oophorus during oocyte maturation, ovulation, and fertilization. *Mol Reprod Dev*. 2002; 61(3):414–24.
12. Eppig JJ. Coordination of nuclear and cytoplasmic oocyte maturation in eutherian mammals. *Reprod Fertil Dev*. 1996;8(4): 485–9.

13. De La Fuente R. Chromatin modifications in the germinal vesicle (GV) of mammalian oocytes. *Dev Biol*. 2006;292(1): 1–12.

14. Conti M, Franciosi F. Acquisition of oocyte competence to develop as an embryo: integrated nuclear and cytoplasmic events. *Hum Reprod Update.* 2018; 24(3): 245-266.

15. Jansen RPS. Germline passage of mitochondria: quantitative considerations and possible embryological sequelae. *Hum Reprod.* 2000;15(suppl 2): 112–28.

16. Van Blerkom J. Mitochondria in human oogenesis and preimplantation embryogenesis: engines of metabolism, ionic regulation and developmental competence. *Reproduction*.2004;128(3): 269–80.

17. Gandolfi TA, Gandolfi F. The maternal legacy to the embryo: cytoplasmic components and their effects on early development. *Theriogenology*. 2001;55(6):1255–76.

18. Telford NA, Watson AJ, Schultz GA. Transition from maternal to embryonic control in early mammalian development: A comparison of several species. *Mol Reprod Dev*. 1990;26(1): 90–100.

19. Leung PCK, Armstrong DT. Interactions of Steroids and Gonadotropins in the Control of Steroidogenesis in the Ovarian Follicle. *Annu Rev Physiol*. 1980;42(1): 71–82.

20. Azhar S, Tsai L, Medicherla S, Chandrasekher Y, Giudice L, Reaven E. Human Granulosa Cells Use High Density Lipoprotein Cholesterol for Steroidogenesis. *J Clin Endocrinol Metab*. 1998;83(3):983–91.

21. Rajan VP, Menon KMJ. Cholesterol Flux between High Density Lipoproteins and Cultured Rat Luteal Cells. *Endocrinology*. 1989;124(4):1857–62.

22. Sutton-McDowall ML, Gilchrist RB, Thompson JG. The pivotal role of glucose metabolism in determining oocyte developmental competence. *Reproduction.* 2010;139(4):685–95.

23. Biggers JD, Whittingham DG, Donahue RP. The pattern of energy metabolism in the mouse oöcyte and zygote. *Proc Natl Acad Sci*. 1967;58(2):560–7.

24. Dunning KR, Russell DL, Robker RL. Lipids and oocyte developmental competence:

the role of fatty acids and β-oxidation. *Reproduction*. 2014; 148(1):R15–27.

25. Kim JY, Kinoshita M, Ohnishi M, Fukui Y. Lipid and fatty acid analysis of fresh and frozen-thawed immature and in vitro matured bovine oocytes. *Reproduction.* 2001;122(1):131–8.

26. Wallace WHB. Oncofertility and preservation of reproductive capacity in children and young adults. *Cancer.* 2011;117(S10):2301–10.

27. Resetkova N, Hayashi M, Kolp LA, Christianson MS. Fertility Preservation for Prepubertal Girls: Update and Current Challenges. *Curr Obstet and Gynecol Rep.* 2013; 2(4): 218–25.

28. Dieleman SJ, Hendriksen PJM, Viuf D, Thomsen PD, Hyttel P, Knijn HM, et al. Effects of in Vivo Prematuration and In Vivo Final Maturation on Developmental Capcity and Quality of Pre-imaplanation Embryos. *Theriogenology*. 2002;57(1):5-20.

29. Krisher RL. The effect of oocyte quality on development. *J Anim Sci*. 2004;82(suppl_13):14–23.

30. De Matos DG, Nogueira D, Cortvrindt R, Herrera C, Adriaenssens T, Pasqualini RS, et al. Capacity of adult and prepubertal mouse oocytes to undergo embryo development in the presence of cysteamine. *Mol Reprod Dev.* 2003;64(2):214–8.

31. Khatun M, Bhuiyan MMU, Ahmed JU, Haque A, Rahman MB, Shamsuddin M. In vitro maturation and fertilization of prepubertal and pubertal black Bengal goat oocytes. *J Vet Sci.* 2011;12(1):75–82.

32. Telfer EE, McLaughlin M. Strategies to support human oocyte development in vitro. *Int J Dev Biol*. 2012;56(10-12):901–7.

33. Palmerini MG, Nottola SA, Leoni GG, Succu S, Borshi X, Berlinguer F, et al. In vitro maturation is slowed in prepubertal lamb oocytes: ultrastructural evidences. Reprod Biol Endocrinol. 2014;12(1):115.

34. Marchal R, Feugang JM, Perreau C, Venturi E, Terqui M, Mermillod P. Meiotic and developmental competence of prepubertal and adult swine oocytes. *Theriogenology.* 2001;56(1):17–29.

35. O'Brien J, Dwarte D, Ryan J, Maxwell W, Evans G. Developmental capacity, energy metabolism and ultrastructure of mature oocytes from prepubertal and adult sheep.

Reprod Fertil Dev; 1996;8(7):1029.

36. Velilla E, Izquierdo D, Rodríguez-González E, López-Béjar M, Vidal F, Paramio MT. Distribution of prepubertal and adult goat oocyte cortical granules during meiotic maturation and fertilisation: Ultrastructural and cytochemical study. *Mol Reprod Dev.* 2004;68(4):507–14.

37. Salamone DF, Damiani P, Fissore RA, Robl JM, Duby RT. Biochemical and Developmental Evidence That Ooplasmic Maturation of Prepubertal Bovine Oocytes Is Compromised Biochemical and Developmental Evidence That Ooplasmic Maturation of Prepubertal Bovine Oocytes Is Compromised 1. *Biol Reprod.* 2001;64:1761–8.

38. Combelles CMH, Cekleniak NA, Racowsky C, Albertini DF. Assessment of nuclear and cytoplasmic maturation in in-vitro matured human oocytes. *Hum Reprod*. Oxford University Press; 200;17(4):1006–16.

39. Romar R, De Santis T, Papillier P, Perreau C, Thélie A, Dell'Aquila M, et al. Expression of Maternal Transcripts During Bovine Oocyte In Vitro Maturation is Affected by Donor Age. *Reprod Domest Anim*. 2011; 46(1):e23–30.

40. Paczkowski M, Yuan Y, Fleming-Waddell J, Bidwell CA, Spurlock D, Krisher RL. Alterations in the transcriptome of porcine oocytes derived from prepubertal and cyclic females is associated with developmental potential. *J Anim Sci.* Oxford University Press; 2011;89(11):3561–71.

41. Broekmans FJ, Knauff EAH, te Velde ER, Macklon NS, Fauser BC. Female reproductive ageing: current knowledge and future trends. *Trends Endocrinol Metab*. 2007;18(2):58–65.

42. Navot D, Bergh PA, Williams MA, Garrisi GJ, Guzman I, Sandler B, et al. Poor oocyte quality rather than implantation failure as a cause of age-related decline in female fertility. *Lancet.* 1991;337(8754):1375–7.

43. Fertilisation H, Authority E. Fertility treatment 2014–2016 Trends and figures. 2018. cited [20 may 2018] available from /https://www.hfea.gov.uk/media/2563/hfea-fertility-trends-and-figures-2017-v2.pdf/.

44. Chiang T, Schultz RM, Lampson MA. Meiotic origins of maternal age-related

aneuploidy. *Biol Reprod*. Oxford University Press; 2012;86(1):1–7.

45. Eichenlaub-Ritter U. Parental age-related aneuploidy in human germ cells and offspring: A story of past and present. *Environ Mol Mutagen.* 1996;28(3):211–36.

46. Freeman SB, Yang Q, Allran K, Taft LF, Sherman SL. Women with a Reduced Ovarian Complement May Have an Increased Risk for a Child with Down Syndrome. *Am J Hum Genet.* 2000;66(5):1680–3.

47. Bartmann AK, Romão GS, Ramos E da S, Ferriani RA. Why do older women have poor implantation rates? A possible role of the mitochondria. *J Assist Reprod Genet.* 2004;21(3):79–83.

48. Broekmans FJ, Soules MR, Fauser BC. Ovarian aging: Mechanisms and clinical consequences. *Endocr Rev.* 2009;30(5):465–93.

49. Steuerwald N, Cohen J, Herrera RJ, Sandalinas M, Brenner CA. Association between spindle assembly checkpoint expression and maternal age in human oocytes. *Mol Hum Reprod.* 2001;7(1):49–55.

50. Carter MG, Hamatani T, Sharov A a., Carmack CE, Qian Y, Aiba K, et al. In situ-synthesized novel microarray optimized for mouse stem cell and early developmental expression profiling. *Genome Res.* 2003;13:1011–21.

51. Hamatani T, Falco G, Carter MG, Akutsu H, Stagg CA, Sharov AA, et al. Age-associated alteration of gene expression patterns in mouse oocytes. *Hum Mol Genet*. 2004;13(19):2263–78.

52. Pan H, Ma P, Zhu W, Schultz RM. Age-associated increase in aneuploidy and changes in gene expression in mouse eggs. *Dev Biol.* 2008;316(2):397–407.

53. Steuerwald NM, Bermúdez MG, Wells D, Munné S, Cohen J. Maternal age-related differential global expression profiles observed in human oocytes. *Reprod Biomed Online.* 2007;14(6):700–8.

54. Grøndahl ML, Yding Andersen C, Bogstad J, Nielsen FC, Meinertz H, Borup R. Gene expression profiles of single human mature oocytes in relation to age. *Hum Reprod.* Oxford University Press; 2010;25(4):957–68.

55. Labrecque R, Sirard M-A. The study of mammalian oocyte competence by transcriptome analysis: progress and challenges. *Mol Hum Reprod.* 2014;20(2)

:10316.

56. Duncan FE, Jasti S, Paulson A, Kelsh JM, Fegley B, Gerton JL. Age-associated dysregulation of protein metabolism in the mammalian oocyte. Aging Cell. 2017 ;16(6):1381-1393.

57. Eppig JJ, Peters AHFM, And EET, Wigglesworth- K. Production of Cumulus Expansion Enabling Factor by Mouse Oocytes Grown In Vitro: Preliminary Characterization of the Factor. *Mol Reprod Dev*. 1993;34:45–56.

58. Wigglesworth K, Lee K-B, Emori C, Sugiura K, Eppig JJ. Transcriptomic diversification of developing cumulus and mural granulosa cells in mouse ovarian follicles. *Biol Reprod*. 2015;92(1):23.

59. Hutt KJ, Albertini DF. An oocentric view of folliculogenesis and embryogenesis. *Reprod Biomed Online*. 2007;14(6):758–64.

60. Kedem-Dickman A, Maman E, Yung Y, Yerushalmi GM, Hemi R, Hanochi M, et al. Anti-Müllerian hormone is highly expressed and secreted from cumulus granulosa cells of stimulated preovulatory immature and atretic oocytes. *Reprod Biomed Online*. 2012;24(5):540–6.

61. Salustri A, Yanagishita M, Hascall VC. Synthesis and accumulation of hyaluronic acid and proteoglycans in the mouse cumulus cell-oocyte complex during follicle-stimulating hormone-induced mucification. *J Biol Chem*. 1989;264(23):13840–7.

62. Eppig2 JJ. Maintenance of Meiotic Arrest and the Induction of Oocyte Maturation in Mouse Oocyte-Granulosa Cell Complexes Developed In Vitro from Preantral Follicles1. Biol Reprod.1991;45(6):824-30.

63. Pincus G, Enzmann E V. The Comparative Behavior of Mammalian Eggs In Vivo and In Vitro I. The Activation of Ovarian Egg. *J Exp Med*. The Rockefeller University Press; 1935;62(5):665–75.

64. Eppig JJ, Downs SM. Chemical signals that regulate mammalian oocyte maturation. *Biol Reprod.* 1984;30(1):1–11.

65. Chen L, Russell PT, Larsen WJ. Functional significance of cumulus expansion in the mouse: Roles for the preovulatory synthesis of hyaluronic acid within the cumulus mass. *Mol Reprod Dev.* 1993;34(1):87–93.

66. Relucenti M, Heyn R, Correr S, Familiari G. Cumulus oophorus extracellular matrix in the human oocyte: a role for adhesive proteins. *Ital J Anat Embryol*. 2005;110 :219–224.

67. Høst E, Gabrielsen A, Lindenberg S, Smidt-Jensen S. Apoptosis in human cumulus cells in relation to zona pellucida thickness variation, maturation stage, and cleavage of the corresponding oocyte after intracytoplasmic sperm injection. *Fertil Steril.* 2002;77(3):511–5.

68. Yang Y, Zhang Y, Li Y. Ultrastructure of human oocytes of different maturity stages and the alteration during in vitro maturation. *Fertil Steril.* 2009 ;92(1):396.e1–6.

69. Wells D, Fragouli E, Bianchi V, Borini A, Patrizio P. Identification of novel non-invasive biomarkers of oocyte aneuploidy. *Fertil Steril.* 2008 ;90:S35.

70. Wang Q, Frolova AI, Purcell S, Adastra K, Schoeller E, Chi MM, et al. Mitochondrial Dysfunction and Apoptosis in Cumulus Cells of Type I Diabetic Mice. *PLoS One.* 2010;5(12).

71. Yang Y, Zhang Y, Li Y. Ultrastructure of human oocytes of different maturity stages and the alteration during in vitro maturation. *Fertil Steril.* 2009;92(1):396.e1–396.e6.

72. Gisè Le Ouandaogo Z, Haouzi D, Assou S, Dechaud H, Kadoch IJ, De Vos J, et al. Human Cumulus Cells Molecular Signature in Relation to Oocyte Nuclear Maturity Stage. *PLoS One.* 2011;6(11):e27179.

73. Feuerstein P, Puard V, Chevalier C, Teusan R, Cadoret V, Guerif F, et al. Genomic Assessment of Human Cumulus Cell Marker Genes as Predictors of Oocyte Developmental Competence: Impact of Various Experimental Factors. Panepucci RA, editor. PLoS One. *Public Library of Science*; 2012 ;7(7):e40449.

74. Pacella L, Zander-Fox DL, Armstrong DT, Lane M. Women with reduced ovarian reserve or advanced maternal age have an altered follicular environment. *Fertil Steril.* 2012; 98(4):986-94.

75. Edwards RG. Follicular fluid. *J Reprod Fertil* . 1974;37(1):189–219.

76. Al-Edani T, Assou S, Ferrières A, Bringer Deutsch S, Gala A, Lecellier C-H, et al. Female Aging Alters Expression of Human Cumulus Cells Genes that Are Essential for Oocyte Quality. *Biomed Res Int.* 2014;2014:1–10.

77. Molinari E, Bar H, Pyle AM, Patrizio P. 1.55105264. *Mol Hum Reprod*. 2016;22(8):866–76.

78. Kocabas AM, Crosby J, Ross PJ, Otu HH, Beyhan Z, Can H, et al. The transcriptome of human oocytes. *Proc Natl Acad Sci*. 2006;103(38):14027–32.

79. Zhao S, Fung-Leung W-P, Bittner A, Ngo K, Liu X. Comparison of RNA-Seq and Microarray in Transcriptome Profiling of Activated T Cells. *PLoS* One. Public Library of Science; 2014;9(1):e78644.

80. Zhao S, Fung-Leung W-P, Bittner A, Ngo K, Liu X, Carstens J. Comparison of RNA-Seq and Microarray in Transcriptome Profiling of Activated T Cells. *PLoS One*. 2014 ;9(1):e78644.

81. Okoniewski MJ, Miller CJ. Hybridization interactions between probesets in short oligo microarrays lead to spurious correlations. *BMC Bioinformatics.* 2006 ;7(1):276.

82. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. 2009;10(1):57–63.

83. Li L, Xiong M. Dynamic Model for RNA-seq Data Analysis. 2015; Biomed Res Int. 2015; 916352.

84. Stegle O, Teichmann SA, Marioni JC. Computational and analytical challenges in single-cell transcriptomics. *Nat Rev Genet*. 2015;16:133–45.

85. Metzker ML. Sequencing technologies — the next generation. *Nat Rev Genet*. 2010;11(1):31–46.

86. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods.* 2008;5(7):621–8.

87. Marguerat S, Rg Bähler J. RNA-seq: from technology to biology. *Cell Mol Life Sci.* 2010 ;67(4):569-79

88. Metzker ML. Sequencing technologies- the next generation. *Nat Rev Genet.* 2010;11(1):31–46.

89. Cloonan N, Forrest ARR, Kolle G, Gardiner BBA, Faulkner GJ, Brown MK, et al. Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nat Methods*.

2008;5(7):613–9.

90. Barbazuk WB, Emrich SJ, Chen HD, Li L, Schnable PS. SNP discovery via 454 transcriptome sequencing. *Plant J.* 2007;51(5):910–8.

91. Vera JC,Wheat CW, Fescemyer HW, Frilander MJ,Ceawford DL et al. Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Mol Ecol*. 2008;17(7):1636–47.

92. Emrich SJ, Barbazuk WB, Li L, Schnable PS. Gene discovery and annotation using LCM-454 transcriptome sequencing. *Genome Res*. 2006;17(1):69–73.

93. Li Y, Thompson H, Hemphill C, Hong F, Forrester J, Johnson RH, et al. An improved one-tube RT-PCR protocol for analyzing single-cell gene expression in individual mammalian cells. *Anal Bioanal Chem*.2010;397(5):1853-9.

94. White AK, VanInsberghe M, Petriv OI, Hamidi M, Sikorski D, Marra MA, et al. High-throughput microfluidic single-cell RT-qPCR. *Proc Natl Acad Sci.* 2011;108(34):13999–4004.

95. Hartmann CH, Klein CA. Gene expression profiling of single cells on large-scale oligonucleotide arrays. *Nucleic Acids Res*. Oxford University Press; 2006;34(21):e143–e143.

96. Eberwine J, Yeh H, Miyashiro K, Cao Y, Nair S, Finnell R, et al. Analysis of gene expression in single live neurons. *Neurobiology.* 1992;89:3010–4.

97. Chiang M-K, Melton DA. Single-cell transcript analysis of pancreas development. *Dev Cell*. 2003;4(3):383–93.

98. Trimarchi JM, Stadler MB, Cepko CL. Individual Retinal Progenitor Cells Display Extensive Heterogeneity of Gene Expression. *PLoS One*. 2008;3(2):e1588.

99. Ramos CA, Bowman TA, Boles NC, Merchant AA, Zheng Y, Parra I, et al. Evidence for Diversity in Transcriptional Profiles of Single Hematopoietic Stem Cells. *PLoS Genet.* 2006 ;2(9):e159.

100. Dalerba P, Kalisky T, Sahoo D, Rajendran PS, Rothenberg ME, Leyrat AA, et al. Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat Biotechnol.* 2011;29(12):1120–7.

101. Ståhlberg A, Bengtsson M. Single-cell gene expression profiling using reverse transcription quantitative real-time PCR. *Methods.* 2010:50(4):282-8.

102. Ozsolak F, Milos PM. RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet*. 2011; 12(2):87–98.

103. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods.* 2009; 6(5):377–82.

104. Islam S, Kjällquist U, Moliner A, Zajac P, Fan J-B, Lönnerberg P, et al. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res.* 2011; 21(7):1160–7.

105. Saliba A-E, Westermann AJ, Gorski SA, Vogel J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res*. 2014; 42(14):8845-60.

106. Shapiro E, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet.* 2013 ;14(9):618–30.

107. Yamada-Fukunaga T, Yamada M, Hamatani T, Chikazawa N, Ogawa S, Akutsu H, et al. Age-associated telomere shortening in mouse oocytes. *Reprod Biol Endocrinol.* 2013;11:108.

108. Hirshfield AN. Heterogeneity of cell populations that contribute to the formation of primordial follicles in rats. *Biol Reprod*. 1992;47(3):466–72.

109. Zheng W, Zhang H, Gorre N, Risal S, Shen Y, Liu K. Two classes of ovarian primordial follicles exhibit distinct developmental dynamics and physiological functions. *Hum Mol Genet.* 2014;23(4):920–8.

110. Liu M, Yin Y, Ye X, Zeng M, Zhao Q, Keefe DL, et al. Resveratrol protects against age-associated infertility in mice. *Hum Reprod.* 2013;28(3):707–17.

111. Yuan R, Peters LL, Paigen B. Mice as a mammalian model for research on the genetics of aging. *ILAR J.* 2011; 52(1):4–15.

112. Kõks S, Dogan S, Tuna BG, González-Navarro H, Potter P, Vandenbroucke RE. Mouse models of ageing and their relevance to disease. *Mech Ageing Dev*. 2016;160:41–53.

113. Hamatani T. Age-associated alteration of gene expression patterns in mouse oocytes.

Hum Mol Genet. 2004;13(19):2263–78.

114. Shomper M, Lappa C, FitzHarris G. Kinetochore microtubule establishment is defective in oocytes from aged mice. *Cell Cycle.* 2014;13(7):1171–9.

115. Lister LM, Kouznetsova A, Hyslop LA, Kalleas D, Pace SL, Barel JC, et al. Age-Related Meiotic Segregation Errors in Mammalian Oocytes Are Preceded by Depletion of Cohesin and Sgo2. *Curr Biol*. 2010 (17):1511–21.

116. Duncan FE, Hornick JE, Lampson M a., Schultz RM, Shea LD, Woodruff TK. Chromosome cohesion decreases in human eggs with advanced maternal age. *Aging Cell.* 2012;11:1121–4.

117. Chiang T, Duncan FE, Schindler K, Schultz RM, Lampson MA. Evidence that Weakened Centromere Cohesion Is a Leading Cause of Age-Related Aneuploidy in Oocytes. *Curr Biol.* 2010;20(17):1522–8

118. Ramsköld D, Luo S, Wang Y-C, Li R, Deng Q, Faridani OR, et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol.* 2012;30(8):777–82.

119. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods.* 2009 ;6(5):377–82.

120. Sasagawa Y, Nikaido I, Hayashi T, Danno H, Uno KD, Imai T, et al. Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. *Genome Biol.* 2013;14(4):3097.

121. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A survey of best practices for RNA-seq data analysis. *Genome Biol.* 2016;17(1):13.

122. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013; 29(1):15–21.

123. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009; 25(16):2078–9.

124. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014; 30(7):923–30.

125. Liao Y, Smyth GK, Shi W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* 2013; 41(10):e108.

126. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010; 26(1):139–40.

127. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2009;10(1):57–63.

128. Marguerat S, Bähler J. RNA-seq: from technology to biology. *Cell Mol Life Sci* 2010; 67(4):569–79.

129. Liu S, Trapnell C. Single-cell transcriptome sequencing: recent advances and remaining challenges. F1000Res. 2016;17;Rev-182.

130. 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(2):103–14.

131. Tang F, Barbacioru C, Nordman E, Li B, Xu N, Bashkirov VI, et al. RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat Protoc.* 2010; 5(3):516–35.

132. Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. *Nat Rev Genet* 2009; 10(2):94–108.

133. Taft RJ, Glazov EA, Cloonan N, Simons C, Stephen S, Faulkner GJ, et al. Tiny RNAs associated with transcription start sites in animals. *Nat Genet*.2009; 41(5):572–8.

134. Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, Flynn RA, et al. Divergent transcription from active promoters. *Science*. 2008; 322(5909):1849–51.

135. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A survey of best practices for RNA-seq data analysis. *Genome Biol.* 2016; 17(1):13.

136. Zhong JF, Chen Y, Marcus JS, Scherer A, Quake SR, Taylor CR, et al. A microfluidic processor for gene expression profiling of single human embryonic stem cells. *Lab Chip* .2008; 8(1):68–74.

137. Wang Z, Gerstein M, Snyder M. RNA-Seq : a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2010;10(1):57–63.

138. Saliba A-E, Westermann AJ, Gorski SA, Vogel J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res*. 2014;42(14):8845–60.

139. 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(2):103–14.

140. Griffin J, Emery BR, Huang I, Peterson M, Carrell DT. Comparative analysis of follicle morphology and oocyte diameter in four mammalian species (mouse, hamster, pig, and human). *J Exp Clin Assist Reprod.* 2006; 3:2.

141. Hartshorn C, Anshelevich A, Wangh LJ. Rapid, single-tube method for quantitative preparation and analysis of RNA and DNA in samples as small as one cell. *BMC Biotechnol.* 2005;5(1):2.

142. Pan X, Durrett RE, Zhu H, Tanaka Y, Li Y, Zi X, et al. Two methods for full-length RNA sequencing for low quantities of cells and single cells. *Proc Natl Acad Sci U S A*. 2013;110(2):594–9.

143. Head SR, Komori HK, Lamere SA, Whisenant T, Nieuwerburgh F Van, Salomon DR, et al. Library construction for next-generation sequencing: Overviews and challenges. *Biotechniques.* 2014;56(2):61-4

144. Bhargava V, Head SR, Ordoukhanian P, Mercola M, Subramaniam S. Technical Variations in Low-Input RNA-seq Methodologies. *Sci Rep.* 2015;4(1):3678.

145. Aird D, Ross MG, Chen W-S, Danielsson M, Fennell T, Russ C, et al. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol.* 2011;12(2):R18.

146. Oyola SO, Otto TD, Gu Y, Maslen G, Manske M, Campino S, et al. Optimizing illumina next-generation sequencing library preparation for extremely at-biased genomes. *BMC Genomics.* 2012;13(1):1.

147. Head SR, Komori HK, LaMere SA, Whisenant T, Van Nieuwerburgh F, Salomon DR, et al. Library construction for next-generation sequencing: overviews and challenges. *Biotechniques*.2014;56(2):61–4, 66, 68.

148. Adiconis X, Borges-Rivera D, Satija R, DeLuca DS, Busby MA, Berlin AM, et al. Comparative analysis of RNA sequencing methods for degraded or low-input samples. *Nat Methods.* 2013;10(7):623–9.

149. Picelli S, Faridani OR, Björklund ÅK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc.* 2014;9(1):171–81.

150. Pan W, Lin J, Le CT. How many replicates of arrays are required to detect gene expression changes in microarray experiments? A mixture model approach. *Genome Biol*. 2002;3(5). research0022.1–research0022.10

151. Kvam VM, Liu P, Si Y. A comparison of statistical methods for detecting differentially expressed genes from RNA-seq data. *Am J Bot*. 2012;99(2):248–56.

152. Liu Y, Zhou J, White KP. RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics.* 2014;30(3):301–4.

153. Fang Z, Cui X. Design and validation issues in RNA-seq experiments. *Brief Bioinform* 2011;12(3):280–7.

154. Peters H. The development of the mouse ovary from birth to maturity. *Acta Endocrinol.* 1969;62(1):98–116.

155. Eppig JJ. Coordination of nuclear and cytoplasmic oocyte maturation in eutherian mammals. *Reprod Fertil Dev.* 1996;8(4):485–9.

156. De Smedt V, Crozet N, Gall L. Morphological and functional changes accompanying the acquisition of meiotic competence in ovarian goat oocyte. *J Exp Zool.* 1994;269(2):128–39.

157. Romaguera R, Moll X, Morató R, Roura M, Palomo MJ, Catalá MG, et al. Prepubertal goat oocytes from large follicles result in similar blastocyst production and embryo ploidy than those from adult goats. *Theriogenology*. 2011;76(1):1–11.

158. Bagg MA, Nottle MB, Armstrong DT, Grupen CG. Relationship between follicle size and oocyte developmental competence in prepubertal and adult pigs. *Reprod Fertil Dev*. 2007;19(7):797–803.

159. Romaguera R, Moll X, Morató R, Roura M, Palomo MJ, Catalá MG, et al. Prepubertal goat oocytes from large follicles result in similar blastocyst production and embryo ploidy than those from adult goats. *Theriogenology.* 2011:76(1):1-11.

160. Kong Q, Xie B, Li J, Huan Y, Huang T, Wei R, et al. Identification and characterization of an oocyte factor required for porcine nuclear reprogramming. *J Biol Chem* 2014;289(10):6960–8.

161. Ptak G, Matsukawa K, Palmieri C, Salda L Della, Scapolo PA, Loi P. Developmental and functional evidence of nuclear immaturity in prepubertal oocytes. *Hum Reprod.* 2006 ;21(9):2228–37.

162. Hodgman R, Tay J, Mendez R, Richter JD. CPEB phosphorylation and cytoplasmic polyadenylation are catalyzed by the kinase IAK1/Eg2 in maturing mouse oocytes. *Development* . 2001;128(14):2815–22.

163. Presicce GA, Jiang S, Simkin M, Zhang L, Looney CR, Godke RA, et al. Age and hormonal dependence of acquisition of oocyte competence for embryogenesis in prepubertal calves. *Biol Reprod*. 1997;56(2):386–92.

164. Hayakawa K, Ohgane J, Tanaka S, Yagi S, Shiota K. Oocyte-specific linker histone H1foo is an epigenomic modulator that decondenses chromatin and impairs pluripotency. *Epigenetics*. 2012;7(9):1029–36.

165. Su Y-Q, Wu X, O'Brien MJ, Pendola FL, Denegre JN, Matzuk MM, et al. Synergistic roles of BMP15 and GDF9 in the development and function of the oocyte–cumulus cell complex in mice: genetic evidence for an oocyte–granulosa cell regulatory loop. *Dev Biol.* 2004;276(1):64–73.

166. Liang L-F, Dean J. Conservation of Mammalian Secondary Sperm Receptor Genes Enables the Promoter of the Human Gene to Function in Mouse Oocytes. *Dev Biol.* 1993;156(2):399–408.

167. Wu X, Viveiros MM, Eppig JJ, Bai Y, Fitzpatrick SL, Matzuk MM. Zygote arrest 1 (Zar1) is a novel maternal-effect gene critical for the oocyte-to-embryo transition. *Nat Genet.* 2003;33(2):187–91.

168. Demiray SB, Yilmaz O, Goker ENT, Tavmergen E, Calimlioglu N, Sezerman U, et al. Expression of the Bone Morphogenetic Protein-2 (BMP2) in the Human Cumulus Cells as a Biomarker of Oocytes and Embryo Quality. *J Hum Reprod Sci* . 2017;10(3):194–200.

169. Nakanishi M, Ando H, Watanabe N, Kitamura K, Ito K, Okayama H, et al.

Identification and characterization of human Wee1B, a new member of the Wee1 family of Cdk-inhibitory kinases. *Genes Cells.* 2000 ;5(10):839–47.

170. Saskova A, Solc P, Baran V, Kubelka M, Schultz RM, Motlik J. Aurora kinase A controls meiosis I progression in mouse oocytes. Cell Cycle. 2008 ;7(15):2368–76.

171. Qu Y, Lu D, Jiang H, Chi X, Zhang H. EZH2 is required for mouse oocyte meiotic maturation by interacting with and stabilizing spindle assembly checkpoint protein BubRI. *Nucleic Acids Res.* 2016 ;44(16):7659–72.

172. Tung JJ, Hansen D V., Ban KH, Loktev A V., Summers MK, Adler JR, et al. A role for the anaphase-promoting complex inhibitor Emi2/XErp1, a homolog of early mitotic inhibitor 1, in cytostatic factor arrest of Xenopus eggs. *Proc Natl Acad Sci.* 2005;102(12):4318–23.

173. Solc P, Saskova A, Baran V, Kubelka M, Schultz RM, Motlik J. CDC25A phosphatase controls meiosis I progression in mouse oocytes. *Dev Biol.* NIH Public Access; 2008;317(1):260–9.

174. Masciarelli S, Horner K, Liu C, Park SH, Hinckley M, Hockman S, et al. Cyclic nucleotide phosphodiesterase 3A-deficient mice as a model of female infertility. J Clin Invest [Internet]. American Society for Clinical Investigation; 2004 Jul [cited 2018 Jan 8];114(2):196–205.

175. Lucifero D, La Salle S, Bourc'his D, Martel J, Bestor TH, Trasler JM. Coordinate regulation of DNA methyltransferase expression during oogenesis. *BMC Dev Biol*. 2007;7(1):36.

176. Dorji, Ohkubo Y, Miyoshi K, Yoshida M. Gene expression profile differences in embryos derived from prepubertal and adult Japanese Black cattle during in vitro development. *Reprod Fertil Dev.* 2012;24(2):370.

177. Wang F, Zhang L, Duan X, Zhang G-L, Wang Z-B, Wang Q, et al. RhoA-mediated FMNL1 regulates GM130 for actin assembly and phosphorylates MAPK for spindle formation in mouse oocyte meiosis. *Cell Cycle.* 2015;14(17):2835-43.

178. Shibuya EK, Boulton TG, Cobb MH, Ruderman J V. Activation of p42 MAP kinase and the release of oocytes from cell cycle arrest. *EMBO J.* European Molecular Biology Organization. 1992;11(11):3963–75.

179. Mcginnis LA, Lee HJ, Robinson DN, Evans JP. MAPK3/1 (ERK1/2) and Myosin Light Chain Kinase in Mammalian Eggs Affect Myosin-II Function and Regulate the Metaphase II State in a Calcium-and Zinc-Dependent Manner. *Biol Reprod*. 2015 ;92146(6):1–14.

180. Tang P, Sharpe CR, Mohun TJ, Wylie CC. Vimentin expression in oocytes, eggs and early embryos of Xenopus laevis. *Development*. 1988;103:279–87.

181. Fénichel P, Durand-Clément M. Role of integrins during fertilization in mammals. *Hum Reprod*. 1998;13 Suppl 4:31–46.

182. Moreno-Layseca P, Streuli CH. Signalling pathways linking integrins with cell cycle progression. *Matrix Biol*. 2014;34:144–53.

183. Liu Z, Shimada M, Richards JS. The involvement of the Toll-like receptor family in ovulation. *J Assist Reprod Genet.* 2008 ;25(6):223–8.

184. Sasaki K1, Tanaka A1, Nagatomo H1, Ogawa H2, Kobayashi K et al. Characterization of Toll-Like Receptor 9 expression during Mouse Preimplantation Development. *J Genit Syst Disord.* 2016;05(01).

185. Das BC, Thapa P, Karki R, Das S, Mahapatra S, Liu T-C, et al. Retinoic acid signaling pathways in development and diseases. *Bioorg Med Chem*. 2014;22(2):673–83.

186. Bowles J, Feng C-W, Miles K, Ineson J, Spiller C, Koopman P. ALDH1A1 provides a source of meiosis-inducing retinoic acid in mouse fetal ovaries. *Nat Commun.* 2016;7:10845.

187. Li H, Clagett-Dame M. Vitamin A Deficiency Blocks the Initiation of Meiosis of Germ Cells in the Developing Rat Ovary In Vivo. *Biol Reprod*. Oxford University Press; 2009. 81(5):996–1001.

188. Downs SM, Mosey JL, Klinger J. Fatty acid oxidation and meiotic resumption in mouse oocytes. *Mol Reprod Dev.* 2009;76(9):844–53.

189. Zolfaghari R, Ross AC. Recent advances in molecular cloning of fatty acid desaturase genes and the regulation of their expression by dietary vitamin A and retinoic acid. Prostaglandins, Leukot Essent Fat Acids. 2003 ;68(2):171–9.

190. Mogas T, Palomo M, Paramio M. Developmental capacity of in vitro matured and fertilized oocytes from prepuperal and adult goats. *Theriogenology*. 47(6):1189-203.

191. Te Velde ER, Pearson PL. The variability of female reproductive ageing. *Hum Reprod Update.* 2002;8(2):141–54.

192. Ottolenghi C1, Uda M, Hamatani T, Crisponi L, Garcia JE, Ko M et al Aging of Oocyte, Ovary, and Human Reproduction. *Ann N Y Acad Sci.* 2004 ;1034(1):117–31.

193. Duncan FE, Chiang T, Schultz RM, Lampson MA. Evidence that a defective spindle assembly checkpoint is not the primary cause of maternal age-associated aneuploidy in mouse eggs. *Biol Reprod*. 2009 ;81(4):768–76.

194. Chiang T, Duncan FE, Schindler K, Schultz RM, Lampson MA. Evidence that weakened centromere cohesion is a leading cause of age-related aneuploidy in oocytes. *Curr Biol.* 2010;20(17):1522–8.

195. Fragouli E, Wells D, Whalley KM, Mills J a, Faed MJW, Delhanty JD a. Increased susceptibility to maternal aneuploidy demonstrated by comparative genomic hybridization analysis of human MII oocytes and first polar bodies. *Cytogenet Genome Res.* 2006;114(1):30–8.

196. Grøndahl ML, Yding Andersen C, Bogstad J, Nielsen FC, Meinertz H, Borup R. Gene expression profiles of single human mature oocytes in relation to age. *Hum Reprod.* 2010;25(4):957–68.

197. Su Y-Q, Sugiura K, Woo Y, Wigglesworth K, Kamdar S, Affourtit J, et al. Selective degradation of transcripts during meiotic maturation of mouse oocytes. *Dev Biol.* 2007;302(1):104–17.

198. Griswold MD, Hogarth CA, Bowles J, Koopman P. Initiating Meiosis: The Case for Retinoic Acid1. *Biol Reprod*. 2012;86(2)35.

199. Thompson JN. Vitamin A in Development of the Embryo. *Am J Clin Nutr.* 1969; 22(8):1063–9.

200. Ulven SM, Gundersen TE, Weedon MS, Landaas VØ, Sakhi AK, Fromm SH, *et al.* Identification of Endogenous Retinoids, Enzymes, Binding Proteins, and Receptors during Early Postimplantation Development in Mouse: Important Role of Retinal Dehydrogenase Type 2 in Synthesis of All-trans-Retinoic Acid. *Dev Biol.* 2000 15;220(2):379-91.

201. Livingston TE, The Effect of Retinol During In Viv0 and In Vitro Oocyte Maturation

and Embryonic Development. 2003.

202. Adolfi MC, Herpin A, Regensburger M, Sacquegno J, Waxman JS, Schartl M. Retinoic acid and meiosis induction in adult versus embryonic gonads of medaka. *Sci Rep.* 2016;6(1):34281.

203. Kikuchi K, Naito K, Noguchi J, Kaneko H, Tojo H. Maturation/M-Phase Promoting Factor Regulates Aging of Porcine Oocytes Matured *In Vitro*. *Cloning Stem Cells* 2002;4(3):211–22.

204. Kumar S, Sandell LL, Trainor PA, Koentgen F, Duester G. Alcohol and aldehyde dehydrogenases: retinoid metabolic effects in mouse knockout models. Biochim *Biophys Acta.* 2012;1821(1):198–205.

205. Molotkov A, Deltour L, Foglio MH, Cuenca AE, Duester G. Distinct retinoid metabolic functions for alcohol dehydrogenase genes Adh1 and Adh4 in protection against vitamin A toxicity or deficiency revealed in double null mutant mice. *J Biol Chem*. 2002 ;277(16):13804–11.

206. Billings SE, Pierzchalski K, Butler Tjaden NE, Pang X-Y, Trainor PA, Kane MA, et al. The retinaldehyde reductase DHRS3 is essential for preventing the formation of excess retinoic acid during embryonic development. *FASEB J.* 2013;27(12):4877–89.

207. Billings SE, Pierzchalski K, Butler Tjaden NE, Pang X-Y, Trainor PA, Kane MA, et al. The retinaldehyde reductase DHRS3 is essential for preventing the formation of excess retinoic acid during embryonic development. *FASEB J.* The Federation of American Societies for Experimental Biology; 2013;27(12):4877–89.

208. Hahn S, Backhaus M, Broecker-Preuss M, Tan S, Dietz T, Kimmig R, et al. Retinol-binding protein 4 levels are elevated in polycystic ovary syndrome women with obesity and impaired glucose metabolism. *Eur J Endocrinol*. 2007;157(2):201–7.

209. Makimura H, Wei J, Dolan-Looby SE, Ricchiuti V, Grinspoon S. Retinol-binding protein levels are increased in association with gonadotropin levels in healthy women. *Metabolism*. 2009;58(4):479–87.

210. Livingston T, Eberhardt D, Edwards JL, Godkin J. Retinol improves bovine embryonic development in vitro. *Reprod Biol Endocrinol.* 2004;2:83.

211. Whaley SL, Hedgpeth VS, Farin CE, Martus NS, Jayes FCL, Britt JH. Influence of vitamin A injection before mating on oocyte development, follicular hormones, and ovulation in gilts fed high-energy diets. *J Anim Sci.* 2000;78:1598–607.

212. Almiñana C, Gil MA, Cuello C, Caballero I, Roca J, Vazquez JM, et al. In vitro maturation of porcine oocytes with retinoids improves embryonic development. *Reprod Fertil Dev.* 2008;20(4):483–9.

213. Reader KL, Stanton J-AL, Juengel JL. The Role of Oocyte Organelles in Determining Developmental Competence. *Biology.* 2017; 6(3):35.

214. Van Blerkom J. Mitochondrial function in the human oocyte and embryo and their role in developmental competence. *Mitochondrion.* 2011;11(5):797–813.

215. Nguyen D, Samson SL, Reddy VT, Gonzalez E V., Sekhar R V. Impaired mitochondrial fatty acid oxidation and insulin resistance in aging: novel protective role of glutathione. *Aging Cell.* 2013;12(3):415–25.

216. Green CL, Mitchell SE, Derous D, Wang Y, Chen L, Han J-DJ, et al. The effects of graded levels of calorie restriction: IX. Global metabolomic screen reveals modulation of carnitines, sphingolipids and bile acids in the liver of C57BL/6 mice. *Aging Cell*. 2017;16(3):529–40.

217. Xu X-Y, Hu J-P, Wu M-M, Wang L-S, Fang N-Y. CCAAT/enhancer-binding protein CEBP-2 controls fat consumption and fatty acid desaturation in Caenorhabditis elegans. *Biochem Biophys Res Commun.* 2015;468:312–8.

218. Rahman SM, Janssen RC, Choudhury M, Baquero KC, Aikens RM, De La Houssaye BA *et al.* CCAAT/enhancer binding protein expression regulates dietary-induced inflammation in macrophages and adipose tissue in mice. *J Biol Chem*. 2012 ;287(41):34349-60.

219. Berkholtz CB, Lai BE, Woodruff TK, Shea LD. Distribution of extracellular matrix proteins type I collagen, type IV collagen, fibronectin, and laminin in mouse folliculogenesis. *Histochem Cell Biol.* 2006;126(5):583–92.

220. Lu L, Guo J, Yue Hua |, Huang | Kevin, Magaye R, Cornell J, et al. Cardiac fibrosis in the ageing heart : Contributors and mechanisms. *Clin Exp Pharmacol Physiol*. 2017 ;44:55–63.

221. Hirshfeld-Cytron JE, Duncan FE, Xu M, Jozefik JK, Shea LD, Woodruff TK. Animal age, weight and estrus cycle stage impact the quality of in vitro grown follicles. *Hum Reprod.* 2011;26(9):2473–85.

222. Briley SM, Jasti S, McCracken JM, Hornick JE, Fegley B, Pritchard MT, et al. Reproductive age-associated fibrosis in the stroma of the mammalian ovary. Reproduction. 2016; 152(3):245–60.

223. Khan SN, Shaeib F, Kavdia M, Abu-Soud HM. Myeloperoxidase and activated macrophages have differing effects on oocyte spindle morphology. *Fertil Steril.* 2016;105(2):e10.

224. Rosselli M, Keller R, Dubey R. Role of nitric oxide in the biology, physiology and pathophysiology of reproduction. *Hum Reprod Update.*1998;4(1):3–24.

225. Prasad S, Tiwari M, Pandey AN, Shrivastav TG, Chaube SK. Impact of stress on oocyte quality and reproductive outcome. *J Biomed Sci.* 2016;23:36.

226. Tokmak A, Yıldırım G, Sarıkaya E, Çınar M, Boğdaycıoğlu N, Yılmaz FM, et al. Increased oxidative stress markers may be a promising indicator of risk for primary ovarian insufficiency: a cross-sectional case control study. *Rev Bras Ginecol Obstet.* 2015;37(9):411–6.

227. Khan SN, Shaeib F, Najafi T, Kavdia M, Gonik B, Saed GM, et al. Diffused Intra-Oocyte Hydrogen Peroxide Activates Myeloperoxidase and Deteriorates Oocyte Quality. PLoS One. 2015;10(7):e0132388.

228. Victor VM, Rovira-Llopis S, Bañuls C, Diaz-Morales N, Martinez de Marañon A, Rios-Navarro C, *et al.* Insulin Resistance in PCOS Patients Enhances Oxidative Stress and Leukocyte Adhesion: Role of Myeloperoxidase. PLoS One. Public Library of Science; 2016;11(3):e0151960.

229. Heinecke JW, Goldberg IJ. Myeloperoxidase: a therapeutic target for preventing insulin resistance and the metabolic sequelae of obesity? Diabetes [Internet]. American Diabetes Association; 2014 ;63(12):4001–3.

230. Ademowo H, Dias D, Burton H, Griffiths O. Lipid (per) oxidation in mitochondria: an emerging target in the ageing process? *Biogerontology.* 2017;18(6), 859–87.

231. Eppig JJ, Wigglesworth K, Pendola FL. The mammalian oocyte orchestrates the rate

of ovarian follicular development. *Proc Natl Acad Sci U S A.* 2002;99(5):2890–4.

232. Matzuk MM, Burns KH, Viveiros MM, Eppig JJ. Intercellular communication in the mammalian ovary: oocytes carry the conversation. Science. 2002;296(5576):2178–80.

233. Huang Z, Wells D. The human oocyte and cumulus cells relationship: New insights from the cumulus cell transcriptome. Mol Hum Reprod. 2010;16(10):715–25.

234. Su Y-Q, Sugiura K, Wigglesworth K, O'Brien MJ, Affourtit JP, Pangas SA, *et al.* Oocyte regulation of metabolic cooperativity between mouse cumulus cells and oocytes: BMP15 and GDF9 control cholesterol biosynthesis in cumulus cells. *Development* 2008;135(1):111–21.

235. Fragouli E, Lalioti MD, Wells D. The transcriptome of follicular cells: Biological insights and clinical implications for the treatment of infertility. *Hum Reprod Update.* 2014;20(1):1–11.

236. Lourenço B, Sousa AP, Almeida-Santos T, Ramalho-Santos J. Relation of cumulus cell status with single oocyte maturity, fertilization capability and patient age. *J Reprod Infertil.* 2014;15(1):15–21.

237. Nakahara K, Saito H, Saito T, Ito M, Ohta N, Sakai N, et al. Incidence of apoptotic bodies in membrana granulosa of the patients participating in an in vitro fertilization program. *Fertil Steril.* 1997;67(2):302–8.

238. Tatone C, Amicarelli F. The aging ovary--the poor granulosa cells. *Fertil Steril.* 2013;99(1):12–7.

239. Hirshfield2 AN. Heterogeneity of Cell Populations That Contribute to the Formation of Primordial Follicles in Rats1. *Biol Reprod.* 1992;47:466–72.

240. Viveiros MM, O 'brien M, Wigglesworth K, Eppig JJ. Characterization of Protein Kinase C-ʏ in Mouse Oocytes Throughout Meiotic Maturation and Following Egg Activation 1. *Biol Reprod.* 2003;69:1494–9.

241. Nikitina L, Wenger F, Baumann M, Surbek D, Körner M, Albrecht C. Expression and localization pattern of ABCA1 in diverse human placental primary cells and tissues. 2011; Placenta. 2011;32(6):420-30.

242. Slee RB, Hillier SG, Largue P, Harlow CR, Miele G, Clinton M. Differentiation-

Dependent Expression of Connective Tissue Growth Factor and Lysyl Oxidase Messenger Ribonucleic Acids in Rat Granulosa Cells. *Endocrinology.* 2001 142(3):1082–9.

243. Harlow CR, Bradshaw AC, Rae MT, Shearer KD, Hillier SG. Oestrogen formation and connective tissue growth factor expression in rat granulosa cells. *J Endocrinol*. 2007;192(1):41–52.

244. Nagashima T, Kim J, Li Q, Lydon JP, DeMayo FJ, Lyons KM, et al. Connective Tissue Growth Factor Is Required for Normal Follicle Development and Ovulation. *Mol Endocrinol.* 2011;25(10):1740–59.

245. Perez GI, Tilly JL. Cumulus cells are required for the increased apoptotic potential in oocytes of aged mice. *Hum Reprod.* 1997;12(12):2781–3.

246. Pugh CW, Ratcliffe PJ. Regulation of angiogenesis by hypoxia: role of the HIF system. *Nat Med.* 2003;9(6):677–84.

247. Gérard N, Caillaud M, Martoriati A, Goudet G, Lalmanach A-C. The interleukin-1 system and female reproduction. *J Endocrinol.* 2004;180:203–12.

248. Grøndahl ML, Andersen CY, Bogstad J, Borgbo T, Hartvig Boujida V, Borup R. Specific genes are selectively expressed between cumulus and granulosa cells from individual human pre-ovulatory follicles. *Mol Hum Reprod.* 2012;18(12):572-84.

249. Granot I, Dekel N. Cell-to-cell communication in the ovarian follicle: developmental and hormonal regulation of the expression of connexin43. *Hum Reprod*. 1998;13(4):85-97.

250. Løset M, Mundal SB, Johnson MP, Fenstad MH, Freed KA, Lian IA, et al. A transcriptional profile of the decidua in preeclampsia. *Am J Obstet Gynecol.* 2011 ;204(1):84.e1–84.e27.

251. Tello D, Balsa E, Acosta-Iborra B, Fuertes-Yebra E, Elorza A, Ordóñez Á, et al. Induction of the mitochondrial NDUFA4L2 protein by HIF-1α decreases oxygen consumption by inhibiting Complex I activity. *Cell Metab.* 2011;14(6):768–79.

252. Mcreynolds S, Dzieciatkowska M, Mccallie BR, Mitchell SD, Stevens J, Hansen K, et al. Impact of maternal aging on the molecular signature of human cumulus cells. Fertil Steril. 2012;(6):1574-80.e5

253. Minton DR, Fu L, Mongan NP, Shevchuk MM, Nanus DM, Gudas LJ. Biology of Human Tumors Role of NADH Dehydrogenase (Ubiquinone) 1 Alpha Subcomplex 4-Like 2 in Clear Cell Renal Cell Carcinoma. *Clin Cancer Res.* 2016;22(11):2791-801.

254. Nolan T, Hands RE, Bustin SA. Quantification of mRNA using real-time RT-PCR. *Nat Protoc*. 2006;1(3):1559–82.

255. Steuerwald N, Cohen J, Herrera RJ, Brenner CA. Quantification of mRNA in single oocytes and embryos by real-time rapid cycle fluorescence monitored RT-PCR. *Mol Hum Reprod.* 2000;6(5):448–53.

256. Larkin JE, Frank BC, Gavras H, Sultana R, Quackenbush J. Independence and reproducibility across microarray platforms. Nat Methods [Internet]. 2005;2(5):337–44.

257. Morrison T, Hurley J, Garcia J, Yoder K, Katz A, Roberts D, et al. Nanoliter high throughput quantitative PCR. *Nucleic Acids Res.* 2006;34(18):e123.

258. Grigorenko E, Fisher C, Patel S, Winkelman V, Williamson P, Chancey C, et al. Highly Multiplex Real-Time PCR–Based Screening for Blood-Borne Pathogens on an OpenArray Platform. *J Mol Diagnostics.* 2017;19(4):549–60.

259. Almeida TA, Quispe-Ricalde A, Montes de Oca F, Foronda P, Hernández MM. A high-throughput open-array qPCR gene panel to identify housekeeping genes suitable for myometrium and leiomyoma expression analysis. *Gynecol Oncol.* 2014;134(1):138–43.

260. Gendelman M, Roth Z. Seasonal Effect on Germinal Vesicle-Stage Bovine Oocytes Is Further Expressed by Alterations in Transcript Levels in the Developing Embryos Associated with Reduced Developmental Competence1. *Biol Reprod.* 2012;86(1):1–9.

261. Mamo S, Carter F, Lonergan P, Leal CL, Al Naib A, McGettigan P, et al. Sequential analysis of global gene expression profiles in immature and in vitro matured bovine oocytes: potential molecular markers of oocyte maturation. *BMC Genomics.* 2011;16;12(1):151.

262. Mamo S, Gal AB, Bodo S, Dinnyes A. Quantitative evaluation and selection of

reference genes in mouse oocytes and embryos cultured in vivo and in vitro. *BMC Dev Biol.* 2007;7:14.

263. Paynton B V, Rempel R, Bachvarova R. Changes in state of adenylation and time course of degradation of maternal mRNAs during oocyte maturation and early embryonic development in the mouse. *Dev Biol.* 1988;129(2):304–14.

264. Biase FH, Fonseca Merighe GK, Santos Biase WKF, Martelli L, Meirelles FV. Global poly(A) mRNA expression profile measured in individual bovine oocytes and cleavage embryos. *Zygote.* 2008;16(01):29–38.

265. Lloyd RE, Romar R, Matas C, Gutierrez-Adan A, Holt W V, Coy P. Effects of oviductal fluid on the development, quality, and gene expression of porcine blastocysts produced in vitro. Reproduction [Internet]. 2009 Apr 1 [cited 2018 Apr 8];137(4):679–87.

266. Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods.* 2001;25(4):402–8.

267. Jeong YJ, Shin HS, Choi HW, Cui XS, Kim NH, Jun JH. Optimization of real time RT-PCR methods for the analysis of gene expression in single oocytes and embryos. *Fertil Steril.* 2004;82:S276–7.

268. Stå A, Kubista M. The workflow of single-cell expression profiling using quantitative real-time PCR. *Expert Rev Mol Diagn.* 2014;14(3):323–31.

269. Huggett J, Dheda K, Bustin S, Zumla A. Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun.* 2005;6(4):279–84.

270. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 2002;3(7):RESEARCH0034.

271. Broeders S, Huber I, Grohmann L, Berben G, Taverniers I, Mazzara M, et al. Guidelines for validation of qualitative real-time PCR methods. *Trends Food Sci Technol.* 2014 ;37(2):115–26.

272. Ståhlberg A, Kubista M. The workflow of single-cell expression profiling using quantitative real-time PCR. *Expert Rev Mol Diagn.* 2014 ;14(3):323.

273. Fragouli E, Bianchi V, Patrizio P, Obradors a., Huang Z, Borini a., et al. Transcriptomic profiling of human oocytes: Association of meiotic aneuploidy and altered oocyte gene expression. *Mol Hum Reprod.* 2010;16(8):570–82.

274. Grøndahl ML, Andersen CY, Bogstad J, Borgbo T, Hartvig Boujida V, Borup R. Specific genes are selectively expressed between cumulus and granulosa cells from individual human pre-ovulatory follicles. *MHR Basic Sci Reprod Med*. 2012;18(12):572–84.

275. Bunel A, Nivet AL, Blondin P, Vigneault C, Richard FJ, Sirard MA. Cumulus cell gene expression associated with pre-ovulatory acquisition of developmental competence in bovine oocytes. *Reprod Fertil Dev.* 2014;26(6):855.

276. O'Connor T, Wilmut I, Taylor J. Quantitative Evaluation of Reference Genes for Real-Time PCR During In Vitro Maturation of Ovine Oocytes. *Reprod Domest Anim*. 2013;48(3):477–83.

277. Taylor S, Wakem M, Dijkman G, Alsarraj M, Nguyen M. A practical approach to RT-qPCR— A practical approach to RT-qPCR-Publishing data that conform to the MIQE guidelines. *Methods.* 2010;50(4):S1–5.

278. Zhang R, Yu C, Wu R, Zhang L, Zhu L, Xu A, et al. RNA-Seq-Based Transcriptome Analysis of Changes in Gene Expression Linked to Human Pregnancy Outcome After In Vitro Fertilization-Embryo Transfer. *Reprod Sci.* 2016;23(1):134-45.

279. Liu XM, Wang YK, Liu YH, Yu XX, Wang PC, Li X, et al. Single-cell transcriptome sequencing reveals that cell division cycle 5- like protein is essential for porcine oocyte maturation. *J Biol Chem.* 2018;293(5):1767-1780.

280. Yin X, Cheng G, Guo H, Wang Q, Li Y, Zhang H. Single cell transcriptome profiling revealed differences in gene expression during oocyte maturation in Haimen white goats. *Genet Mol Res Mol Res.* 2017;16(161).

281. Wells D, Patrizio P. Gene expression profiling of human oocytes at different maturational stages and after in vitro maturation. *Am J Obstet Gynecol.* 2008;198(4):455.e1–9; discussion 455.e9–11.

282. Best MW, Wu J, Pauli SA, Kane MA, Pierzchalski K, Session DR, et al. A role for retinoids in human oocyte fertilization: regulation of connexin 43 by retinoic acid in

cumulus granulosa cells. *MHR Basic Sci Reprod Med.* 2015;21(6):527–34.

283. Suwa H, Kishi H, Imai F, Nakao K, Hirakawa T, Minegishi T. Retinoic acid enhances progesterone production via the cAMP/PKA signaling pathway in immature rat granulosa cells. *Biochem Biophys Rep.* 2016;12(8):62-67.

284. Bonet ML, Ribot J, Palou A. Lipid metabolism in mammalian tissues and its control by retinoic acid. Biochim Biophys Acta. *Mol Cell Biol Lipids.* 2012;1821(1):177–89.

285. Mok HJ, Shin H, Lee JW, Lee G-K, Suh CS, Kim KP, et al. Age-Associated Lipidome Changes in Metaphase II Mouse Oocytes. *PLoS One.* 2016;11(2):e0148577.

286. Phillipson OT. Management of the aging risk factor for Parkinson's disease. *Neurobiol Aging.* 2014;35(4):847–57.

287. Bonet ML, Ribot J, Palou A. Lipid metabolism in mammalian tissues and its control by retinoic acid. Biochim Biophys Acta - *Mol Cell Biol Lipids.* 2012;1821(1):177–89.

288. Sturmey R, Reis A, Leese H, McEvoy T. Role of Fatty Acids in Energy Provision During Oocyte Maturation and Early Embryo Development. *Reprod Domest Anim.* 2009;44(s3):50–8.

289. Dunning KR, Russell DL, Robker RL. Lipids and oocyte developmental competence: the role of fatty acids and β-oxidation. Reproduction. 2014;148(1):R15–27.

290. Charlier C, Montfort J, Chabrol O, Brisard D, Nguyen T, Le Cam A, et al. Oocyte-somatic cells interactions, lessons from evolution. *BMC Genomics.* 2012; 13:560.

291. Sanchez-Lazo L, Brisard D, Elis S, Maillard V, Uzbekov R, Labas V, et al. Fatty acid synthesis and oxidation in cumulus cells support oocyte maturation in bovine. *Mol Endocrinol.* 2014;28(9):1502–21.

292. Bouniol-Baly C, Hamraoui L, Guibert J, Beaujean N, Szöllösi MS, Debey P. Differential transcriptional activity associated with chromatin configuration in fully grown mouse germinal vesicle oocytes. *Biol Reprod*. 1999;60(3):580–7.

293. Eppig JJ, Schultz RM, O'Brien M, Chesnel F. Relationship between the Developmental Programs Controlling Nuclear and Cytoplasmic Maturation of Mouse Oocytes. *Dev Biol.* 1994;164(1):1–9.

294. Swain JE, Pool TB. ART failure: oocyte contributions to unsuccessful fertilization. *Hum Reprod Update.* 2008;14(5):431–46.

295. Dumesic DA, Meldrum DR, Katz-Jaffe MG, Krisher RL, Schoolcraft WB. Oocyte environment: follicular fluid and cumulus cells are critical for oocyte health. *Fertil Steril.* 2015;103:303–16.