8-19-2015

Gene Expression of Bovine Embryonic Development and Nuclear Reprogramming in the Mouse

Zongliang Jiang
University of Connecticut - Storrs, zongliang.jiang@uconn.edu

Gene Expression of Bovine Embryonic Development and Nuclear Reprogramming in the Mouse

Zongliang Jiang, Ph.D.

University of Connecticut, 2015

Abstract

Understanding of transcriptional machinery of early embryo development and epigenetic mechanisms of reprogramming is essential for embryogenesis and the success of animal biotechnology. This dissertation presents (1) transcriptional profiles of bovine in vivo pre-implantation development (Chapter Two); (2) bovine imprinted gene expression patterns across mammalian species (Chapter Three); (3) effects of High Hydrostatic Pressure (HHP) on expression profiles of in vitro produced embryos (Chapter Four); (4) roles of BAF complex in reprogramming (Chapter Five).

In the first section, we reported comprehensive transcriptome dynamics of single matured bovine oocytes and pre-implantation embryos developed in vivo using the RNA-seq technologies. Subsequently, using weighted gene co-expression network analysis, we found 12 stage-specific modules of co-expressed genes and identified potential master regulators of embryo development. Finally, we conducted the first comparison of embryonic expression profiles across three mammalian species, human, mouse and bovine. We found that the three species share more maternally deposited genes than embryonic genome activated genes and demonstrated that bovine embryos are better models for human embryonic development.

With the "gold standards" in vivo embryo transcriptome datasets, we characterized all currently known bovine imprinted genes and compared them to their counterparts in humans, mice and pigs in the second section. We found interesting expression patterns such as high levels of paternally expressed and low levels of maternally expressed genes in bovine oocytes. We also report drastic variations of species-specific levels of expression of imprinted genes.

In the third section, we evaluated the effects of HHP on the transcriptome of bovine IVF blastocysts by the DNA Microarray technology. We found a positive effect of HHP on bovine IVF blastocysts mediated by specific gene expression changes.

In the fourth section, we aimed to better understand the mechanism of reprogramming and to further improve its efficiency by studying the roles of somatic BAF components, *Brm* and *Baf170,* during reprogramming in the mouse model. We found that *Brm* and *Baf170* inhibit reprogramming in a stage-specific fashion. We further showed that inhibiting somatic BAF improves complete reprogramming by facilitating the activation of the "pluripotency circuitry" and by constituting to the Stat3-regulated epigenetic network during pluripotency establishment.

Title Page

Gene Expression of Bovine Embryonic Development and Nuclear Reprogramming in the Mouse

Zongliang Jiang

B.S., Guangxi University, P.R. China, 2008

M.S., University of Connecticut, U.S.A, 2013

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2015

Approval Page

Doctor of Philosophy Dissertation

Gene Expression of Bovine Embryonic Development and Nuclear Reprogramming in the Mouse

Presented by

Zongliang Jiang, B.S., M.S.

Major Advisor _____
Xiuchun (Cindy) Tian

Associate Advisor _____
Rachel J. O'neill

Associate Advisor _____
Robert A. Milvae

Associate Advisor _____
Young Tang

University of Connecticut

2015

Acknowledgments

My deepest gratitude goes to my amazing major advisor, Dr. Xiuchun (Cindy) Tian for her guidance, motivation and support. Her creativity and ingenuity have motivated me to come up with and explore my own ideas, which helped me grow as an independent scientist. I would like to say " I truly thank you for all that you have done for me, Dr. Tian, you have a place in my heart forever!"

I am grateful to my graduate committee members, Dr. Rachel O'neill, Dr. Robert Milvae and Dr. Young Tang for incredibly valuable advice and support. Their feedback and comments have been so valuable to me.

I would like to extend my sincerest thanks to my present colleagues members, Dr. Rashid Ali, Ellie Duan, Kaleigh Flock and Mingyuan Zhang for being always supportive, encouraging and celebrating my research. I am deeply indebted to many of my previous lab mates, Dr. Chih-Jen Lin, Sharon Kish, Dr. Mark Carter, Dr. Sadie Marjani, Ling Wang, Dr. Xueming Zhao and Yan Luo, who had either worked with me or trained me for the experiments. A special thank you to Dr. Young Tang, who mentored and worked with me throughout my graduate work. Graduate life was made even more memorable because of all of wonderful colleagues I got to share all the experiences with.

Foremost, I am indebted to faculty members at Guangxi University, especially Prof. Ming Zhang and Prof. Xi Gan, for inspiring me, for shaping who I am, and for paving the way for my graduate study.

I am blessed to have the kindest and most caring family. To my mom, she is a great single mom who raise me up and never doubt my decision to move 10 thousands miles away from home, despite still not fully understanding what it is that I actually do. She is my anchor

and wind that sustains my flight. I am truly blessed to be her son. Also I thank to my brother, sister-in-law, my lovely niece, and nephew. They have been staying by my side through all the challenges. All of my best qualities are due to their all.

Table of Contents

List of Common Abbreviations

AP: alkaline phosphatase

ART: assisted reproductive technology

BAF: Brg/Brahma-associated factors

BWS: Beckwith-Wiedemann Syndrome

cDNA: complementary deoxyribonucleic acid

DIC: differential interference contrast microscopy

DNA: deoxyribonucleic acid

Dnmt: DNA methyltransferase

EB: embryoid body

EGA: embryonic genome activation

ESC: embryonic stem cell

FACS: fluorescence-activated cell sorting

FPKM: fragments per kilobase of exon per million fragments mapped

GEO: gene expression Omnibus

GFP: green fluorescence protein

GFP: green fluorescence protein

GO: gene ontology

HHP: high hydrostatic pressure

HSP: heat shock protein

IETS: international embryo transfer society

iPSC: induced pluripotent stem cell

IVF: in vitro fertilization

IVP: in vitro produced

Jak: Janus kinase

Jaki: Jak inhibitor I

Klf4: Kruppel-like factor 4

LIF: leukemia inhibitory factor

LOS: large offspring syndrome

MEF: mouse embryonic fibroblast

mRNA: messenger ribonucleic acid

Oct4: octamer-binding transcription factor 4, also known as POU5F1

OKS: Oct4, Klf4, and Sox2

OKSM: Oct4, Klf4, Sox2 and c-Myc

PCA: Principal component analyses

pre-iPSC: partially reprogrammed iPSC

qRT-PCR: quantitative real time polymerase chain reaction

RNA-seq: RNA sequencing

RNA: ribonucleic acid

RPKM: reads per kilobase of exon per million fragments mapped

SD: standard deviation

shRNA: short hairpin RNA

Sox2: SRY (sex determining region Y)-box 2

Stat3: signal transducer and activator of transcription 3

SWI/SNF: SWItch/Sucrose NonFermentable

WGCNA: weighted gene co-expression network analysis

List of Tables and Figures

List of Tables

List of Figures

List of Supplementary Tables

List of Supplementary Figures

Chapter One

Introduction

1.1. Embryonic Pre-implantation Development.

Mammalian pre-implantation development is a complex process including fertilization, cleavage divisions, compaction, and blastulation. During this process, massive degradation of oocyte-stored maternal RNAs/proteins and gradual activation of the embryonic genome take place [1]. Embryonic genome activation (EGA) is initiated at species-specific stages. This is at 2-cell stage in the mouse [2-4] and between 4- and 8-cell stages in humans [5]. However, the exact timing of EGA in bovine is still debated [6-11]. The characterization of normal maternal and embryonic transcriptomes is essential for the understanding embryogenesis and has important implications for assisted reproductive technology (ART).

In vivo derived embryos from livestock species such as cattle are valuable models because, first, cattle is an economically important species for which ART has been well-developed [12]. The characterization of in vivo developed embryos can provide a reference base for those generated by various in vitro techniques. This enhances our ability to improve assisted embryo biotechnologies such as embryo culture conditions. Second, bovine embryos are increasingly used as an alternative to those of the mouse for the analysis of human pre-implantation development that is limited by a scarcity of material, ethical and legal confines.

To date, few molecular and cellular studies have been conducted using in vivo embryos due to high expense and small sample size. Through a collaboration with Xinjiang Academy of Animal Sciences, we generated bovine in vivo produced embryos at different stages and profiled the transcriptomes of these embryos by RNA-seq. The data are presented in the second chapter of this dissertation.

1.2. Transcriptome Profiling of Pre-implantation Development by Microarray and RNA Sequencing (RNA-seq) Technology

Various technologies including hybridization or sequence-based approaches have been developed to deduce and quantify embryonic transcriptomes. During the past decade and until recently, hybridization-based DNA microarray had been the most successful and widely used transcriptome analysis method [13-15]. However, microarray has significant limitations including (1) detecting only printed genes; (2) relative low detection limit; (3) variations from hybridization; and (4) inability to detect alternative splicing and novel transcripts. The introduction of state-of-the-art technology, high throughput next generation sequencing (NGS) technologies [16] such as RNA sequencing (RNA-seq) revolutionized functional genomics and potentially overcame all problems associated with the DNA microarray [17]. RNA-seq allows single-base resolution [18], unlimited capture of gene expression levels depending only on the depth of sequencing [19], gene fusion detection [20], small RNA characterization [21] and detection of alternative splicing events [22,23].

It is essential to capture transcriptome of few or even single-cell in study of early embryos. Previously, researchers have integrated either in vitro transcription-based (IVT) linear amplification [3,4,24,25], PCR-based exponential amplification [26,27], or a combination of both [28,29] into DNA microarray for small amounts of starting materials, or even a single cell. More recently, such amplification approaches have been combined with the RNA-seq system [19,30], promoting rapid advances in the field early embryonic development [31,32], and revealing the molecular constituents and pathways of totipotency and pluripotency of the early embryos.

Moreover, a number of either hybridization- or sequencing-based technologies are now emerging that allows reliable transcriptomics from minute cell quantities including

hybridization-based, the NanosString nCounter System [33] and sequencing-based, nanoCAGE [34]. In addition, PCR array and Fluidigm [35,36] offer platforms that perform quantitative real-time polymerase chain reaction (qRT-PCR) on gene panels in a multiplexed manner and has been used to profile single cells. These approaches are advantageous for the determination of the expression levels of a subset of transcripts across cells of interest.

1.3. Transcriptome and Post-transcriptome Modifications

With the complete catalogue of transcripts and spatio-temporal expression patterns of each gene during different stages of embryo development, identifying and dissecting complex biological networks that drive physiological functions are now possible. To this extent, we are set forth to mining specific categories of gene activities in order to understand their roles in early embryos. These specific subsets of genes include those associated with genomic imprinting, X-chromosome compensation,, oxidative phosphorylation, epigenetic modifications of histones, etc..

Genomic imprinting, parent-of-origin-dependent gene expression [37], plays important roles in embryonic and placental development as well as in maternal behavior [38]. However, the identification of imprinted genes in farm species and their roles in development lag behind those in the mouse and human [39]. Because imprints are established during gametogenesis and embryo development [40], environmental factors such as in vitro culture and manipulations of oocytes and embryos affect their expression which can last through the entire life of the animals [41]. Qualitative and quantitative expression profiles of imprinted genes from in vivo pre-implantation embryos are essential gold standards for embryos produced from biotechnologies. In the Chapter three of this dissertation , we characterized all currently known bovine imprinted genes and compared them to their counterparts in humans, mice and pigs in the second section.

1.4. Environment Effects on Embryonic Development

Development depends on both genetics and the environment. Early embryos undergo dynamic changes of gene activities and are particularly sensitive to environmental factors. Although the use of in vitro fertilization (IVF) in cattle and other species such as humans has been valuable to both science and industry, the optimal conditions for fertilization, embryo culture and preservation are still yet to be identified. Negative consequences on IVF derived embryos, fetuses, placentas, and offspring differ significantly in morphology and developmental competence compared with those produced in vivo [42]. Cryopreservation of gametes and embryos is an artificial manipulation that can introduce stress from osmotic, oxidative, cold/heat shock, nutritional and mechanic changes. A relatively recent invention indicated that treating embryos with a well-defined and properly applied sublethal high hydrostatic pressure (HHP) can induce general adaptation and increase tolerance of the embryos to various in vitro procedures [43]. For example, HHP treatment has been shown to improve the survival rates, fertilizing ability and developmental competence of cryopreserved oocytes [44,45], sperm [46,47], embryos [48-50] and embryonic stem cells [51]. Re-expansion and hatching rates of embryos after vitrification-warming were also found to be affected by the duration between termination of HHP treatment and the initiation of vitrification [49,52]. When perfected, such invention can be applied to human embryos to improve IVF success. However, the molecular mechanisms of HHP action have yet to be ascertained. When few clues are available for such effects, a throughput gene expression approach such as the DNA microarray is the method of choice to interrogate all possible molecular pathways involved. In Chapter four of this dissertation, we evaluated the effects of HHP on the transcriptome of bovine IVF blastocysts by the DNA Microarray technology.

1.5. Reprogramming Somatic Cells to Stem Cells of Embryonic Properties

Induced pluripotent stem cells (iPSCs) are embryonic stem cell (ESC) - like cells reprogrammed using ectopic transcription factors *Oct4, Sox2, Klf4*, and *c-Myc* (OKSM) [53,54]. This technology holds great potential for regenerative medicine in producing patient-specific pluripotent cells while bypassing controversial embryo manipulations. iPSCs have been successfully generated from mice and humans by using various combination of reprogramming factors [53-55] in the form of modified RNA [56], proteins [57] and small molecules [58]. A variety of chromatin modifiers have also been reported to facilitate epigenetic changes leading to authentic iPSC reprogramming [59,60] as reprogramming is achieved by overcoming a series of epigenetic barriers [61]. Despite these advances, the reprogramming efficiency of somatic cells remains extremely low [62].

Despite the historic success with generating iPSCs in mice and humans, derivation of pluripotent stem cells from agricultural species, especially cattle, has been extremely challenging for its ill-defined molecular mechanisms and optimized culture conditions in the previous reported incompletely reprogrammed bovine iPSCs [63-67], which negatively impact their potential applications, such as genetic manipulations by using cutting-edge genomic editing technologies of TALENs [68,69] and CRISPRs [70,71].

Acquisition of induced pluripotency requires an intricate interplay among specialized transcriptional circuitries, signaling pathways and chromatin remodeling. In addition to DNA and histone modifications, ATP-dependent enzymes that remodel chromatin are important controllers of chromatin structure and assembly, and are major contributors to regulations of gene expression [72,73]. The SWI/SNF (SWItch/Sucrose NonFermentable or BAF, Brg/Brahma-associated factors) complexes are epigenetic modifiers of chromatin structure and undergo

progressive changes in subunit composition during cellular differentiation. For example, in ESCs esBAF contains *Brg1* and *Baf155*, while their homologs, *Brm* and *Baf170,* are present in BAF of somatic cells. Components of esBAF have been shown to be important in both maintenance of mouse ESCs and iPSC induction. Deficiency in *Brg1, Baf47, Baf155,* or *Baf250* impaired the ability of mESCs to proliferate and to maintain pluripotency [74-79]. *Brg1* and *Baf155*, combined with *Oct4*, *Sox2*, *Klf4* and *c-Myc*, can synergistically increase reprogramming efficiency [80]. It is likely that progressive subunit changes of BAF complex occur while cells transit from differentiated to pluripotent states. The mechanisms of acquisition of components of esBAF and disposition of the somatic BAF during pluripotency establish, however, remain unclear.

Our understanding of epigenetic mechanisms of reprogramming using mouse model will provide insights to bypassing hurdle of bovine iPSCs generation in the future and further for transgenic animal work in agriculture and disease model establishment. In Chapter five of this dissertation, we determined roles of somatic BAF components, *Brm* and *Baf170,* during reprogramming by using shRNA-mediated knockdown studies in the mouse model.

1.6. References

1 Schultz, R. M. The molecular foundations of the maternal to zygotic transition in the preimplantation embryo. *Human reproduction update* **8**, 323-331 (2002).

2 Schultz, R. M. Regulation of zygotic gene activation in the mouse. *BioEssays : news and reviews in molecular, cellular and developmental biology* **15**, 531-538, doi:10.1002/bies.950150806 (1993).

3 Hamatani, T., Carter, M. G., Sharov, A. A. & Ko, M. S. Dynamics of global gene expression changes during mouse preimplantation development. *Developmental cell* **6**, 117-131 (2004).

4 Wang, Q. T. *et al.* A genome-wide study of gene activity reveals developmental signaling pathways in the preimplantation mouse embryo. *Developmental cell* **6**, 133-144 (2004).

5 Braude, P., Bolton, V. & Moore, S. Human gene expression first occurs between the four- and eight-cell stages of preimplantation development. *Nature* **332**, 459-461, doi:10.1038/332459a0 (1988).

6 Telford, N. A., Watson, A. J. & Schultz, G. A. Transition from maternal to embryonic control in early mammalian development: a comparison of several species. *Molecular reproduction and development* **26**, 90-100, doi:10.1002/mrd.1080260113 (1990).

7 Memili, E. & First, N. L. Developmental changes in RNA polymerase II in bovine oocytes, early embryos, and effect of alpha-amanitin on embryo development. *Molecular reproduction and development* **51**, 381-389, doi:10.1002/(SICI)1098-2795(199812)51:4<381::AID-MRD4>3.0.CO;2-G (1998).

8 Misirlioglu, M. *et al.* Dynamics of global transcriptome in bovine matured oocytes and preimplantation embryos. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 18905-18910, doi:10.1073/pnas.0608247103 (2006).

9 Xie, D. *et al.* Rewirable gene regulatory networks in the preimplantation embryonic development of three mammalian species. *Genome research* **20**, 804-815, doi:10.1101/gr.100594.109 (2010).

10 Graf, A. *et al.* Fine mapping of genome activation in bovine embryos by RNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 4139-4144, doi:10.1073/pnas.1321569111 (2014).

11 Kues, W. A. *et al.* Genome-wide expression profiling reveals distinct clusters of transcriptional regulation during bovine preimplantation development in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 19768-19773, doi:10.1073/pnas.0805616105 (2008).

12 Fadiel, A., Anidi, I. & Eichenbaum, K. D. Farm animal genomics and informatics: an update. *Nucleic Acids Res* **33**, 6308-6318, doi:10.1093/nar/gki931 (2005).

13 Lockhart, D. J. *et al.* Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nature biotechnology* **14**, 1675-1680, doi:10.1038/nbt1296-1675 (1996).

14 Duggan, D. J., Bittner, M., Chen, Y., Meltzer, P. & Trent, J. M. Expression profiling using cDNA microarrays. *Nature genetics* **21**, 10-14, doi:10.1038/4434 (1999).

15 Lipshutz, R. J., Fodor, S. P., Gingeras, T. R. & Lockhart, D. J. High density synthetic oligonucleotide arrays. *Nature genetics* **21**, 20-24, doi:10.1038/4447 (1999).

16 Metzker, M. L. Sequencing technologies - the next generation. *Nature reviews. Genetics* **11**, 31-46, doi:10.1038/nrg2626 (2010).
17 Ozsolak, F. & Milos, P. M. RNA sequencing: advances, challenges and opportunities. *Nature reviews. Genetics* **12**, 87-98, doi:10.1038/nrg2934 (2011).
18 Cloonan, N. & Grimmond, S. M. Transcriptome content and dynamics at single-nucleotide resolution. *Genome biology* **9**, 234, doi:10.1186/gb-2008-9-9-234 (2008).
19 Tang, F. *et al.* RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nature protocols* **5**, 516-535, doi:10.1038/nprot.2009.236 (2010).
20 Maher, C. A. *et al.* Transcriptome sequencing to detect gene fusions in cancer. *Nature* **458**, 97-101, doi:10.1038/nature07638 (2009).
21 Linsen, S. E. *et al.* Limitations and possibilities of small RNA digital gene expression profiling. *Nature methods* **6**, 474-476, doi:10.1038/nmeth0709-474 (2009).
22 Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470-476, doi:10.1038/nature07509 (2008).
23 Ameur, A., Wetterbom, A., Feuk, L. & Gyllensten, U. Global and unbiased detection of splice junctions from RNA-seq data. *Genome biology* **11**, R34, doi:10.1186/gb-2010-11-3-r34 (2010).
24 Zeng, F., Baldwin, D. A. & Schultz, R. M. Transcript profiling during preimplantation mouse development. *Developmental biology* **272**, 483-496, doi:10.1016/j.ydbio.2004.05.018 (2004).
25 Smith, S. L. *et al.* Global gene expression profiles reveal significant nuclear reprogramming by the blastocyst stage after cloning. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 17582-17587, doi:10.1073/pnas.0508952102 (2005).
26 Iscove, N. N. *et al.* Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA. *Nature biotechnology* **20**, 940-943, doi:10.1038/nbt729 (2002).
27 Hartmann, C. H. & Klein, C. A. Gene expression profiling of single cells on large-scale oligonucleotide arrays. *Nucleic Acids Res* **34**, e143, doi:10.1093/nar/gkl740 (2006).
28 Kurimoto, K. *et al.* An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. *Nucleic Acids Res* **34**, e42, doi:10.1093/nar/gkl050 (2006).
29 Kurimoto, K., Yabuta, Y., Ohinata, Y. & Saitou, M. Global single-cell cDNA amplification to provide a template for representative high-density oligonucleotide microarray analysis. *Nature protocols* **2**, 739-752, doi:10.1038/nprot.2007.79 (2007).
30 Islam, S. *et al.* Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome research* **21**, 1160-1167, doi:10.1101/gr.110882.110 (2011).
31 Yan, L. *et al.* Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nature structural & molecular biology* **20**, 1131-1139, doi:10.1038/nsmb.2660 (2013).
32 Xue, Z. *et al.* Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* **500**, 593-597, doi:10.1038/nature12364 (2013).

33 Geiss, G. K. *et al.* Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature biotechnology* **26**, 317-325, doi:10.1038/nbt1385 (2008).
34 Plessy, C. *et al.* Linking promoters to functional transcripts in small samples with nanoCAGE and CAGEscan. *Nature methods* **7**, 528-534, doi:10.1038/nmeth.1470 (2010).
35 Wheeler, A. R. *et al.* Microfluidic device for single-cell analysis. *Anal Chem* **75**, 3581-3586 (2003).
36 Marcus, J. S., Anderson, W. F. & Quake, S. R. Microfluidic single-cell mRNA isolation and analysis. *Anal Chem* **78**, 3084-3089, doi:10.1021/ac0519460 (2006).
37 Ferguson-Smith, A. C. Genomic imprinting: the emergence of an epigenetic paradigm. *Nature reviews. Genetics* **12**, 565-575, doi:10.1038/nrg3032 (2011).
38 Lawson, H. A., Cheverud, J. M. & Wolf, J. B. Genomic imprinting and parent-of-origin effects on complex traits. *Nature reviews. Genetics* **14**, 609-617, doi:10.1038/nrg3543 (2013).
39 Tian, X. C. Genomic imprinting in farm animals. *Annual review of animal biosciences* **2**, 23-40, doi:10.1146/annurev-animal-022513-114144 (2014).
40 Reik, W. & Walter, J. Genomic imprinting: parental influence on the genome. *Nature reviews. Genetics* **2**, 21-32, doi:10.1038/35047554 (2001).
41 Mann, M. R. *et al.* Selective loss of imprinting in the placenta following preimplantation development in culture. *Development* **131**, 3727-3735, doi:10.1242/dev.01241 (2004).
42 Farin, P. W., Piedrahita, J. A. & Farin, C. E. Errors in development of fetuses and placentas from in vitro-produced bovine embryos. *Theriogenology* **65**, 178-191, doi:10.1016/j.theriogenology.2005.09.022 (2006).
43 Pribenszky, C., Molnar, M., Cseh, S. & Solti, L. Improving post-thaw survival of cryopreserved mouse blastocysts by hydrostatic pressure challenge. *Animal reproduction science* **87**, 143-150, doi:10.1016/j.anireprosci.2004.09.007 (2005).
44 Du, Y. *et al.* High hydrostatic pressure: a new way to improve in vitro developmental competence of porcine matured oocytes after vitrification. *Reproduction* **135**, 13-17, doi:10.1530/REP-07-0362 (2008).
45 Pribenszky, C., Du, Y., Molnar, M., Harnos, A. & Vajta, G. Increased stress tolerance of matured pig oocytes after high hydrostatic pressure treatment. *Animal reproduction science* **106**, 200-207, doi:10.1016/j.anireprosci.2008.01.016 (2008).
46 Pribenszky, C. *et al.* Stress preconditioning of boar spermatozoa: a new approach to enhance semen quality. *Reproduction in domestic animals = Zuchthygiene* **46 Suppl 2**, 26-30, doi:10.1111/j.1439-0531.2011.01812.x (2011).
47 Huang, S. Y. *et al.* Hydrostatic pressure pre-treatment affects the protein profile of boar sperm before and after freezing-thawing. *Animal reproduction science* **112**, 136-149, doi:10.1016/j.anireprosci.2008.04.016 (2009).
48 Bock, I. *et al.* Stress Tolerance and Transcriptional Response in Mouse Embryos Treated with High Hydrostatic Pressure to Enhance Cryotolerance. *Cryoletters* **31**, 401-412 (2010).
49 Siqueira, E. *et al.* Vitrification of bovine blastocysts pretreated with sublethal hydrostatic pressure stress: evaluation of post-thaw in vitro development and gene expression. *Reprod Fert Develop* **23**, 585-590 (2011).

50 Trigal, B. *et al.* Cell counts and survival to vitrification of bovine in vitro produced blastocysts subjected to sublethal high hydrostatic pressure. *Reproduction in domestic animals = Zuchthygiene* **48**, 200-206, doi:10.1111/j.1439-0531.2012.02131.x (2013).
51 Pribenszky, C. *et al.* Stress for stress tolerance? A fundamentally new approach in mammalian embryology. *Biology of reproduction* **83**, 690-697, doi:10.1095/biolreprod.110.083386 (2010).
52 Pribenszky, C. *et al.* 89 Improved post-warming developmental cometence of open pulled straw-vitrified in vitro produced bovine blastocysts by sublethal hydrostatic pressure pretreatment. *Reproduction, Fertility and Development* **20**, 125-125, doi:http://dx.doi.org/10.1071/RDv20n1Ab89 (2008).
53 Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676, doi:10.1016/j.cell.2006.07.024 (2006).
54 Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313-317, doi:10.1038/nature05934 (2007).
55 Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917-1920, doi:10.1126/science.1151526 (2007).
56 Warren, L. *et al.* Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell stem cell* **7**, 618-630, doi:10.1016/j.stem.2010.08.012 (2010).
57 Kim, D. *et al.* Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell stem cell* **4**, 472-476, doi:10.1016/j.stem.2009.05.005 (2009).
58 Hou, P. *et al.* Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science* **341**, 651-654, doi:10.1126/science.1239278 (2013).
59 Ang, Y. S. *et al.* Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* **145**, 183-197, doi:10.1016/j.cell.2011.03.003 (2011).
60 Mansour, A. A. *et al.* The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature* **488**, 409-413, doi:10.1038/nature11272 (2012).
61 Apostolou, E. & Hochedlinger, K. Chromatin dynamics during cellular reprogramming. *Nature* **502**, 462-471, doi:10.1038/nature12749 (2013).
62 Hanna, J. H., Saha, K. & Jaenisch, R. Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. *Cell* **143**, 508-525, doi:10.1016/j.cell.2010.10.008 (2010).
63 Han, X. *et al.* Generation of induced pluripotent stem cells from bovine embryonic fibroblast cells. *Cell research* **21**, 1509-1512, doi:10.1038/cr.2011.125 (2011).
64 Huang, B. *et al.* A virus-free poly-promoter vector induces pluripotency in quiescent bovine cells under chemically defined conditions of dual kinase inhibition. *PloS one* **6**, e24501, doi:10.1371/journal.pone.0024501 (2011).
65 Sumer, H. *et al.* NANOG is a key factor for induction of pluripotency in bovine adult fibroblasts. *Journal of animal science* **89**, 2708-2716, doi:10.2527/jas.2010-3666 (2011).

66 Cao, H. *et al.* Characterization of bovine induced pluripotent stem cells by lentiviral transduction of reprogramming factor fusion proteins. *Int J Biol Sci* **8**, 498-511, doi:10.7150/ijbs.3723 (2012).
67 Deng, Y. *et al.* Generation of induced pluripotent stem cells from buffalo (Bubalus bubalis) fetal fibroblasts with buffalo defined factors. *Stem Cells Dev* **21**, 2485-2494, doi:10.1089/scd.2012.0018 (2012).
68 Miller, J. C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nature biotechnology* **29**, 143-148, doi:10.1038/nbt.1755 (2011).
69 Zhang, F. *et al.* Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nature biotechnology* **29**, 149-153, doi:10.1038/nbt.1775 (2011).
70 Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821, doi:10.1126/science.1225829 (2012).
71 Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature biotechnology* **31**, 233-239, doi:10.1038/nbt.2508 (2013).
72 Jaenisch, R. & Young, R. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* **132**, 567-582, doi:10.1016/j.cell.2008.01.015 (2008).
73 Ho, L. & Crabtree, G. R. Chromatin remodelling during development. *Nature* **463**, 474-484, doi:10.1038/nature08911 (2010).
74 Gao, X. *et al.* ES cell pluripotency and germ-layer formation require the SWI/SNF chromatin remodeling component BAF250a. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 6656-6661, doi:10.1073/pnas.0801802105 (2008).
75 Yan, Z. *et al.* BAF250B-associated SWI/SNF chromatin-remodeling complex is required to maintain undifferentiated mouse embryonic stem cells. *Stem Cells* **26**, 1155-1165, doi:10.1634/stemcells.2007-0846 (2008).
76 Ho, L. *et al.* An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 5187-5191, doi:10.1073/pnas.0812888106 (2009).
77 Ho, L. *et al.* An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 5181-5186, doi:10.1073/pnas.0812889106 (2009).
78 Schaniel, C. *et al.* Smarcc1/Baf155 couples self-renewal gene repression with changes in chromatin structure in mouse embryonic stem cells. *Stem Cells* **27**, 2979-2991, doi:10.1002/stem.223 (2009).
79 Ho, L. *et al.* esBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating polycomb function. *Nature cell biology* **13**, 903-913, doi:10.1038/ncb2285 (2011).
80 Singhal, N. *et al.* Chromatin-Remodeling Components of the BAF Complex Facilitate Reprogramming. *Cell* **141**, 943-955, doi:10.1016/j.cell.2010.04.037 (2010).

Chapter Two

Transcriptional Profiles of Bovine In Vivo Pre-implantation Development

Published in BMC Genomics, 2014 Sep 4;15:756

(Jiang Z, Sun J, Dong H, Luo O, Zheng X, Obergfell C, Tang Y, Bi J, O'Neill R, Ruan Y, Chen J &Tian XC)

2.1. Abstract

During mammalian pre-implantation embryonic development dramatic and orchestrated changes occur in gene transcription. The identification of the complete changes has not been possible until the development of the Next Generation Sequencing Technology. Here we report comprehensive transcriptome dynamics of single matured bovine oocytes and pre-implantation embryos developed in vivo. Surprisingly, more than half of the estimated 22,000 bovine genes, 11,488 to 12,729 involved in more than 100 pathways, is expressed in oocytes and early embryos. Despite the similarity in the total numbers of genes expressed across stages, the nature of the expressed genes is dramatically different. A total of 2,845 genes were differentially expressed among different stages, of which the largest change was observed between the 4- and 8-cell stages, demonstrating that the bovine embryonic genome is activated at this transition. Additionally, 774 genes were identified as only expressed/highly enriched in particular stages of development, suggesting their stage-specific roles in embryogenesis. Using weighted gene co-expression network analysis, we found 12 stage-specific modules of co-expressed genes that can be used to represent the corresponding stage of development. Furthermore, we identified conserved key members (or hub genes) of the bovine expressed gene networks. Their vast association with other embryonic genes suggests that they may have important regulatory roles in embryo development; yet, the majority of the hub genes are relatively unknown/under-studied in embryos. We also conducted the first comparison of embryonic expression profiles across three mammalian species, human, mouse and bovine, for which RNA-seq data are available. We found that the three species share more maternally deposited genes than embryonic genome activated genes. More importantly, there are more similarities in embryonic transcriptomes between bovine and humans than between humans and mice, demonstrating that bovine embryos

are better models for human embryonic development. This study provides a comprehensive examination of gene activities in bovine embryos and identified little-known potential master regulators of pre-implantation development.

Key words: pre-implantation development; embryonic genome activation; RNA-seq; stage specific module; hub genes, bovine, human, mouse

2.2. Introduction

Mammalian pre-implantation embryonic development is a complex process including fertilization, cleavage divisions, compaction, and blastulation. During this process, massive degradation of oocyte-stored maternal RNA/proteins and gradual activation of the embryonic genome take place [1]. Earlier studies employing RNA polymerase II inhibitor suggested that the timing of EGA is correlated with the speed of embryonic development. For example, a-amanitin halted embryo development at the 2-cell stage in mice [2-4] and between 4- and 8-cell stages in humans [5]. However, the exact timing of EGA in bovine is still debated. Developmental block of cultured bovine embryos occur between the 8- and 16-cell stages [6-8], suggesting EGA at this transition. Similar conclusion was reached by studying expression profiles of bovine in vitro embryos using microarray [9] or the RNA sequencing (RNA-seq) technology [10]. However, a microarray study utilizing pooled in vivo bovine embryos suggested that bovine EGA occurs between the 4- and 8-cell stages [11].

To date, the most comprehensive transcriptome profiling of bovine in vivo embryos was carried out using the Affymetrix GeneChip Bovine Genome Array. Although this microarray contains roughly 23,000 bovine transcripts, these represent only 12,752 genes (personal communications with Affymetrix), approximately half of the mammalian genes. Previous data are also impacted by hybridization variations of microarray and the use of pooled embryos [8,11]. Although more comprehensive data were obtained using RNA-seq, they were restricted to bovine blastocysts only [12,13] or with the use of in vitro embryos [10]. As a result, the complete descriptions of gene activities during bovine in vivo embryonic development are still not available. Moreover, the timing of EGA should not be established using expression data from half of the genome or in vitro embryos. The RNA-seq technology provides unique benefits for

studying gene expression with high resolutions and reproducibility, as well as for detecting novel transcripts and alternative splicing events [14,15]. Here we applied the Solid RNA-seq platform on single in vivo matured oocytes and in vivo developed embryos from the 2-cell to the blastocyst stages and obtained their comprehensive transcriptome dynamics. The identification of highly connected, yet relatively little known or completely unknown hub genes that are potentially master regulators of gene expression opens up unprecedented opportunities for further understanding of early development. Furthermore, we used RNA-seq datasets recently generated in humans and mice and carried out a comprehensive stage-specific comparison across the three mammalian species. We found that the three species shared more maternally deposited genes than embryonic genome activated genes. More significantly, there are more similarities between bovine and human embryonic transcriptomes than those between humans and mice. The data obtained here will function as a comprehensive reference base for embryos generated from reproductive biotechnologies in the bovine as well as in the human for which the use of in vivo embryos is highly limited.

2.3. Materials and Methods

2.3.1. Ethics statement.

Oocytes and embryos were obtained from healthy Holstein cows in the Institute of Animal Science, Xinjiang Academy of Animal Science, Urumqi, Xinjiang, P. R. China. The animal protocol was approved by the Animal Care and Use Committee of Xinjiang Academy of Animal Science (Research license 200815).

2.3.2. Collection of In Vivo Matured Oocytes and Pre-implantation Embryos

Multiparous Holstein cows (n= 10) were synchronized and superovulated as described [16,17]. Artificial insemination using semen from one of three bulls with proven fertility was conducted at 12 and 24 hours post standing heat (Day 0). Donor animals were sacrificed at 30 hours, and 2-4 days after estrus to collect in vivo developed oocytes and embryos at the 2- to 16-cell stages by oviductal flushing. Early morulae, compact morulae and blastocysts were collected by routine non-surgical uterine flushing on Days 5, 6 and 7. All oocytes and embryos were examined and staged under light microscopy. Only morphologically intact embryos meeting the standards of Grade 1 by the International Embryo Transfer Society were included in this study. Embryos were washed twice in D-PBS before frozen and stored individually in RNAlater (Ambion, Grand Island, NY) in liquid nitrogen.

2.3.3. RNA Isolation, Linear Amplification, Library Construction and Sequencing

Following the reproducible procedures of RNA extraction and linear amplification from our previous study [18], we isolated total RNA from individual oocytes/embryos using TRIzol (Invitrogen, Grand Island, NY) and co-precipitated the RNA with linear acrylamide (Ambion). The quality of the total RNA was examined with the Aglient RNA 6000 Pico kit (Aglient

Technologies, Santa Clara, CA) using the Aglient Bioanalyzer 2100. RNA was then amplified twice using the TargetAmp 2-round aminoallyl-aRNA amplification kit 1.0 (Epicentre, Madison, WI) according to the manufacturer's instructions. 500 ng of amplified RNA (aRNA) were used to construct the sequencing library following the manufacturer's instructions by SOLiD™ Total RNA-seq Kit (Life Technologies, Grand Island, NY). After the sequencing library was prepared, we used an Agilent 2100 bioanlyzer to analyze the quality of the libraries. The sequencing libraries were then barcoded, multiplexed, and sequenced on a 5500xl Genetic Analyzer at the Center for Applied Genetics and Technology, University of Connecticut. We obtained 430 million sequencing reads with a read length of 75-bp from 16 single oocytes and embryos. The high correlation coefficients between samples of the same development stage demonstrated the reproducibility of the method (Table S2.1).

2.3.4. Mapping, Assembly and Gene Expression Analysis

Sequencing adapters were trimmed using Cutadapt (https://code.google.com/p/cutadapt/) and sequencing reads of low quality were pre-filtered by FASTX-Toolkit before mapping (http://hannonlab.cshl.edu/fastx_toolkit/), using the options "fastq_quality_trimmer-Q 33-v-t20-l 30-I". The quality of reads after filtering was examined using 'fastQC' (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Filtered reads were mapped to the Btau_4.6.1 assembly using Tophat [19]. Individual mapped reads were fed to Cufflinks [19] to construct transcriptome models. Any novel genes and transcripts that did not fit the supplied gene models (NCBI RefGene) were also assembled. Cuffmerge [19] was used to converge individual transcriptome to produce a master gene model. Genes and transcripts mapped to uncertain chromosomes and contigs were eliminated.

The merged gene model and mapping result BAM files from each RNA-seq library were used to quantify the expression levels of all genes using the Python package, HTSeq [20]. A matrix of Pearson correlation coefficient was created using R, which was in turn used to create the heatmap. Principle component analysis (PCA) was analyzed by using R. Differentially expressed genes between two consecutive developmental stages were identified using default parameters in DESeq [20]. In each comparison, only genes whose sum of expression across all compared samples was greater than the 25th percentile were used. Genes were deemed differentially expressed between subsequent developmental stages if they showed a P-value of less than 0.05 (Negative Binomial Distribution). Expression pattern clusters were generated by the K-means clustering algorithm using R. For gene expression patterns, correlations between pattern indicators and tested genes were calculated. P-values associated with correlations were also calculated and the Bonferroni correction was applied to adjust the P-value for multiple testing. Genes with adjusted P-value of less than 0.05 were considered to have followed the corresponding expression pattern.

2.3.5. Detection of Co-expressed Gene Modules

The R package for weighted gene co-expression network analysis (WGCNA) [21] was used to detect co-expressed gene modules. A weighted gene co-expression network was first constructed, in which genes were nodes and connected with weighted edges. The connection weights between any pair of two nodes i and j were computed by $A(i,j) = (\frac{1}{2} + \frac{1}{2} corr(i,j))^{12}$, where $corr(i,j)$ is the correlation between the expression levels of nodes i and j across all stages. The topological overlap matrix W which measures the topological similarity of every two genes (nodes) was calculated based on the connection weights as follows: $W(i,j) = \frac{l(i,j) + A(i,j)}{\min\{k_i, k_j\} + 1 - A(i,j)}$

where $l(i,j) = \sum_d A(i,d)A(d,j)$ with d indexing the nodes that connect to both i and j, and $k_i = \sum_d A(i,d)$ with d indexing the nodes that connect to the node i.

This topological overlap matrix has been shown to produce more biologically meaningful co-expressed gene clusters [22]. We computed the distance matrix D by $D(i,j) = 1 - W(i,j)$. A dendrogram of clusters was obtained by applying a hierarchical clustering algorithm [23] on the matrix D. The dynamic cutting algorithm reported by Langfelder et al. [24] was used to cut the dendrogram to obtain the clusters of co-expressed genes.

An eigengene was calculated for each cluster as the principal component that explained the largest variance of the data in the cluster. It was a weighted sum of expression profiles of all genes in the cluster where the expression profile of a gene is a vector comprising the values of gene expression at the seven different stages. The eigengene served as the representative of the gene expression profiles in the cluster. Then, clusters whose eigengenes were interrelated with correlation of more than 0.7 were merged. The final clusters of genes were referred to as gene co-expression modules.

2.3.6. Stage-specific Module Identification

To detect modules whose eigengene showed high expression levels at a specific stage but low in others, we used a unit vector to indicate each stage. In other words, the entry of this unit vector for the corresponding stage was one, and zero for the others. We then computed the correlations between each stage-indicator vector and the eigengene of each module, which also yielded a P-value associated with each correlation. Smaller P-values corresponded to more significant correlations. If a module received $P < 0.05$ for the correlation at a particular stage, we labeled the module as specific to that stage.

2.3.7. Module Preservability/Reproducibility

We downloaded and mapped genes from two microarray datasets of the bovine [9,11], and two RNA-seq datasets of the human and mouse [25], as well as our own to the orthologous gene database (http://www.ncbi.nlm.nih.gov/homologene/). After identifying the commonly expressed orthologs, the preservability of a module was measured by the *Z* statistics [26], which characterizes the density and connectivity of genes within a module to those in the validation dataset. The function, module Preservation, in the WGCNA package was used to calculate the *Z* statistics. The categories of preservation were defined as strong if $Z \geq 10$, weak to moderate if $2 \leq Z < 10$, and no evidence of preservation if $Z < 2$, as suggested by an early simulation study [26].

2.3.8. Cross-Species Module Overlapping Analysis

To study if the development of functional modules conserves across species, we compared the gene co-expression modules of the bovine, mouse and human. The same module detection analysis was performed on the human and mouse datasets by Xue et al. [25]. The number of overlapping genes in any two modules each from a different species was counted. Fisher's exact test was conducted to show whether or not the degree of overlapping was simply due to a random chance, which yielded a P-value reflecting the statistical significance of the overlap.

2.3.9. Module Hub Gene Identification and Validation

The membership of a gene in a module was measured by the correlation between that gene and the eigengene of the module. Genes in a module that are highly correlated with the module eigengene are defined as hub genes for the module. We used all genes with correlation to their module eigengene of greater than 0.9 as the hub genes. To explore the connections among hub genes, we examined the top 200 connections of the top 150 hubgenes for each stage

specific module and visualized them in VisANT [27]. To validate the hub genes, we used the raw datasets from two previously published microarray studies in the bovine and one RNA-seq study in the human and mouse. These data were subjected to WGCNA and stage-specific modules and lists of hub genes (kME > 0.9) were generated for each dataset. We then determined the overlap of hub genes from each stage-specific module of the same developmental stage in different datasets.

2.3.10. Gene Ontology Analysis

Functional annotation enrichment analysis for Gene Ontology (GO) was conducted by topGo package in R [28]. Database for Annotation, Visualization and Integrated Discovery [29] Bioinformatics Resource [30] was used for pathway analyses. We summarized all similar sub-GO terms and pathways into an overarching term, and P-values are shown for the representative terms.

2.3.11. Validation of RNA-seq Data

Quantitative real-time PCR (qRT-PCR) was performed to validate differential expression of 10 selected genes using embryos at the 4- and 8-cell stages (n = 3). Among these, five genes (*GATA6, GNB2L1, BAD, H2AFZ* and *NANOG*) were up-regulated and five (*GDP9*, *DNMT1*, *ZP2*, *STAT3* and *OOER*) were down-regulated between these two stages. Amplified RNA from individual embryos was reverse transcribed to cDNA by SuperScript III Reverse Transcriptase (Invitrogen) and amplified with specific primers (Table S2). The qRT-PCR was performed using SYBR Green PCR Master Mix (ABI) and the ABI 7500 Fast instrument. Data were analyzed using the 7500 software version 2.0.2 provided with the instrument. All values were normalized to the internal control, *GAPDH*. The oocytes and embryos from 2-cell to blastocyst stages were pooled and used as the calibrator sample. The relative gene expression values were calculated

using the $2^{-\Delta\Delta Ct}$ method. The mean for each stage was determined and compared for an overall fold change.

2.4. Results

2.4.1. Expression Profiles of Bovine In Vivo Matured Oocytes and Pre-implantation Embryos

Using linearly amplified RNA from single oocytes/embryos, we obtained approximately 430 million sequencing reads from duplicate samples of bovine in vivo pre-implantation embryos at 8 stages of development (Table S2.3). The raw FASTQ files and normalized read counts per gene are available at Gene Expression Omnibus (GEO) (www.ncbi.nlm.nih.gov/geo) under the accession number GSE59186. High Pearson correlation coefficients were found among biological replicates of the same developmental stage (Figure 2.1A, Table S2.1), demonstrating the reproducibility of sample preparation and the sequencing technology. Instead of hundreds of expressed genes reported earlier [8,11], the total numbers of detectable genes ranged from 11,488 to 12,729 from oocytes to blastocysts in our study with the use of RNA-seq (Table 2.1 and Table S2.4). For the first time, it is revealed that the bovine oocytes and early embryos expressed roughly 50% of the total estimated 22,000 genes in the bovine genome.

Pearson correlation coefficients and principal component analyses (PCA) of all detected genes revealed two distinct segmentations of bovine pre-implantation development: the first from the oocyte to the 4-cell stage; and the second from the 8-cell to the blastocyst stage (Figure 2.1A, B). This segmentation demonstrated that EGA in bovine occurs between the 4- and 8-cell stages. This timing concurred with the conclusion by Kues et al. [10,11] yet contrasted with those of all other studies [6-8,10]. Two additional segmentations were also worth-noting: the first from oocyte to 2-/4-cell stages, and the second from 8-cell/morula to the blastocyst stage (Figure 2.1B). These divisions were likely results of dramatic degradation of maternal RNAs and early differentiation, respectively.

2.4.2. Differentially Expressed Genes during Bovine In Vivo Pre-implantation Development

Although the total numbers of genes expressed by embryos of different stages did not vary much, the actual genes expressed during early development are dramatically different. A total of 2,845 unique genes were identified to be differentially expressed between all consecutive stages of development ($P < 0.05$). Similar to Pearson correlations on all detected genes, hierarchal clustering of the differentially expressed genes also partitioned pre-implantation development into two distinct clusters (Figure 2.1C), that from oocytes to 4-cell and from 8-cell to blastocysts, which confirms the timing of EGA. The majority of the differentially expressed genes, 2,031, were found between the 4- to 8-cell stages, providing another confirmation that bovine EGA occurs at this transition. Among these genes, 1,086 and 945 were down- and up-regulated, respectively (Figure 2.1D). The down-regulated genes are involved in reproduction, transcription and cell cycle regulations. Conversely, the up-regulated genes, representing those transcribed from the embryonic genome, are involved in translation, ATP metabolic process as well as RNA processing (Table S2.5). The second largest change in gene expression occurred from compact morula to blastocyst, with 829 genes differentially expressed (Figure 2.1D), suggesting that specific genes were necessary during the blastulation and early differentiation processes. The biological processes significantly represented at this transition are including cell proliferation, transport and early differentiation (Table S2.6). This burst of protein production and cell division may be necessary to prepare the blastocyst for the upcoming coordinated differentiation.

Two minor EGA events were also identified in addition to the major EGA between the 4- and 8-cell stages: one from the oocyte to the 2-cell stage and the other from the 16-cell to the early morula stage, with 324 and 413 genes differentially expressed, respectively (Figure 2.1D).

Between the oocytes and the 2-cell embryos, 166 of the 324 differentially expressed genes were down-regulated. These represent rapid degradation of the maternally stored transcripts. Gene ontology analysis indicated significant over-representation of elements involved in cell cycle and mitosis II (Table S2.7), suggesting that the 2-cell embryos reprogrammed its cell cycle regulation from that of an arrested state to an active mode of cell division through changes of gene expression. The second minor EGA, between the 16-cell embryo and early morula, included 413 differentially expressed genes (Figure 2.1D). These genes may play important roles during the development of tight junctions and other processes of compaction. Interestingly at this transition, we found a high enrichment of genes involved in stem cell maintenance and development, suggesting that genes for pluripotency are active long before the formation of the inner cell mass (Table S2.8).

In spite of the aforementioned differences, there are also wide-spread similarities in the expression profiles between the 2- and 4-cell embryos, the 8- and 16-cell embryos, as well as the early and compact morulae (Figure 2.1D). Specifically, only 97 genes were differentially expressed between the 2- and 4-cell embryos (Table S2.9). Moreover, among the more than 11,000 genes commonly expressed by both the 8- and 16-cell embryos, only 236 genes were differentially expressed (Table S2.10). Likewise, as few as 187 differentially expressed genes were found among the 10,843 commonly expressed genes between the early and compact morulae (Table S2.11).

To confirm the throughput results from RNA-seq, we performed quantitative real-time PCR (qRT-PCR) on 10 genes using in vivo embryos at the 4- and 8-cell stages (n = 3). Among the selected genes, five (*GATA6, GNB2L1, BAD, H2AFZ* and *NANOG*) were up-regulated and

five (*GDP9*, *DNMT1*, *ZP2*, *STAT3* and *OOER*) were down-regulated between these two stages. The qRT-PCR results substantiated those from RNA-seq (Table 2.2).

2.4.3. Cluster Profiles of Differentially Expressed Genes

Although as many as 2,845 differentially expressed genes were identified, the pattern of their dynamic changes can be categorized into as few as 30 distinct clusters, labeled as Clusters 1 to 30 (Figure S2.1; Table S2.12). These clusters can be further assigned to four main groups of different dynamic patterns (Figure 2.2A). The first group, including Clusters 1, 7, 9, 11, 13, 14, 15, 17, 19, 22, 24, 27 and 29, represents genes that increased their expression levels during pre-implantation. All clusters in this group, with the exception of Clusters 11, 17 and 29 showed a dramatic increase at the 8-cell stage, indicating that they are transcribed from the embryonic genome. Genes in these clusters include developmentally important ones such as *GATA6*, *H2AFZ* and *NANOG*. Interestingly, genes in Clusters 11 and 29, including *GATA3* and *DSP,* peaked at the blastocyst stage, suggesting their roles in blastocyst formation and early differentiation. The second dynamic expression pattern, including Clusters 2, 5, 6, 8, 16, 18, 23, 25, 26 and 28, represents genes that underwent an overall trend of decrease, suggesting continued degradation over the pre-implantation period. Of special interest was a sharp decrease during the 4- to 8- cell transition in Clusters 6, 8 and 28, including oocyte-specific genes such as *ZP2* and *WEE2*, demonstrating the lack of their involvement in embryo development. The decrease of genes in this group may also be a pre-requisite for EGA. The third expression pattern, including Clusters 4, 10, 20 and 21, contains genes that first increased and then decreased their expression levels. Among these, Clusters 4 and 10 were up at the 8-cell stages and then declined, suggesting that these genes, such as *BAD, APOPT1* and *GNAT1*, are only involved in the activation of the embryonic genome. The last group, including Clusters 3, 12 and

30, represents genes that maintained relatively constant levels of expression throughout all stages studied. Members from this group such as *ATP1A1*, *ATP5F1* and *RALB* are involved in ion exchange, energy metabolism and signal transduction, suggesting their necessary roles during the entire process of pre-implantation development. It is possible that some of the transcripts in this group were stable maternal RNAs that were never degraded while others were maintained by the early embryos. Nonetheless, genes in this group are good loading control candidates for gene expression quantifications.

In addition to the dynamic changes of gene expression, we also identified genes that are only enriched in one particular stage of development. Specifically, a group of 119 genes were enriched only in the matured oocytes (Figure 2.2B, Table S2.13). These transcripts were degraded after fertilization and remained suppressed during embryo development. Apart from well-known/studied genes such as *H1FOO*, we identified many less known/not annotated genes such as *LOC782175* and *LOC536606*. These genes likely have limited roles in embryo development but are important in maintaining oocytes at the matured stage. Further investigations into their roles in oocytes will enhance our understanding of the mechanism for meiotic arrest. Another group of 234 genes, including *DNMT3A*, *GATA3*, *CD9* and *APOP1*, were only enriched at the blastocyst stage (Figure 2.2B, Table S2.13). Groups of genes were also found to be enriched in a short duration of development such as 310 during oocyte to 4-cell stage and 111 during 8-cell to blastocyst stage (Figure 2.2B, Table S2.13).

2.4.4. Stage-Specific and Cross-Species Gene Expression Comparisons

In addition to analyzing changes of individual genes, we also examined gene-interactions by identifying modules of genes that were co-expressed. Gene co-expression suggests their involvement in a common network of biological processes and functions [31]. Using weighted

gene co-expression network analysis (WGCNA) [21,22] we identified 17 distinct co-expression modules from 13,127 detected genes in our RNA-seq dataset (Figure 2.3A). Twelve of these modules were stage-specific, i.e., these modules included genes that were overexpressed in a particular embryonic stage (Figure 2.3B, Table S2.14) and can be used to represent the corresponding stage of development. Interestingly, analysis of the functions of genes in these modules revealed a sequential progression of stage-specific core gene networks. It migrated from cell cycle in oocytes, to regulation of transcription in 4-cell embryos, to translation in 8-cell embryos, to stem cell development, maintenance and differentiation in morulae, and finally to cell proliferation and protein transport in blastocysts (Figure 2.3C). Such coordinated changes of functional pathways are reflective of the little-known developmental programming.

To explore the conservation and divergence of genes in the 12 stage-specific co-expression modules within and across species, we downloaded the raw datasets from two previously published microarray studies of the bovine [9,11] and one recently published RNA-seq study of the human and mouse [25]. We then identified 8,103, 9,648 and 8,705 commonly expressed orthologs from bovine, humans and mice against our expression dataset. The modulePreservation function of WGCNA was used to calculate the Z-statistics [32], which is a measure of the level and pattern of the connectivity of co-expressed genes. As expected, the 12 stage-specific modules were more significantly preserved within species than between species. Specifically, 5 out of the 12 bovine stage-specific modules were strongly ($Z \geq 10$) preserved in the two published bovine microarray datasets of similar samples, another 5 were weak to moderately preserved ($2 < Z < 10$; Figure 4). To date, there is only one published report [25] of cross-species comparisons using co-expression module data from human and mouse embryos. Here we conducted the first study assessing the cross-species preservation using data from the three

available mammalian species, bovine, human and mouse. Remarkably, most bovine stage-specific modules were at least moderately preserved with those of the human but less with those of the mouse (Figure 2.4), suggesting that the human not only share more similarities with the bovine than with the mouse in genome sequences [33,34] but also in embryonic gene-expression patterns, and thus supporting the notion that bovine embryos are better models for human embryonic development than their mouse counterparts.

To further characterize the conservation and variation of functional modules among species, we conducted module analysis using WGCNA on all detected genes, 14,766 and 13,879, respectively, from the RNA-seq datasets of the human and mouse [19]. We then compared the gene lists within the co-expressed modules of the three species. Intriguingly, there are significant overlaps of genes in the bovine and human stage-specific modules ($P<10^{-4}$; Figure 2.5A, Table S2.15). Of note, many genes in modules specific to the bovine oocyte, 4-cell, 8-cell and morula stages overlapped with those at the corresponding stages in humans. For example, 513 genes ($P<10^{-60}$) overlapped between bovine and human oocytes alone. These genes are involved in protein transport and cell cycle processes. Similarly, highly significant overlap in module genes were observed during bovine and human late pre-implantation development (from 8-cell to blastocyst) although some bovine stage-specific module genes can be found in multiple stages of human development, and vice versa (Figure 2.5A). GO analysis of these overlapped module genes indicate significant over-representation of translation, RNA processing, generation of precursor metabolites and energy. To our surprise, as many as 95 genes ($P<10^{-9}$) in the bovine oocyte-specific module were found in the human 4-cell specific module, giving the apparent suggestion that the progression of embryo development in the bovine may be more rapid than that of humans. This certainly contradicts with the established observations that bovine in vivo

embryo development is 8 days (oocytes to blastocysts) [35] while in humans it is 5 days [36]. One possible explanation to this is the diversity of embryonic programming, e.g., human embryos prepare for an invasive type of implantation while bovine embryos only attach to the uterus [37]. Consistent with this possibility, GO analysis showed that these overlapped genes are related to signal transduction such as Ras and small GTPase signaling. Together, these results suggested that bovine and humans share many core transcriptional programming, while differ in stage-specificity and timing. In the contrary, overlaps of genes between bovine and mouse stage-specific modules were only observed before EGA and after morula formation (Figure 2.5B, Table S2.16). For example, genes in the mouse oocyte-specific module overlapped significantly with those specific to bovine oocyte ($P<10^{-10}$), 2-cell ($P<10^{-2}$) and 4-cell stages ($P<10^{-4}$). Meanwhile, genes specific to the mouse morula module significantly overlapped with those in the bovine 8-cell ($P<10^{-3}$), compact morula ($P<10^{-2}$) and blastocyst modules ($P<10^{-12}$). These results showed that mouse early and late pre-implantation genes are spread over a large period of the bovine development, consistent with the speed of embryo development in these two species.

Collectively, our results show that the three mammalian species share more maternally deposited genes than those expressed after EGA. Based on overlapped genes found in modules prior to EGA, bovine maternal detritus (RNA and protein) occurs later than that in the mouse but slightly earlier than that in the human, despite the longer bovine pre-implantation development.

2.4.5. Identification, Visualization and Validation of Hub Genes

In order to identify genes that are central and highly-connected within the stage-specific modules, we conducted hub gene identification analysis. Hub genes are highly correlated within the stage-specific modules and are conceptual and concrete representatives of the corresponding modules. For each stage-specific module, we assigned all genes with Pearson correlation

coefficients greater than 0.9 as its hub genes (Table 2.3). Furthermore, to explore the connections among hub genes, we examined the top 200 connections of the top 150 hub genes (highly correlated hub genes) for each stage-specific module and visualized them in VisANT (Table 2.4, Figure 2.6). The full lists of these genes can be found in Table S16. Although there are well-studied genes such as *RALB* in oocytes and *DNMT3A* in blastocysts, many of these genes are surprisingly either under-studied in embryonic development or un-annotated, and are thus less known/unknown in this process. For example, *LOC100137763* and *LOC100849216* were highly correlated, un-annotated hub genes in bovine mature oocytes and blastocysts, respectively (Table 2.4, Figure S2.2). The highly correlated hub genes reported here are likely key players for their specific stage(s) of development and may function as "master regulators" of gene expression and stage transition in early development. Further investigation into their identities and functions will greatly enhance our understanding of embryo development and our ability to manipulate embryos through biotechnologies.

Cross-species analysis showed higher degrees of hub gene validation at the oocyte and blastocyst stages than at other stages. For example, 32% (211 genes) and 48% (132 genes) of all hub genes in the oocyte_1 and blastocyst_1 modules, respectively, were validated in at least one dataset (Table 2.3, Table S2.16). Fewer hub genes from the 2-cell to morula stage were successfully validated among species. For example, only 4% and 8% of hub genes were validated at the 16-cell and 8-cell stages, respectively. This low degree of validation reflected the divergence in the stage-specificity and timing of transcriptional programming. Unexpectedly, we observed relatively low number/percentage of validated hub genes against the two previously published bovine datasets [9,11] (Table 2.3), likely because of the low coverage of microarray and/or the relatively low resolution of microarray data.

2.4.6. Pathways in Stage-Specific Modules during Bovine Pre-implantation Development

Pathway analysis revealed essential signaling and metabolic networks in embryonic development. We found more than 100 pathways involved in a sequential order relative to bovine pre-implantation development, most of which were represented in oocytes, major EGA transition (4-cell to 8-cell) and blastocysts (Table S2.17). Components of cell cycle, RNA degradation and progesterone-mediated oocyte maturation pathways were highly enriched before the 4-cell stage, while ribosome, spliceosome and proteasome pathways were highly represented after the 8-cell stage. Interestingly, pathways for oxidative phosphorylation, glycolysis, pyruvate metabolism, pentose phosphate and the citrate cycle (TCA cycle), which are critical not only for cell proliferation, but also for maintenance of pluripotency [38], were uniquely found in blastocysts. Additionally, many well-known pathways including MAPK, insulin, ErbB, Wnt, mTOR and TGF-beta signaling were operative in bovine before the 8-cell stage. It is noteworthy that the most prominent changes in biological networks occurred from oocyte to 4-cell stage and blastocyst stage reflecting major functional transitions.

2.5. Discussion

The development of RNA sequencing technologies permits the study of gene regulation at an unprecedented level. Here, we provide the first comprehensive description of gene activities during in vivo bovine embryonic development. Most such studies had been conducted in the mouse [3,4,9,25,39]. However, mouse data have limited utility in human embryogenesis due to the large differences in gene expression and genome sequences as shown here and in earlier studies. It is therefore important to establish the full expression profile database from an alternative species. To date, all expression profile studies using bovine embryos were either conducted on in vitro embryos and/or using the microarray [8-13,40]. The few studies employing the RNA-seq technology involved blastocyst stage only [12,13,40] except for a recent RNA-seq study using in vitro embryos of multiple stages [10]. None of these reports, however, can be used as the complete "gold standards" for bovine embryo development because in vitro developed embryos have wide-spread gene expression anomalies and the DNA microarray technology limits gene detection to only those printed [8,11] in addition to variations from hybridization. In this study, we applied the RNA-seq technology and revealed the transcriptomes of bovine in vivo pre-implantation development in a very high-throughput and quantitative manner [14]. For the first time, the bovine matured oocytes and early embryos were shown to transcribe more than half of all bovine genes [34].

The timing of EGA in bovine has long been accepted to be between the 8- and 16-cell stages [6-8]. Using in vivo embryos and microarray containing approximately half of the bovine genome, Kues et al. proposed a new timing: between the 4- to 8-cell stage when the largest number of differentially expressed genes were found [11]. This result was confirmed in our study, also employing in vivo embryos but with all bovine expressed genes, and a more powerful

throughput technology, the RNA-seq. However, two prior expression profile studies both using in vitro bovine embryos [8,10], maintained EGA at the 8- to 16-cell transition. It has been shown through the use of RT-PCR that in vitro vs. in vivo embryos have step-wise differences in mRNA expression [41-43]. The difference in EGA timing among the aforementioned studies therefore further demonstrated that in vitro embryos are not suitable for establishing reference base of early development [44,45].

Another important contribution of this study was the discovery of patterns of gene expression and their correlation to milestones of embryo development. Four waves of transcriptional changes, between oocyte and 2-cell, between 4- and 8-cell, between 16-cell to early morula, and between compact morula to blastocyst, were each correlated to degradation of maternal RNA, major EGA, compaction and blastulation. These, together with the identification of transient, stage-exclusive gene expression, provide directions of future research in embryogenesis.

Also for the first time, we identified a number of stage-specific modules in bovine pre-implantation development. They not only represent the corresponding stage of embryogenesis, but reveal an interesting progression of core gene networks from cell cycle (oocyte), to regulation of transcription (4-cell), translation (8-cell), stem cell development, maintenance and differentiation (morula), and finally to cell proliferation and protein transport (blastocyst). The identification of these orchestrated functional changes is among the first step to unveil the little-known embryonic programming, and is important in enhancing our ability to improve assisted embryo biotechnologies such as embryo culture conditions. For example, metabolic pathways unique to the bovine blastocysts, such as glycolysis, pyruvate metabolism and the pentose phosphate pathway, were identified. Their presence is compatible with the “Warburg effect”

commonly found in cancer cells [46]. In this unique pattern of metabolism, glycolytic end products of glucose enter the pentose phosphate pathway instead of the TCA cycle [47]. Such variation from the somatic cells' metabolism of the TCA cycle [47] not only allows for rapid cell proliferation, but also maintains the pluripotency of the bovine blastocyst [38]. Using this feature of the blastocyst, specific medium that encourages the pentose pathway may be developed to increase the proportion of good embryos in the in vitro production system.

Our cross-species analysis demonstrated that human embryos share more similarity to those of the bovine than the mouse in transcriptomes during early embryonic development. The expression profiles established in this report can therefore serve as a reference base for embryos from assisted technology from both cattle and humans. Interestingly, gene expression profiles unveiled unique developmental programming of embryos in the three species analyzed. At the superficial level differences in stage-specific modules appear to suggest that the bovine embryos progress slower than those of the mouse, but more rapid than those of the human. While this is consistent with the in vivo embryo development between the mouse (3.5 days) [39] and the bovine (8 days) but not between the bovine and the human (5 days) [36]. The inconsistency of gene expression at similar stages of development between humans and bovine suggest that the early embryos employ different pathways to prepare themselves for the upcoming different process of implantation. Moreover, our results showed that the three mammalian species share more maternally deposited genes than EGA-activated genes, concurring with the conclusion from a microarray study using the Bayesian clustering method [9] and again revealing species differences in the programming of embryo development.

The cellular and molecular mechanisms governing mammalian pre-implantation development are still poorly understood. Here we identified a number of hub genes that are

critical connectors to other expressed components within each embryonic stage of development in the bovine. Their importance is "validated" by those that have been studied previously. For example, *RALB* is a highly correlated hub gene in oocytes and has key roles in both bovine [48] and Xenopus embryo development [49]. It is also implicated in tumorigenesis and cell proliferation in mice [50]. Similarly, *DNMT3A* is a hub gene in blastocysts. Studies in mice have demonstrated that *DNMT3A* is essential for de novo methylation and embryo development [51,52]. *DNMT3A* is also likely essential in the bovine blastulation process. The observations that *DNMT3A* is significantly reduced in cloned bovine embryos [53] and that lower pregnancy/calving rates and abnormal development are commonplace in cloned fetuses are "functional validation" of the hub gene status for this important regulator of epigenetic modifications [54].

Most identified hub genes, however, haven't been studied or annotated, albeit their potential important roles in embryo development. For example, *LOC100137763* and *LOC100849216* are highly correlated, yet un-annotated hub genes identified in oocytes and blastocysts, respectively. The hub genes identified here represent the unprecedented opportunities and insights offered by the RNA-seq technology and bioinformatics. Collectively, our inventories of all hub genes provide a valuable resource for further studies of the molecular mechanisms of pre-implantation development.

Although we had to conduct linear RNA amplification in order to yield sufficient materials from single oocytes/embryos, the highly reproducible protocol employed here has been previously validated [18] and does not alter the abundance of RNA from the original samples [18]. Readers are cautioned, however, that the in vivo oocytes/embryos used here were generated after superovulation treatment. Although superovulation can affect gene expression of oocytes/embryos [55-58], it is frequently used in both research and production [59] because naturally

ovulated/developed oocytes/embryos from single-ovulatory, large animals such as cattle are not very feasible. Nonetheless, the ultimate "gold standards" for gene expression during bovine pre-implantation development can only be established using naturally ovulated/developed oocytes/embryos.

Overall, this study provides comprehensive examinations of gene activities in in vivo bovine oocytes and pre-implantation embryos. Cross-species analysis revealed that bovine pre-implantation transcriptional profiles share more similarity to those of the human than the mice. The data presented here can be used to assess the impact of various assisted reproductive techniques in both bovine and human reproduction.

Table 2.1. The numbers of genes detected in bovine in vivo matured oocytes and each stage of in vivo embryonic development.

Stage	No. of Genes (FPKM>0.1)
Oocyte	11488
2-cell	12678
4-cell	12718
8-cell	12729
16-cell	11973
Early morula	11670
Compact morula	11910
Blastocyst	11924

FPKM: fragments per kilobase of exon per million fragments mapped.

Table 2.2. Quantitative real-time RT-PCR (qRT-PCR) results of 10 selected genes between 4- and 8-cell stage embryos.

Comparison	Gene symbol	Log (fold change) RNA-Seq	Expression	Log (fold change)* qRT-PCR
4- vs. 8-cell embryos	*GATA6*	8.4	Up	10.6
	GNB2L1	8.2	Up	9.2
	BAD	6.4	Up	8.4
	H2AFZ	8.2	Up	10.3
	NANOG	11.5	Up	7.8
	GDP9	-3.8	Down	-2.1
	DNMT1	-3.2	Down	-2.6
	ZP2	-3.5	Down	-4.4
	STAT3	-2.8	Down	-2.3
	OOER	-2.7	Down	-2.0

***Fold change is expressed as the ratios of the values of the 4-cell embryos (n=3) divided by those of the 8-cell embryos (n=3). Real time RT-PCR results substantiated the differential gene expression patterns from RNA-seq.**

Table 2.3. The numbers and validation results of hub genes in bovine in vivo oocytes and embryos.

Stage-Specific Modules	Total No. of Genes /Module	Total No. of Hub Genes/Module	No. (%) Hub Genes (Validated in at Least One Dataset)*	No. (%) Hub Genes (Validated in Kues et al. 2008)
Oocyte_1	2347	650	211 (32%)	117 (18%)
Oocyte_2	815	125	24 (19%)	6 (5%)
2-cell	444	112	12 (11%)	11 (10%)
4-cell	1868	299	54 (18%)	19 (6%)
8-cell_1	229	52	7 (13%)	5(10%)
8-cell_2	640	172	14 (8%)	13 (8%)
16-cell	247	27	1 (4%)	0 (0)
Early morula	120	15	2 (13%)	2 (13%)
Compact morula	311	41	7 (17%)	7 (17%)
Blastocyst_1	1049	274	132 (48%)	116 (42%)
Blastocyst_2	1366	366	59 (16%)	44 (12%)
Blastocyst_3	1010	118	65 (55%)	44(37%)

More than one modules were found for oocytes, 8-cell embryos and blastocysts.
*: two bovine microarray [9,11], one human and one mouse RNA-seq datasets [25] were used.

Table 2.4. Highly correlated hub genes in bovine stage-specific modules.

Stage-specific module	Hub Genes
Oocyte_1	*SRPX, NAA30*
Oocyte_2	*LOC100137763, PAX3, RALB, SMC1B, UNC13C, VANGL1*
2-cell	*CAPRIN2, LACC1, LOC616167, NLRP9, ZGLP1, POL*
4-cell	*CNTNAP2, TPM3*
8-cell	*LOC519952, LOC789391, LYSMD3, TBXAS1, THAP8*
16-cell	*ARL10, FAM84B, LOC790411, CCDC39*
Early morula	*LGALS9, STAC, LOC100140626*
Compact morula	*APOBR, GALNTL1, LRP8, PCDH10, RGS20, HOXA11, LOC781048*
Blastocyst_1	*DNMT3A, ATP6V0A4, FAM115C, LGALS1, SLC9A3R1*
Blastocyst_2	*BCAM, BPIFA1, LOC100849216, PLXNA3, SHROOM2, SLC16A7*
Blastocyst_3	*EEF2, RPL10A, RPL38*

Multiple modules exist for oocytes and blastocysts.

Figure 2.1. Correlation and hierarchical analyses of transcriptomes of bovine in vivo developed oocytes and embryos.

(A). Heatmap of duplicate samples of the same stages of bovine embryonic development. The color spectrum, ranging from red through white to blue, represents Pearson correlation coefficients ranging from 1 to 0.53, indicating high to low correlations. All duplicate samples are highly correlated in Pearson coefficients demonstrating the reproducibility of the procedures.

(B). Principal component analysis (PCA) of the transcriptomes for seven stages of in vivo developed bovine embryos and oocytes. Embryos from the same stage are shown by symbols of the same shape. The arrows indicate the developmental direction of the embryos. PC1, PC2 and PC3 represent the top three dimensions of the differentially expressed genes among the preimplantation embryos.

(C). Hierarchical clustering of differentially expressed genes in in vivo developed bovine oocytes and embryos. Two major clusters are shown, one consisted of the matured oocytes and embryos at the 2- and 4-cell stages. The second cluster is consisted of embryos at the 8-, 16-cell, early morula, compact morula and blastocyst stages. The clear separation of embryos into two groups demonstrated the timing of EGA in cattle at the 4- to 8-cell transition. The color spectrum, ranging from red through yellow to blue, indicates normalized levels of gene expression.

(D). The numbers of differentially expressed genes in consecutive stages of bovine in vivo pre-implantation development (P<0.05).

Fig. 2.1

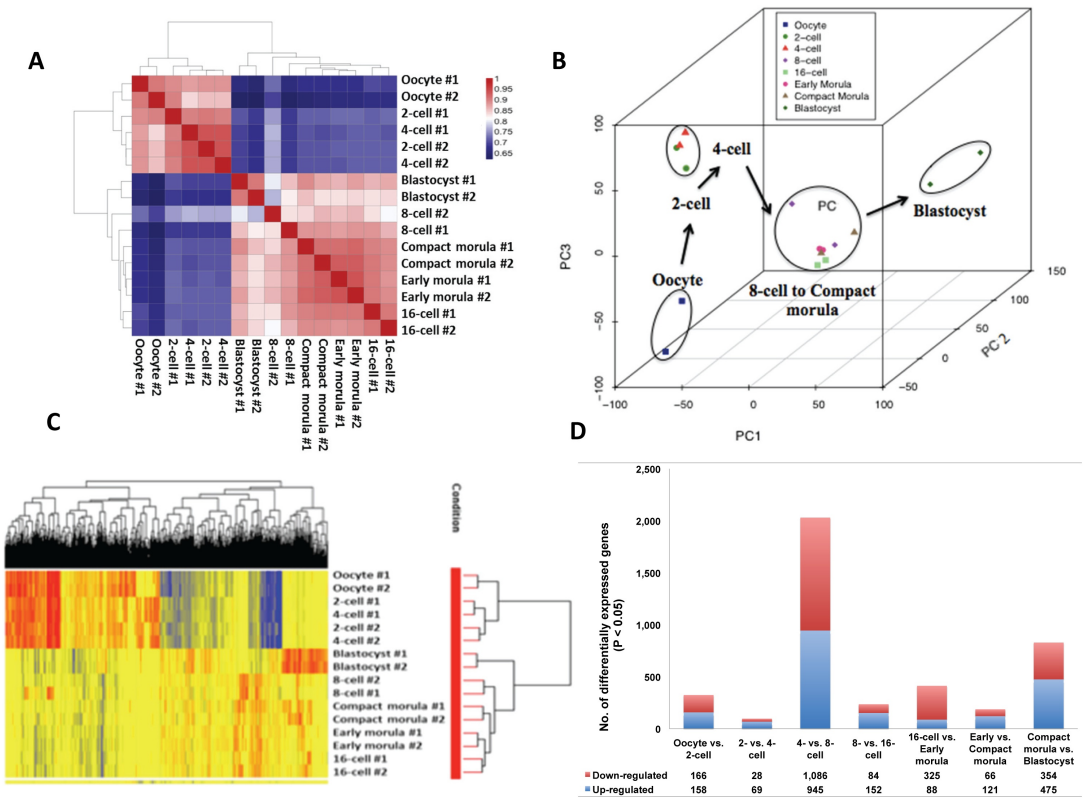


Figure 2.2. Distinct patterns and dynamics of gene expression during bovine in vivo pre-implantation development.

(A). Representative clusters of expression dynamics during early development. Genes were clustered to be increased (a), decreased (b), increased first and then decreased (c), and maintained relatively constant levels of expression (d).

(B). Identification of stage-specific/enriched genes by cluster analysis. Groups of genes were found to be only expressed in oocytes and blastocysts, enriched in oocytes to 4-cell embryos, and 8-cell to blastocysts. The color spectrum, ranging from red to white, represents high to low levels of gene expression.

Fig. 2.2

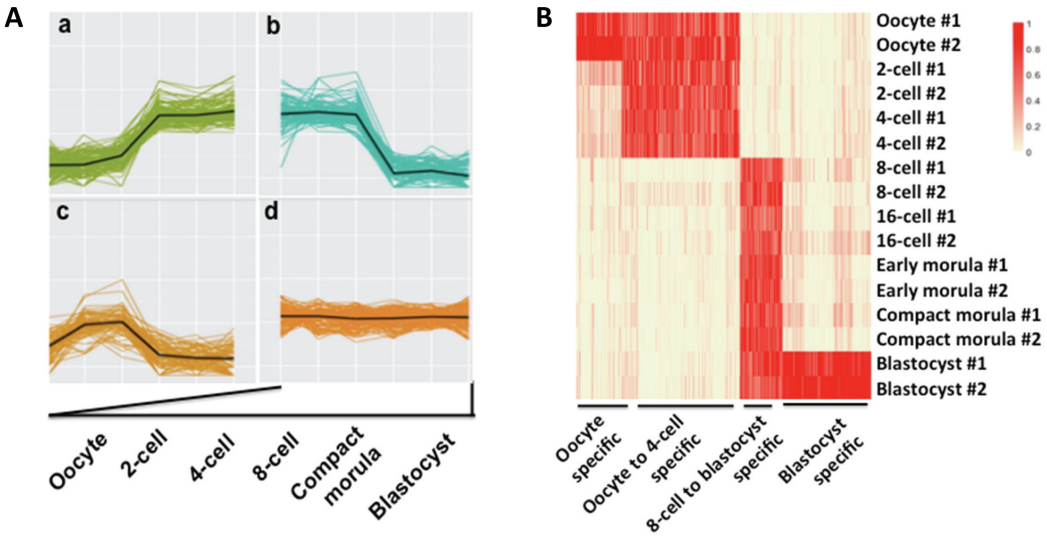


Figure 2.3. Co-expression network analysis of bovine pre-implantation development.
(A). Hierarchical cluster tree showing modules of co-expressed genes identified by WGCNA. A total of 17 co-expressed modules were found during bovine pre-implantation development and were represented by branches and labeled by different colors to the right of the tree. The height (X-axis) indicates levels of correlation.
(B). Heatmap of correlations (and corresponding P-values) between co-expressed modules and stage of development. The color scheme, from blue through white to red, indicates the levels of correlation, from low to high. The stage-specific modules identified are highly correlated (i.e. over-expressed) with distinct developmental stages (columns).
(C). Functional terms of stage-specific modules of co-expressed genes during bovine pre-implantation development. A systematic and sequential changes in functions of co-expressed genes were observed as the embryos progress through early development.

Fig. 2.3

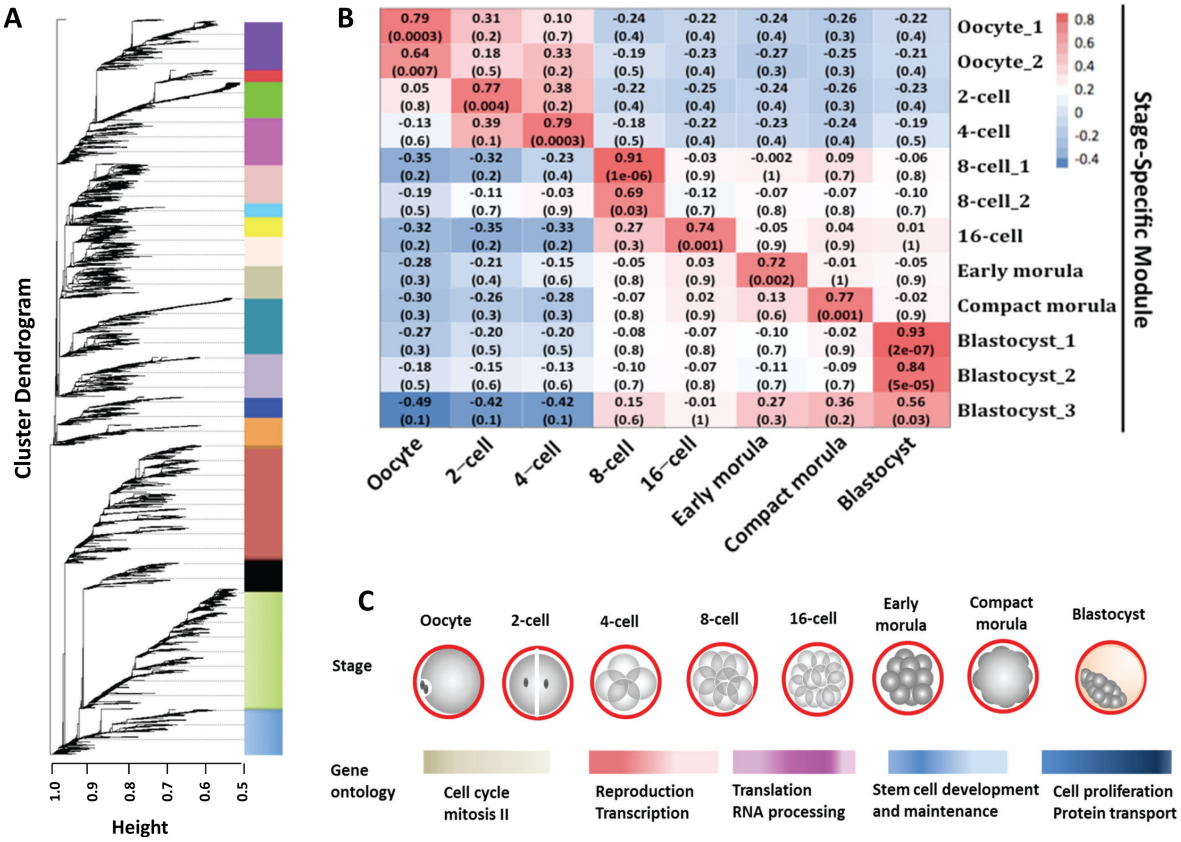


Figure 2.4. Heatmap of module preservation of stage-specific gene co-expression among bovine, human and mouse oocytes and embryos. Commonly expressed orthologs from the present study and those published previously as indicated on the X-axis, including two from bovine microarray studies [9,11] and an RNA-seq study in humans and mice [25] were identified and included. The labels on the Y-axis are stage-specific modules of co-expressed genes. The color scheme, from white to red, indicates low to high levels of preservation.

Fig. 2.4

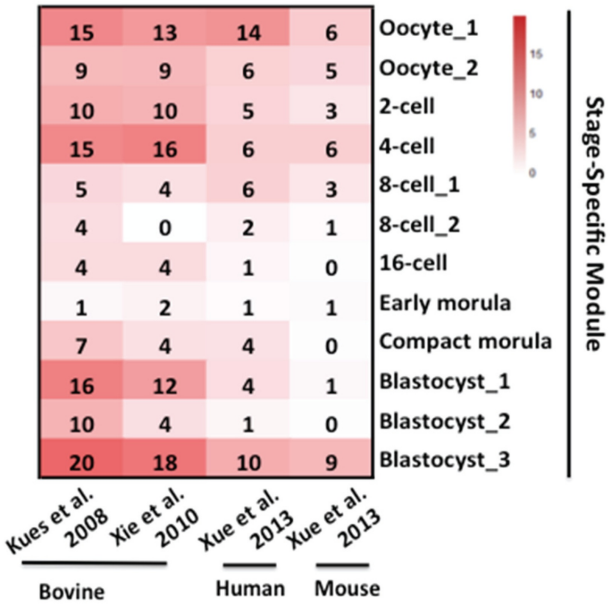


Figure 2.5. Divergence of stage-specific gene co-expression among bovine, human and mouse oocytes and embryos.

(A). Heatmap of gene overlap between independently constructed bovine and human modules. The X- and Y-axes show human (n=9) and bovine stage-specific modules (n=12), respectively. Each cell contains the number of intersecting genes and the corresponding P-value (–log10) of the intersection. Significant gene overlaps were found in nearly all stages of development between the human and bovine.

(B). Heatmap of gene overlap between independently constructed bovine and mouse modules. The X- and Y-axes show mouse (n=9) and bovine stage-specific modules (n=12), respectively. Each cell contains the number of intersecting genes and the corresponding P-value (–log10) of the intersection. Significant gene overlaps were only observed in oocytes and morula/blastocysts between the bovine and mouse.

Fig. 2.5

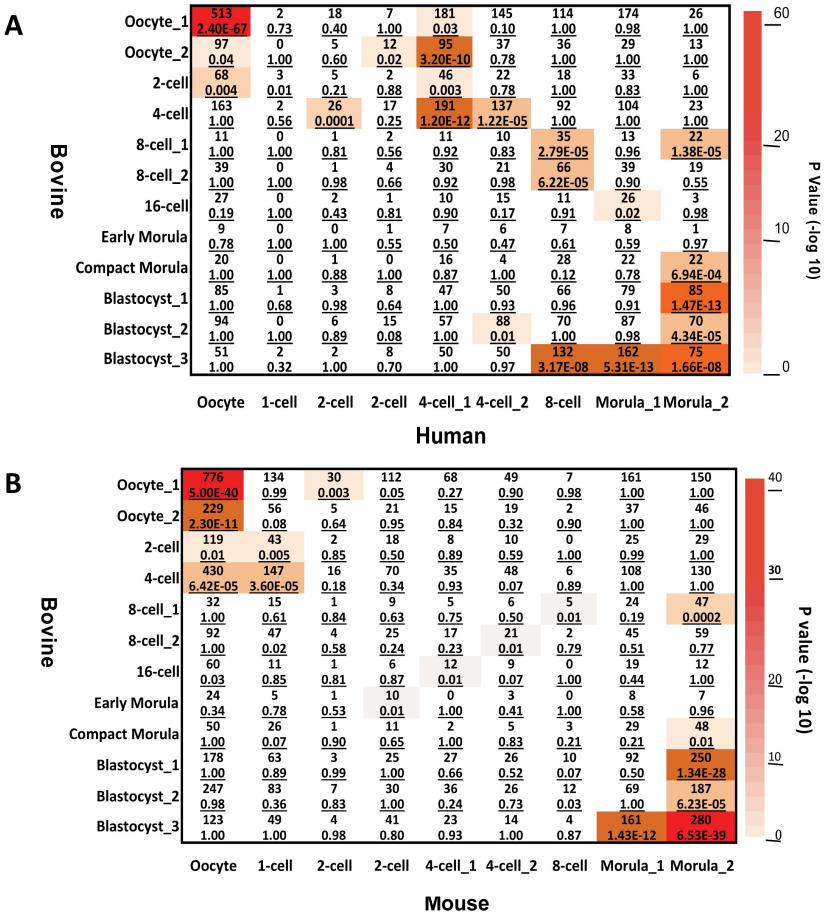
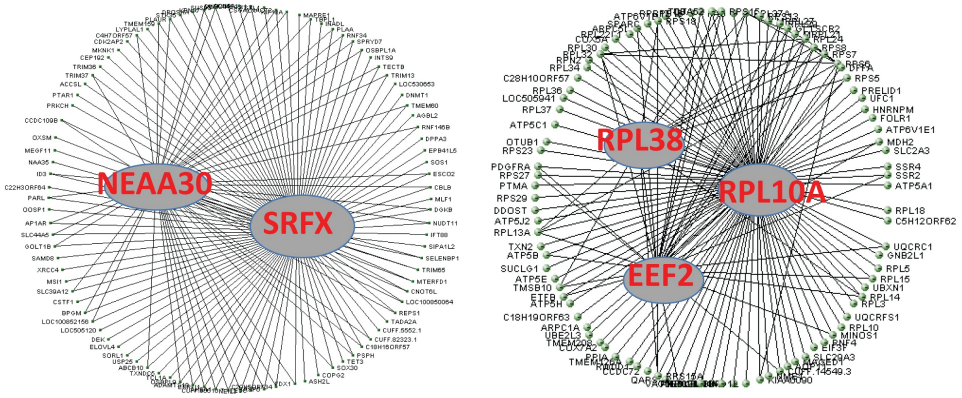


Figure 2.6. Representative, highly correlated hub genes in bovine oocytes (*NEAA30, SRFX*) and blastocysts (*RPL38, EEF2, RPL10A*).

Fig. 2.6



2.6. Supplementary Information

Table S2.1: Pearson correlation coefficients of duplicate bovine oocytes and embryos of the same stage.

	Oocyte #1	Oocyte #2	2-cell #1	2-cell #2	4-cell #1	4-cell #2	8-cell #1	8-cell #2	16-cell #1	16-cell #2	Early morula #1	Early morula #2	Compact morula #1	Compact morula #2	Blastocyst #1	Blastocyst #2
Oocyte #1	1	0.91	0.89	0.88	0.88	0.88	0.66	0.72	0.70	0.69	0.68	0.68	0.68	0.68	0.66	0.65
Oocyte #2	0.91	1	0.90	0.85	0.84	0.85	0.63	0.69	0.67	0.67	0.66	0.66	0.65	0.66	0.64	0.62
2-cell #1	0.89	0.90	1	0.90	0.90	0.90	0.69	0.75	0.72	0.72	0.72	0.72	0.72	0.71	0.71	0.69
2-cell #2	0.88	0.85	0.90	1	0.93	0.93	0.68	0.76	0.72	0.71	0.71	0.72	0.71	0.71	0.70	0.68
4-cell #1	0.88	0.84	0.90	0.93	1	0.93	0.68	0.77	0.71	0.71	0.71	0.71	0.71	0.70	0.69	0.68
4-cell #2	0.88	0.85	0.90	0.93	0.93	1	0.69	0.76	0.72	0.71	0.71	0.71	0.70	0.71	0.69	0.68
8-cell #1	0.66	0.63	0.69	0.68	0.68	0.69	1	0.83	0.89	0.86	0.88	0.88	0.89	0.88	0.85	0.82
8-cell #2	0.72	0.69	0.75	0.76	0.77	0.76	0.83	1	0.83	0.81	0.85	0.86	0.84	0.86	0.80	0.77
16-cell #1	0.70	0.67	0.72	0.72	0.71	0.72	0.89	0.83	1	0.91	0.90	0.91	0.90	0.90	0.86	0.83
16-cell #2	0.69	0.67	0.72	0.71	0.71	0.71	0.86	0.81	0.91	1	0.89	0.89	0.89	0.89	0.87	0.84
Early morula #1	0.68	0.66	0.72	0.71	0.71	0.71	0.88	0.85	0.90	0.89	1	0.94	0.91	0.93	0.86	0.83
Early morula #2	0.68	0.66	0.72	0.72	0.71	0.71	0.88	0.86	0.91	0.89	0.94	1	0.92	0.93	0.88	0.84
Compact morula #1	0.68	0.65	0.72	0.71	0.71	0.70	0.89	0.84	0.90	0.89	0.91	0.92	1	0.91	0.89	0.85
Compact morula #2	0.68	0.66	0.71	0.71	0.70	0.71	0.88	0.86	0.90	0.89	0.93	0.93	0.91	1	0.87	0.83
Blastocyst #1	0.66	0.64	0.71	0.70	0.69	0.69	0.85	0.80	0.86	0.87	0.86	0.88	0.89	0.87	1	0.90
Blastocyst #2	0.65	0.62	0.69	0.68	0.68	0.68	0.82	0.77	0.83	0.84	0.83	0.84	0.85	0.83	0.90	1

Table S2.2: Primers for real time qRT-PCR.

Genes	Primer sequences (5' - 3')	Annealing temperature (°C)	Fragment size (bp)	Accession number
GATA6	GAAGACCCAGGGAGAAGAGG			
	GCAGTTCTGGTTTCAGCACA	60	152	XM_002697727.1
GNB2L1	ATCTCACAACGGGCACTACC			
	CGGTTGTCAGAGGAGAAAGC	60	82	XM_005209237.1
BAD	GAGGATGAGCGACGAGTTTC			
	TCAACCAGGACTGGAGGAAG	60	119	XM_005227186.1
H2AFZ	GGTAAGGCTGGGAAGGACTC			
	TTCAAGTACCTCTGCGGTGA	60	195	NM_174809.2
NANOG	TAGGGAATCTTCACCCATGC			
	GGGACCGTCTCTTCCTTCTC	60	204	NM_001025344.1
GDF9	TCTGGTTCCAGCTTCATTCA			
	CTGGAGAGCCATACCGATGT	60	200	NM_174681.2
DNMT1	AGTGGGGGACTGTGTTTCTG			
	CTGCATGTCCTCACACTCGT	60	199	NM_182651.2
ZP2	ATGTATCTCCTGGGCAGACG			
	GAGGCCATTTGCTATTTCCA	60	200	NM_173973.2
STAT3	GGCCATCTTGAGCACTAAGC	60	206	NM_001012671.2

	TGGTGGCATCCATGATCTTA			
OOER	AGGAATCCGCTGGTGTTCTT			
	TGGCTCAGAAGCACACTCTT	60	200	NM_001077869.2
GAPDH 4	AGATGGTGAAGGTCGGAGTG			
	GAAGGTCAATGAAGGGGTCA	60	117	NM_001034034.2

Table S2.3: Summary of sequence read alignments to the reference genome.

Stage	Total reads	Total mapped reads	Mapped reads (%)	Unmapped reads
Oocyte #1	33,710,645	21,088,863	62.56	12,621,782
Oocyte #2	27,340,216	19,042,322	69.65	8,297,894
2-cell #1	24,604,359	12,271,310	49.87	12,333,049
2-cell #2	48,722,568	27,125,963	55.67	21,596,605
4-cell #1	25,442,215	13,834,869	54.38	11,607,346
4-cell #2	35,703,218	19,172,270	53.70	16,530,948
8-cell #1	18,788,513	12,270,210	65.31	6,518,303
8-cell #2	21,290,475	14,667,771	68.89	6,622,704
16-cell #1	29,993,256	18,528,062	61.77	11,465,194
16-cell #2	13,799,407	7,732,281	56.03	6,067,126
Early morula #1	23,813,659	15,778,377	66.26	8,035,282
Early morula #2	32,821,097	21,557,091	65.68	11,264,006
Compact Morula #1	20,822,446	12,464,723	59.86	8,357,723
Compact Morula #2	26,674,386	17,364,014	65.10	9,310,372
Blastocyst #1	23,772,162	15,102,607	63.53	8,669,555
Blastocyst #2	21,989,584	13,098,519	59.57	8,891,065
Total	**429,288,206**	**261,099,252**	**60.82**	**168,188,954**

Table S2.4: Normalized read counts (expression) of genes in bovine oocytes and embryos.
http://www.ncbi.nlm.nih.gov/pubmed/25185836

Table S2.5: Differentially expressed genes between the 4- and 8-cell embryos. Spreadsheet 1: all differentially expressed genes; Spreadsheet 2: genes down-regulated; Spreadsheet 3: genes up-regulated; Spreadsheet 4: GO analysis output of down-regulated genes; Spreadsheet 5: GO analysis output of up-regulated genes.
http://www.ncbi.nlm.nih.gov/pubmed/25185836

Table S2.6: Differentially expressed genes between compact morulae and blastocysts. Spreadsheet 1: all differentially expressed genes; Spreadsheet 2: genes down-regulated; Spreadsheet 3: genes up-regulated; Spreadsheet 4: GO analysis output of down-regulated genes; Spreadsheet 5: GO analysis output of up-regulated genes.
http://www.ncbi.nlm.nih.gov/pubmed/25185836

Table S2.7: Differentially expressed genes between oocytes and the 2-cell embryos. Spreadsheet 1: all differentially expressed genes; Spreadsheet 2: genes down-regulated; Spreadsheet 3: genes up-regulated; Spreadsheet 4: GO analysis output of down-regulated genes; Spreadsheet 5: GO analysis output of up-regulated genes.
http://www.ncbi.nlm.nih.gov/pubmed/25185836

Table S2.8: Differentially expressed genes between the 16-cell embryos and early morulae. Spreadsheet 1: all differentially expressed genes; Spreadsheet 2: genes down-regulated; Spreadsheet 3: genes up-regulated; Spreadsheet 4: GO analysis output of down-regulated genes; Spreadsheet 5: GO analysis output of up-regulated genes.
http://www.ncbi.nlm.nih.gov/pubmed/25185836

Table S2.9: Differentially expressed genes between the 2- and 4-cell embryos. Spreadsheet 1: all differentially expressed genes; Spreadsheet 2: genes down-regulated; Spreadsheet 3: genes up-regulated; Spreadsheet 4: GO analysis output of down-regulated genes; Spreadsheet 5: GO analysis output of up-regulated genes.
http://www.ncbi.nlm.nih.gov/pubmed/25185836

Table S2.10: Differentially expressed genes between the 8- and 16-cell embryos. Spreadsheet 1: all differentially expressed genes; Spreadsheet 2: genes down-regulated; Spreadsheet 3: genes up-regulated; Spreadsheet 4: GO analysis output of down-regulated genes; Spreadsheet 5: GO analysis output of up-regulated genes.
http://www.ncbi.nlm.nih.gov/pubmed/25185836

Table S2.11: Differentially expressed genes between the early and compact morulae. Spreadsheet 1: all differentially expressed genes; Spreadsheet 2: genes down-regulated; Spreadsheet 3: genes up-regulated; Spreadsheet 4: GO analysis output of down-regulated genes; Spreadsheet 5: GO analysis output of up-regulated genes.
http://www.ncbi.nlm.nih.gov/pubmed/25185836

Table S2.12: Distinct Clusters of gene expression patterns in bovine oocytes and embryos.

http://www.ncbi.nlm.nih.gov/pubmed/25185836

Table S2.13: Stage-specific/enriched genes in bovine oocytes and pre-implantation embryos.
http://www.ncbi.nlm.nih.gov/pubmed/25185836

Table S2.14: Co-expressed genes in stage-specific modules of bovine oocytes and embryos.
http://www.ncbi.nlm.nih.gov/pubmed/25185836

Table S2.15: Gene expression overlap between species. Spreadsheet 1: overlapped genes between bovine and human embryos; Spreadsheet 2: overlapped genes between bovine and mouse embryos.
http://www.ncbi.nlm.nih.gov/pubmed/25185836

Table S2.16: All hub genes identified in bovine oocytes and pre-implantation embryos (genes in blue are validated in at least one dataset of the human, mouse and bovine).
http://www.ncbi.nlm.nih.gov/pubmed/25185836

Table S2.17: Functional pathways in stage-specific modules in bovine oocytes and embryos.

Stage	Pathways	Count	P Value
Oocyte	bta04520:Adherens junction	18	1.40E-04
Oocyte	hsa04360:Axon guidance	13	2.24E-04
Oocyte	bta04120:Ubiquitin mediated proteolysis	27	3.51E-04
Oocyte	bta04530:Tight junction	24	1.15E-03
Oocyte	bta04110:Cell cycle	24	1.29E-03
Oocyte	hsa05200:Pathways in cancer	20	1.74E-03
Oocyte	bta04012:ErbB signaling pathway	17	2.62E-03
Oocyte	bta04114:Oocyte meiosis	21	3.30E-03
Oocyte	bta03410:Base excision repair	9	6.07E-03
Oocyte	hsa04010:MAPK signaling pathway	16	7.13E-03
Oocyte	bta04540:Gap junction	16	7.49E-03
Oocyte	bta03018:RNA degradation	5	7.95E-03
Oocyte	bta04914:Progesterone-mediated oocyte maturation	16	9.45E-03
Oocyte	bta04320:Dorso-ventral axis formation	7	1.20E-02
Oocyte	bta05220:Chronic myeloid leukemia	14	2.29E-02
Oocyte	bta00130:Ubiquinone and other terpenoid-quinone biosynthesis	4	2.32E-02
Oocyte	hsa04510:Focal adhesion	12	2.41E-02
Oocyte	hsa05222:Small cell lung cancer	7	2.92E-02
Oocyte	bta04310:Wnt signaling pathway	22	2.98E-02
Oocyte	hsa04660:T cell receptor signaling pathway	8	3.06E-02
Oocyte	bta04722:Neurotrophin signaling pathway	19	3.23E-02
Oocyte	bta05216:Thyroid cancer	7	3.36E-02
Oocyte	bta05221:Acute myeloid leukemia	11	3.49E-02
Oocyte	hsa04810:Regulation of actin cytoskeleton	12	3.70E-02
Oocyte	bta05213:Endometrial cancer	10	3.71E-02
Oocyte	bta03420:Nucleotide excision repair	9	3.89E-02
Oocyte	bta05212:Pancreatic cancer	12	5.35E-02
Oocyte	bta04144:Endocytosis	25	5.81E-02
Oocyte	bta03020:RNA polymerase	3	5.96E-02
Oocyte	hsa04620:Toll-like receptor signaling pathway	7	6.23E-02
Oocyte	hsa04910:Insulin signaling pathway	8	8.24E-02
Oocyte	bta05211:Renal cell carcinoma	11	9.10E-02
Oocyte	bta00240:Pyrimidine metabolism	14	9.50E-02
2-cell	bta05014:Amyotrophic lateral sclerosis (ALS)	4	8.35E-02
2-cell	bta04810:Regulation of actin cytoskeleton	8	8.91E-02
2-cell	bta04621:NOD-like receptor signaling pathway	4	9.13E-02
2-cell	bta00534:Heparan sulfate biosynthesis	3	9.63E-02
4-cell	bta04320:Dorso-ventral axis formation	9	4.84E-05
4-cell	bta04120:Ubiquitin mediated proteolysis	22	3.51E-04
4-cell	bta04914:Progesterone-mediated oocyte maturation	16	4.07E-04

4-cell	bta05215:Prostate cancer	16	6.09E-04
4-cell	bta05200:Pathways in cancer	37	9.98E-04
4-cell	bta04916:Melanogenesis	16	1.14E-03
4-cell	bta05210:Colorectal cancer	15	1.85E-03
4-cell	bta04910:Insulin signaling pathway	19	2.47E-03
4-cell	bta00230:Purine metabolism	14	3.76E-03
4-cell	bta04510:Focal adhesion	24	3.79E-03
4-cell	bta04144:Endocytosis	23	6.41E-03
4-cell	bta04150:mTOR signaling pathway	10	8.04E-03
4-cell	bta04114:Oocyte meiosis	16	8.42E-03
4-cell	bta04062:Chemokine signaling pathway	14	1.17E-02
4-cell	bta04664:Fc epsilon RI signaling pathway	12	1.19E-02
4-cell	bta04360:Axon guidance	16	1.26E-02
4-cell	bta05211:Renal cell carcinoma	11	1.40E-02
4-cell	bta04660:T cell receptor signaling pathway	15	1.40E-02
4-cell	bta04722:Neurotrophin signaling pathway	16	1.69E-02
4-cell	bta05213:Endometrial cancer	9	1.70E-02
4-cell	bta05212:Pancreatic cancer	11	1.71E-02
4-cell	bta04520:Adherens junction	11	1.71E-02
4-cell	bta04012:ErbB signaling pathway	12	1.90E-02
4-cell	bta04662:B cell receptor signaling pathway	11	2.08E-02
4-cell	bta05214:Glioma	10	2.25E-02
4-cell	bta04810:Regulation of actin cytoskeleton	22	2.37E-02
4-cell	bta04110:Cell cycle	16	2.39E-02
4-cell	bta04930:Type II diabetes mellitus	8	2.50E-02
4-cell	bta05217:Basal cell carcinoma	9	2.96E-02
4-cell	bta05221:Acute myeloid leukemia	9	3.60E-02
4-cell	bta04310:Wnt signaling pathway	17	3.85E-02
4-cell	bta04710:Circadian rhythm	4	4.67E-02
4-cell	bta05223:Non-small cell lung cancer	8	5.75E-02
4-cell	bta04070:Phosphatidylinositol signaling system	10	5.80E-02
4-cell	bta05410:Hypertrophic cardiomyopathy (HCM)	10	6.24E-02
4-cell	bta04370:VEGF signaling pathway	10	6.70E-02
4-cell	bta00562:Inositol phosphate metabolism	2	7.41E-02
4-cell	bta05412:Arrhythmogenic right ventricular cardiomyopathy (ARVC)	2	8.86E-02
4-cell	bta03018:RNA degradation	8	8.06E-02
4-cell	bta04010:MAPK signaling pathway	25	8.74E-02
4-cell	bta04912:GnRH signaling pathway	11	9.13E-02
4-cell	bta05410:Hypertrophic cardiomyopathy (HCM)	2	9.77E-02
4-cell	bta04350:TGF-beta signaling pathway	10	9.90E-02
8-cell	bta03010:Ribosome	50	3.03E-42
8-cell	bta05016:Huntington's disease	30	1.11E-08

8-cell	bta05012:Parkinson's disease	24	1.30E-07
8-cell	bta03050:Proteasome	14	2.21E-07
8-cell	bta05010:Alzheimer's disease	24	6.19E-06
8-cell	bta03040:Spliceosome	16	1.74E-03
8-cell	bta04260:Cardiac muscle contraction	11	2.74E-03
8-cell	bta03420:Nucleotide excision repair	10	9.53E-03
8-cell	bta00900:Terpenoid backbone biosynthesis	4	3.44E-02
8-cell	bta04320:Dorso-ventral axis formation	6	3.58E-02
8-cell	bta00970:Aminoacyl-tRNA biosynthesis	5	4.15E-02
8-cell	bta03020:RNA polymerase	5	5.13E-02
8-cell	bta04310:Wnt signaling pathway	20	5.24E-02
8-cell	bta04120:Ubiquitin mediated proteolysis	5	5.45E-02
8-cell	bta05213:Endometrial cancer	9	6.39E-02
8-cell	bta04623:Cytosolic DNA-sensing pathway	5	7.69E-02
16-cell	bta00280:Valine, leucine and isoleucine degradation	4	1.38E-02
16-cell	bta00330:Arginine and proline metabolism	4	2.47E-02
16-cell	bta05221:Acute myeloid leukemia	4	2.86E-02
16-cell	bta00471:D-Glutamine and D-glutamate metabolism	2	3.60E-02
16-cell	bta04640:Hematopoietic cell lineage	4	6.04E-02
16-cell	bta05322:Systemic lupus erythematosus	4	6.26E-02
16-cell	bta05200:Pathways in cancer	8	7.24E-02
16-cell	bta04810:Regulation of actin cytoskeleton	6	7.95E-02
16-cell	bta04512:ECM-receptor interaction	4	8.65E-02
16-cell	bta00072:Synthesis and degradation of ketone bodies	2	9.93E-02
Early Morula	bta00051:Fructose and mannose metabolism	2	9.85E-02
Early Morula	bta05211:Renal cell carcinoma	3	9.67E-02
Early Morula	bta04810:Regulation of actin cytoskeleton	7	3.85E-03
Early Morula	bta04660:T cell receptor signaling pathway	4	5.23E-02
Early Morula	bta04530:Tight junction	4	7.50E-02
Compact Morula	bta04530:Tight junction	5	8.63E-02
Blastocyst	bta03010:Ribosome	51	7.22E-40
Blastocyst	bta00190:Oxidative phosphorylation	57	9.83E-13
Blastocyst	bta05012:Parkinson's disease	48	2.78E-08
Blastocyst	bta03040:Spliceosome	26	7.95E-08
Blastocyst	bta05010:Alzheimer's disease	52	8.68E-07
Blastocyst	bta05016:Huntington's disease	54	2.20E-06
Blastocyst	bta04142:Lysosome	37	5.75E-05
Blastocyst	bta00510:N-Glycan biosynthesis	18	1.67E-04
Blastocyst	bta00520:Amino sugar and nucleotide sugar metabolism	16	2.46E-03
Blastocyst	bta00480:Glutathione metabolism	6	3.92E-03
Blastocyst	bta00970:Aminoacyl-tRNA biosynthesis	15	4.21E-03
Blastocyst	bta00563:Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	10	4.55E-03

Blastocyst	bta03050:Proteasome	15	1.08E-02
Blastocyst	bta00150:Androgen and estrogen metabolism	6	1.47E-02
Blastocyst	bta00010:Glycolysis / Gluconeogenesis	17	1.66E-02
Blastocyst	bta01040:Biosynthesis of unsaturated fatty acids	9	1.66E-02
Blastocyst	bta03410:Base excision repair	11	2.11E-02
Blastocyst	bta00020:Citrate cycle (TCA cycle)	11	2.66E-02
Blastocyst	bta03440:Homologous recombination	10	2.81E-02
Blastocyst	bta00140:Steroid hormone biosynthesis	6	5.09E-02
Blastocyst	bta00310:Lysine degradation	12	5.29E-02
Blastocyst	bta04260:Cardiac muscle contraction	18	5.59E-02
Blastocyst	bta00514:O-Mannosyl glycan biosynthesis	3	7.58E-02
Blastocyst	bta00980:Metabolism of xenobiotics by cytochrome P450	4	7.63E-02
Blastocyst	bta00620:Pyruvate metabolism	11	8.13E-02
Blastocyst	bta00983:Drug metabolism	10	9.10E-02
Blastocyst	bta04020:Calcium signaling pathway	8	9.29E-02
Blastocyst	bta02010:ABC transporters	11	9.45E-02
Blastocyst	bta00030:Pentose phosphate pathway	8	9.58E-02
Blastocyst	bta00750:Vitamin B6 metabolism	4	9.87E-02

Figure S2.1: All clusters of expression dynamics during early bovine in vivo embryo development.

Fig. S2.1

Figure S1. All clusters of expression dynamics during early bovine in vivo embryo development.

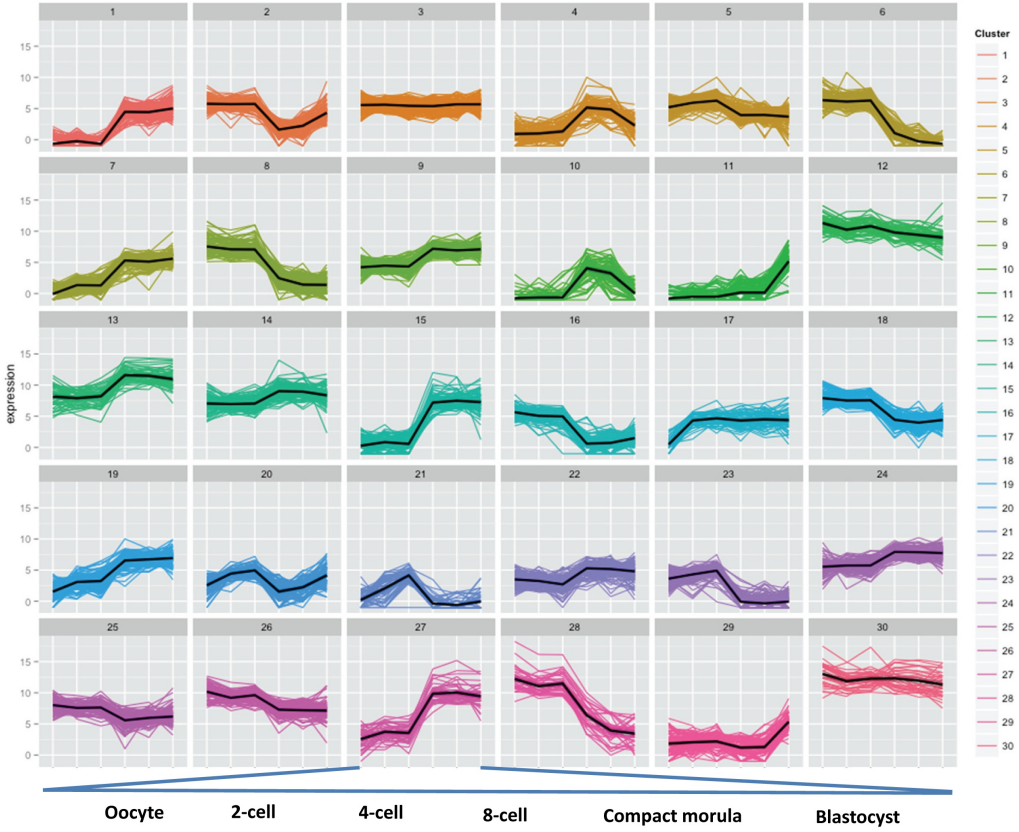
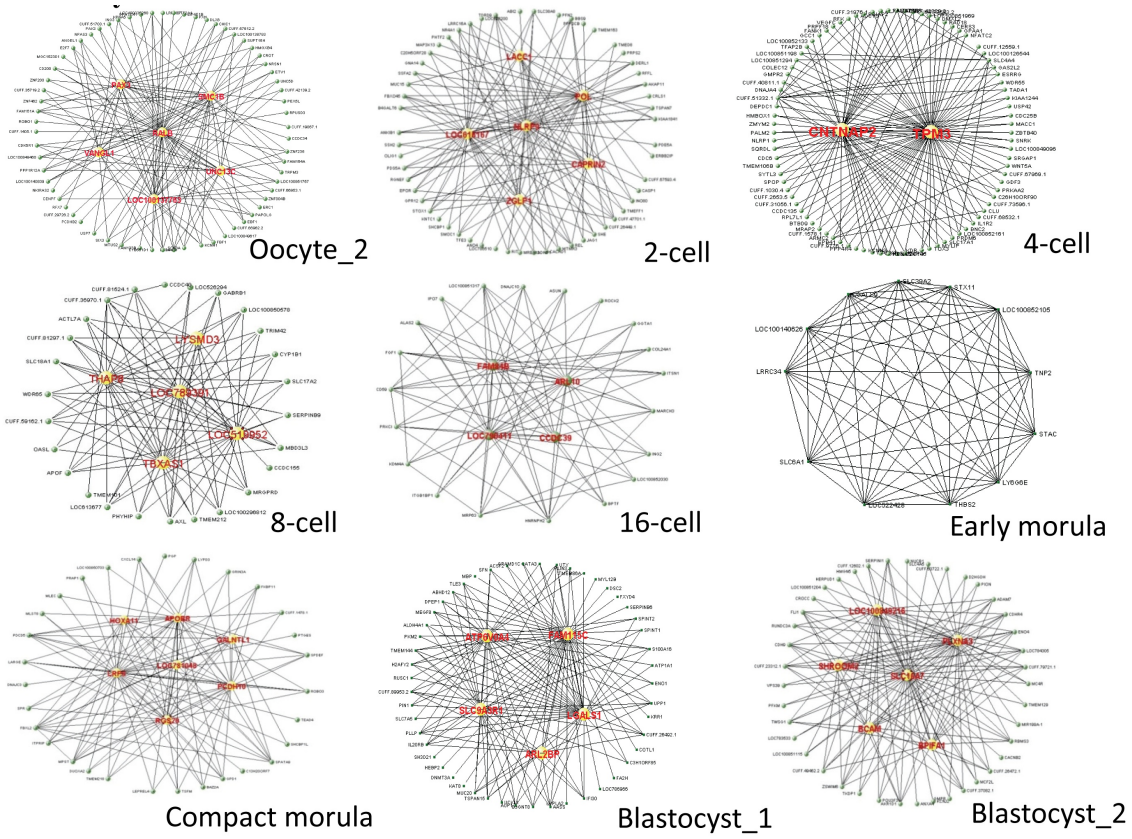


Figure S2.2: Visualization of representative highly correlated hub genes in bovine oocytes and embryos.

Fig. S2.2



Scripts for Bovine RNA-Seq Analysis:

1. Data trimming and quality control

1) Extract .fastq files from the solid raw files
    a. In the SOLiD output files, each lane contain (lane1 to 6) a .sqx file that needs to be decompress. Each lane subfolder also has a multiplex_x.txt file that list which samples are in that lane. For this example, that data was present in lane 5 and 6. And each of the samples, for example B2D, has a read file for lane 5 and for lane 6.
    b. To decompress the data, I used xsqtools that can be downloaded here: http://www.lifetechnologies.com/us/en/home/technical-resources/software-downloads/xsq-software.html. To decompress a lane, type 'convertFromXSQ.sh -c /scratch/solid_data/lane5/Run6_ZJ_YT_rnaSEQ_2013_01_17_1_05.xsq`. A few files will be extracted, but what we need from now on is the .fastq filed for each replicates. You should also rename each files to a more convenient filename. For example for B2D, I renamed the .fastq files from lane 5 and 6 to B2D_05.fastq and B2D_06.fastq.
2) Quality filtering and adapter trimming
    a. For this, you need 'cutadapt' (http://code.google.com/p/cutadapt/) and 'fastq_quality_trimmer' (http://hannonlab.cshl.edu/fastx_toolkit/).
    b. To cut the universal adapter, type 'cutadapt -a CGCCTTGGCCGTACAGCAG -n 4 B2D_05.fastq > tmp1.fastq'
    c. To cut the 5' primer, type 'cutadapt -g CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGGTGAT -n 4 tmp1.fastq > tmp2.fastq'
    d. To filter out low quality bases, type 'fastq_quality_trimmer -Q 33 -v -t 20 -l 30 -i tmp2.fastq -o B2D_05_filtered.fastq'
    e. You can check the quality of your reads after filtering unsing 'fastQC' (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/)
3) Download and index (if needed) reference genomes (Bos Taurus version Baylor Btau_4.6.1/bostau7))
    a. Because the cow genome was not indexed. I downloaded the full genome from http://hgdownload.cse.ucsc.edu/goldenPath/bosTau7/bigZips/bosTau7.fa.gz. Decompress the file and to index the genome with bowtie2 (see previous link), type 'bowtie2-build bosTau7.fa BT7'. BT7 will be the name of the index.
    b. You will also need the gene model file (.gtf) file. You will need also the gene model file (.gtf file). These can be found on the UCSC web server under the 'Table' browser -> your organism build -> refseq track -> table refgene -> output .gtf format
    c. To map the genomes, I used 'tophat2' with the following command line (both lane 5 and 6 are mapped together.: tophat2 -G BT7.gtf -o tophat2_B2D /ref_genomes/BT7 B2D_05_filtered.fastq, B2D_06_filtered.fastq
    d. Repeat for all samples.
4) Differential expression calculation with cuffdiff:
    a. cuffdiff mm10.gtf -o cuffdiff/ tophat2_B2D/accepted_hits.bam,/tophat2_A1D/accepted_hits.bam tophat2_B2J/accepted_hits.bam,/tophat2_A1J/accepted_hits.bam

All these can be performed on Galaxy web server (https://main.g2.bx.psu.edu/). Tutorial for RNAseq on galaxy here: https://docs.uabgrid.uab.edu/wiki/UAB_Galaxy_RNA_Seq_Step_by_Step_Tutorial

2. Mapping

RNA-Seq reads were mapped to Btau 4.6.1 assembly by using Tophat.
Example command line: tophat -p 16 -o /home/jol/JZL_RNA-Seq/cow/00C/1/ -G /home/jol/igenomes/Bos_taurus/NCBI/Btau_4.6.1/Annotation/Genes/genes.gtf /home/jol/igenomes/Bos_taurus/NCBI/Btau_4.6.1/Sequence/Bowtie2Index/genome /home/jol/JZL_RNA-Seq/00C_1.fastq

3. Transcriptome structure assembly

Individual mapped bam files are feed to Cufflinks to construct transcriptome models with the guidiance (-g option in Cufflinks) of Btau 4.6.1 NCBI bovine gene model.

Example command line: cufflinks -p 16 -g /home/jol/igenomes/Bos_taurus/NCBI/Btau_4.6.1/Annotation/Genes/genes.gtf -N -o /home/jol/JZL_RNA-Seq/cow/00C/1/ -L 00C_1 /home/jol/JZL_RNA-Seq/cow/00C/1/accepted_hits.bam
The results of each of these Cufflinks runs will contain a transcript.gtf file, which records all known genes and transcripts from NCBI gene model as well as novel genes and transcripts Cufflinks builds. These individual transcript.gtf files (16 in total) are used to converge to a master gene model in the next step.

4. Converge transcriptome structure from all replicates from all cell stages
Cuffmerge was used to converge individual transcript.gtf files to a master gene/transcript model/
Example command line: cuffmerge -s /home/jol/igenomes/Bos_taurus/NCBI/Btau_4.6.1/Sequence/WholeGenomeFasta/genome.fa -g /home/jol/igenomes/Bos_taurus/NCBI/Btau_4.6.1/Annotation/Genes/genes.gtf -p 16 ./assemblies.txt
The output of this Cuffmerge step is a merged.gtf file containing all known genes from NCBI as well as novel genes and transcripts built from the RNA-Seq data.
Then genes and transcrips mapped to uncertain chromosomes and contigs (i.e. chrUN) are filered out. The results are stored in filtered_merged.gtf after this filtering.

5. Quantify gene expression levels
The merged gene model and mapping result bam files from each RNA-Seq libraries are used to quantify the expression level of all genes (on gene level). This is achieved by using a Python package called HTSeq (http://www-huber.embl.de/users/anders/HTSeq/doc/count.html#count).
Example command line: python -m HTSeq.scripts.count -m intersection-nonempty -i gene_id /home/jol/JZL_RNA-Seq/cow/00C/1/accepted_hits.sam filtered_merged.gtf > 00C_1_merged_gn_readCounts.txt

6. Differential gene expression analyses
The files containing read counts for genes were used as input to DESeq (a R package) to perform differential gene expression analysis. Detailed procedures are recorded in the R log file. Please note, a number of DE analyses were performed by using appropriate read counts files depending on the compared cell stages, and in each comparison, only genes with sum of expression across all compared samples greater than 25% percentile were used in the particular comparison. For more detials about this, please refer to the user manual of DESeq.

7. PCA
```
library(scatterplot3d)
rna_exp <- read.table(file = 'E:/project/embryo_develop/expr/data/bull_expression.txt', header = TRUE, sep = '\t')
datExpr1 <- cbind(data.frame(gene = rna_exp$gene_name), rna_exp[, 2:(ncol(rna_exp) - 1)]) # 27393
datExpr2 <- datExpr1
# remove genes that had no expression at any stage
datExpr2 <- datExpr2[apply(datExpr2[, 2:ncol(datExpr2)], 1, function(x){if(length(which(x == 0)) == length(x))
return(FALSE) else return(TRUE)}), ] # 18318 genes
datExpr3 <- t(datExpr2[, 2:ncol(datExpr2)])
pca_rs <- prcomp(datExpr3, scale = T)
color <- c("blue4", "blue4", "chartreuse4", "chartreuse4", "firebrick1", "firebrick1", "darkviolet", "darkviolet",
"lightgreen", "lightgreen", "deeppink", "deeppink", "burlywood4", "burlywood4", "darkgreen", "darkgreen")
pch <- c(15, 15, 16, 16, 17, 17, 18, 18, 15, 15, 16, 16, 17, 17, 18, 18)
scatterplot3d(pca_rs$x[, 1:3], color = color, pch = pch)
legend(x=-2,y=5,c("Oocyte","2-cell","4-cell", "8-cell", "16-cell", "32-cell", "Compact Morula",
"Blastocyst"),cex=.8, pch=pch[c(1, 3, 5, 7, 9, 11, 13, 15)], col=color[c(1, 3, 5, 7, 9, 11, 13, 15)])
pc1 <- pca_rs$x[, 1]
pc2 <- pca_rs$x[, 2]
plot(pc1, pc2, col="white", xlab="PC1", ylab="PC2")
points(pc1[1:2], pc2[1:2], col="blue4", pch=15)
points(pc1[3:4], pc2[3:4], col="chartreuse4", pch=16)
points(pc1[5:6], pc2[5:6], col="firebrick1", pch=17)
points(pc1[7:8], pc2[7:8], col="darkviolet", pch=18)
points(pc1[9:10], pc2[9:10], col="lightgreen", pch=15)
points(pc1[11:12], pc2[11:12], col="deeppink", pch=16)
```

```
points(pc1[13:14], pc2[13:14], col="burlywood4", pch=17)
points(pc1[15:16], pc2[15:16], col="darkgreen", pch=18)
legend(x=-50,y=130,c("Oocyte","2-cell","4-cell", "8-cell", "16-cell", "32-cell", "Compact Morula",
"Blastocyst"),cex=.8, pch=pch[c(1, 3, 5, 7, 9, 11, 13, 15)], col=color[c(1, 3, 5, 7, 9, 11, 13, 15)])
setwd("~/Desktop/JZL/cow/")
```

8. DESeq

```
library(DESeq), library(ggplot2), library("RColorBrewer"), library("gplots"), library(reshape), library(pheatmap),
library(cluster)
read.table("geneIDTogeneName_uniq.txt", sep=" ", header=F) -> gene.name
rownames(gene.name) <- gene.name$V1
Take Comparison 1 as example.
comparison1.sampleID <- c("00C_1","00C_2","2C_1","2C_2")
comparison1.files <-
c("00C_1_merged_gn_readCounts.txt","00C_2_merged_gn_readCounts.txt","2C_1_merged_gn_readCounts.txt","2
C_2_merged_gn_readCounts.txt")
comparison1.conditions <- c("00C","00C","2C","2C")
comparison1.sample <- data.frame(cbind(comparison1.sampleID, comparison1.files, comparison1.conditions))
comparison1.cds <- newCountDataSetFromHTSeqCount(comparison1.sample, ".")
use <- (rowSums(counts(comparison1.cds)) >= 1)
comparison1.cds <- comparison1.cds[use, ]
use <- (rowSums(counts(comparison1.cds)) > quantile(rowSums(counts(comparison1.cds)), probs=0.25))
comparison1.cds <- comparison1.cds[use, ]
comparison1.cds <- estimateSizeFactors(comparison1.cds)
comparison1.cds <- estimateDispersions(comparison1.cds)
plotDispEsts(comparison1.cds)
comparison1.res <- nbinomTest(comparison1.cds, "00C", "2C")
plotMA(comparison1.res)
hist(comparison1.res$pval, breaks=100, col="skyblue", border="slateblue", main="")
colnames(comparison1.res) <- c("id", "baseMean", "00C", "2C", "foldChange", "log2FoldChange", "pval", "padj")
comparison1.DE <- comparison1.res[comparison1.res$pval <= 0.05, ]
comparison1.DE$gene.name <- gene.name[as.vector(comparison1.DE$id), 2]
write.table(comparison1.DE, "comparison1_DE.txt", quote=F, sep="\t", row.names=F)
comparison1.vsd <- varianceStabilizingTransformation(comparison1.cds)
select <- which(comparison1.res$pval <= 0.05)
hmcol = colorRampPalette(brewer.pal(9, "GnBu"))(100)
heatmap.2(exprs(comparison1.vsd)[select,], col = hmcol, trace="none", margin=c(10, 6))
# All samples
total.sampleID <-
c("00C_1","00C_2","2C_1","2C_2","4C_1","4C_2","8C_1","8C_2","16C_1","16C_2","32C_1","32C_2","CM_1","
CM_2","BL_1","BL_2")
total.files <-
c("00C_1_merged_gn_readCounts.txt","00C_2_merged_gn_readCounts.txt","2C_1_merged_gn_readCounts.txt","2
C_2_merged_gn_readCounts.txt","4C_1_merged_gn_readCounts.txt","4C_2_merged_gn_readCounts.txt","8C_1_m
erged_gn_readCounts.txt","8C_2_merged_gn_readCounts.txt","16C_1_merged_gn_readCounts.txt","16C_2_merge
d_gn_readCounts.txt","32C_1_merged_gn_readCounts.txt","32C_2_merged_gn_readCounts.txt","CM_1_merged_g
n_readCounts.txt","CM_2_merged_gn_readCounts.txt","BL_1_merged_gn_readCounts.txt","BL_2_merged_gn_rea
dCounts.txt")
total.conditions <-
c("00C","00C","2C","2C","4C","4C","8C","8C","16C","16C","32C","32C","CM","CM","BL","BL")
total.sample <- data.frame(cbind(total.sampleID, total.files, total.conditions))
total.cds <- newCountDataSetFromHTSeqCount(total.sample, "./")
use <- (rowSums(counts(total.cds)) > quantile(rowSums(counts(total.cds)), probs=0.25))
#total.cds <- total.cds[use, ]
total.cds <- estimateSizeFactors(total.cds)
```

```
cor(counts(total.cds, normalized=T)) -> m pheatmap(m, color = colorRampPalette(c("navy", "white",
"firebrick3"))(50), fontsize=14) counts(total.cds, normalized=T) -> total.expression
# All samples (drop 2C_2)
total.sampleID <-
c("00C_1","00C_2","2C_1","4C_1","4C_2","8C_1","8C_2","16C_1","16C_2","32C_1","32C_2","CM_1","CM_2",
"BL_1","BL_2")
total.files <-
c("00C_1_merged_gn_readCounts.txt","00C_2_merged_gn_readCounts.txt","2C_1_merged_gn_readCounts.txt","4
C_1_merged_gn_readCounts.txt","4C_2_merged_gn_readCounts.txt","8C_1_merged_gn_readCounts.txt","8C_2_m
erged_gn_readCounts.txt","16C_1_merged_gn_readCounts.txt","16C_2_merged_gn_readCounts.txt","32C_1_merg
ed_gn_readCounts.txt","32C_2_merged_gn_readCounts.txt","CM_1_merged_gn_readCounts.txt","CM_2_merged_
gn_readCounts.txt","BL_1_merged_gn_readCounts.txt","BL_2_merged_gn_readCounts.txt")
total.conditions <- c("00C","00C","2C","4C","4C","8C","8C","16C","16C","32C","32C","CM","CM","BL","BL")
total.sample <- data.frame(cbind(total.sampleID, total.files, total.conditions))
total.cds <- newCountDataSetFromHTSeqCount(total.sample, "./")
use <- (rowSums(counts(total.cds)) > quantile(rowSums(counts(total.cds)), probs=0.25))
#total.cds <- total.cds[use, ]
total.cds <- estimateSizeFactors(total.cds)
cor(counts(total.cds, normalized=T)) -> m
pheatmap(m, color = colorRampPalette(c("navy", "white", "firebrick3"))(50), fontsize=14)
counts(total.cds, normalized=T) -> total.expression
total.expression.df <- data.frame(total.expression)
colnames(total.expression.df) <- total.sampleID
total.expression.df$gene_name <- gene.name[rownames(total.expression.df), 2]
write.table(total.expression.df, "Total_expression.txt", sep="\t", row.names=T, col.names=T, quote=F)
# All samples (taking mean between replicates)
total.sampleID <-
c("00C_1","00C_2","2C_1","2C_2","4C_1","4C_2","8C_1","8C_2","16C_1","16C_2","32C_1","32C_2","CM_1","
CM_2","BL_1","BL_2")
total.files <-
c("00C_1_merged_gn_readCounts.txt","00C_2_merged_gn_readCounts.txt","2C_1_merged_gn_readCounts.txt","2
C_2_merged_gn_readCounts.txt","4C_1_merged_gn_readCounts.txt","4C_2_merged_gn_readCounts.txt","8C_1_m
erged_gn_readCounts.txt","8C_2_merged_gn_readCounts.txt","16C_1_merged_gn_readCounts.txt","16C_2_merge
d_gn_readCounts.txt","32C_1_merged_gn_readCounts.txt","32C_2_merged_gn_readCounts.txt","CM_1_merged_g
n_readCounts.txt","CM_2_merged_gn_readCounts.txt","BL_1_merged_gn_readCounts.txt","BL_2_merged_gn_rea
dCounts.txt")
total.conditions <-
c("00C","00C","2C","2C","4C","4C","8C","8C","16C","16C","32C","32C","CM","CM","BL","BL")
total.sample <- data.frame(cbind(total.sampleID, total.files, total.conditions))
total.cds <- newCountDataSetFromHTSeqCount(total.sample, "./")
#use <- (rowSums(counts(total.cds)) > quantile(rowSums(counts(total.cds)), probs=0.25))
#total.cds <- total.cds[use, ]
# Variance Stablization Transformation
total.cds.blind <- estimateSizeFactors(total.cds)
total.cds.blind <- estimateDispersions(total.cds.blind, method="blind")
pairs(counts(total.cds.blind, normalized=T), upper.panel=panel.cor)
pairs(log10(counts(total.cds.blind, normalized=T)+1), upper.panel=panel.cor)
total.vsd.blind <- varianceStabilizingTransformation(total.cds.blind)
select <- order(rowMeans(counts(total.cds.blind)), decreasing=T)[1:300]
hmcol = colorRampPalette(c("Green","black","red"))(100)
pheatmap(exprs(total.vsd.blind)[select, ], color=hmcol)
pairs(exprs(total.vsd.blind), upper.panel=panel.cor)
#dists <- as.matrix(dist(t(exprs(total.vsd.blind))))
#pheatmap(1-dists, color = colorRampPalette(c("navy", "white", "firebrick3"))(50), fontsize=14)
#pheatmap(cor(exprs(total.vsd.blind)), color = colorRampPalette(c("navy", "white", "firebrick3"))(50), fontsize=14)
# Focus on DE genes in 00C, 2C, 4C, 8C, CM and BL
```

```
t <-
unique(sort(c(as.vector(comparison1.DE$id),as.vector(comparison2.DE$id),as.vector(comparison3.DE$id),as.vecto
r(comparison6.DE$id),as.vector(comparison7.DE$id)))) # length of t is 2952
pheatmap(exprs(total.vsd.blind)[t, ], color = colorRampPalette(c("navy", "white", "firebrick3"))(50), fontsize=14)
pheatmap(exprs(total.vsd.blind)[t, c(1:8,13:16)], color = colorRampPalette(c("navy", "white", "firebrick3"))(50),
fontsize=14)
exprs(total.vsd.blind) -> m
m <-
cbind(apply(m[,1:2],1,mean),apply(m[,3:4],1,mean),apply(m[,5:6],1,mean),apply(m[,7:8],1,mean),apply(m[,9:10],1,
mean),apply(m[,11:12],1,mean),apply(m[,13:14],1,mean),apply(m[,15:16],1,mean))
colnames(m) <- c("00C", "2C", "4C", "8C", "16C", "32C", "CM", "BL")
pheatmap(cor(m), color = colorRampPalette(c("navy", "white", "firebrick3"))(50), fontsize=14) # This heatmap
makes best sense
pheatmap(m[t, ], color = colorRampPalette(c("navy", "white", "firebrick3"))(50), fontsize=14)
pheatmap(m[t, c(1,2,4,7,8)], color = colorRampPalette(c("navy", "white", "firebrick3"))(50), fontsize=14)
# Try hclust, maybe should use pam()
# hc <- hclust(dist(d))
# clusters <- cutree(hc, k=30)
# mat <- melt(d)
# Cluster.ID <- factor(rep(as.vector(clusters), 5), levels=as.character(1:max(clusters)))
# mat <- cbind(mat, Cluster.ID)
# ggplot(mat, aes(X2, value, group=X1, color=Cluster.ID)) + geom_line() +
scale_x_discrete(limits=c("00C","2C","8C","CM","BL")) + xlab("") + ylab("expression") +
scale_color_discrete(name="Cluster", breaks=as.character(1:30)) + facet_wrap(~Cluster.ID, nrow=5)
# ggplot(mat, aes(X2, value, group=X1, color=Cluster.ID)) + geom_line() + stat_summary(aes(x=X2, y=value,
group=1), fun.data=mean_cl_boot, color="black", fill="black", alpha=0.2, size=1.1, geom="smooth") +
scale_x_discrete(limits=c("00C","2C","8C","CM","BL")) + xlab("") + ylab("expression") +
scale_color_discrete(name="Cluster", breaks=as.character(1:30)) + facet_wrap(~Cluster.ID, nrow=5)
# Now try to cluster the 2952 DE genes
d <- m[t, c(1,2,3,4,7,8)]
#kmeans(d, centers=30, iter.max=100) -> k
pam(d, k=30) -> k # this is a better clustering algorithm
mat <- melt(d)
Cluster.ID <- factor(rep(as.vector(k$cluster), 6), levels=as.character(1:max(k$cluster)))
mat <- cbind(mat, Cluster.ID)
ggplot(mat, aes(X2, value, group=X1, color=Cluster.ID)) + geom_line() + stat_summary(aes(x=X2, y=value,
group=1), fun.data=mean_cl_boot, color="black", fill="black", alpha=0.2, size=1.1, geom="smooth") +
scale_x_discrete(limits=c("00C","2C","4C","8C","CM","BL")) + xlab("") + ylab("expression") +
scale_color_discrete(name="Cluster", breaks=as.character(1:30)) + facet_wrap(~Cluster.ID, nrow=5)
```

9. GO enrichment anaylses for each cluster (30 clusters in total)

```
library(topGO)
library(Rgraphviz)
# Prepare gene universe and GO annotation relationship
symbol2go <- read.table("symbol2go_cow.txt", header=F, sep="\t")
symbol2go$V1 <- as.character(symbol2go$V1)
symbol2go$V2 <- as.character(symbol2go$V2)
aggregate(symbol2go, by=list(symbol2go$V1), FUN=c) -> symbol2go.list
background.genes <- as.vector(unique(gene.name[names(k$cluster), 2]))
symbol2go.list[,c(1,3)] -> symbol2go.list
rownames(symbol2go.list) <- symbol2go.list$Group.1
colnames(symbol2go.list) <- c("gene", "go_id")
symbol2go.list <- symbol2go.list[background.genes, ]
gene2go <- as.list(symbol2go.list$go_id)
names(gene2go) <- symbol2go.list$gene
## For 1 cluster
```

```
# No test can be done for cluster 17
for(x in c(1:16,18:30)){
   print(c("Cluster", x))
   names(which(k$cluster==x)) -> tracking.ids
   test.genes <- as.vector(unique(gene.name[tracking.ids, 2]))
   geneList <- factor(as.integer(background.genes %in% test.genes))
   #geneList <- factor(as.integer(names(gene2go) %in% test.genes))
   names(geneList) <- names(gene2go)
   GOdata.BP <- new("topGOdata", ontology = "BP", allGenes = geneList, annot = annFUN.gene2GO, gene2GO =
gene2go)
test.stat <- new("classicCount", testStatistic = GOFisherTest, name = "Fisher test")
resultFisher <- getSigGroups(GOdata.BP, test.stat)
   GenTable(GOdata.BP, classic=resultFisher, topNodes=1000, numChar=500) -> BP.result
   BP.result <- BP.result[BP.result$classic <= 0.01, ]
   significant.genes <- vector()
      for(i in 1:dim(BP.result)[1]){
         term <- BP.result[i, "GO.ID"]
         annotated <- genesInTerm(GOdata.BP, term)[[1]]
         significant <- intersect(annotated, test.genes)
         significant.genes[i] <- paste(significant, collapse=",")
      }
   cbind(BP.result, significant.genes) -> BP.result
   #BP.result <- BP.result[which(p.adjust(BP.result[,"classic"],method="bonferroni")<=0.05),]
   #ggplot(BP.result, aes(Term, -log10(as.numeric(BP.result$classic)))) + geom_bar(stat="identity") +
theme(axis.text.x  = element_text(angle=90, vjust=0.5)) + scale_y_continuous(name="-log10(P-value)") +
scale_x_discrete(limits=BP.result$Term)
   prefix <- paste("Cluster", x, sep="_")
   printGraph(GOdata.BP, resultFisher, firstSigNodes = 30, fn.prefix = prefix, useInfo = "def", pdfSW = TRUE)
   fn <- paste(x, "Cluster_GO.txt", sep="_")
   write.table(BP.result, file=fn, quote=F, row.names=F, sep="\t")
}
# Get genes in each cluster and their expression levels
for(x in 1:30){
counts(total.cds.blind, normalized=T)[names(which(k$cluster==x)),] -> count.table
cbind(as.vector(gene.name[names(which(k$cluster==x)),2]), count.table) -> count.table
colnames(count.table)[1] <- "gene"
fn <- paste(x, "Cluster_expression.txt", sep="_")
write.table(count.table, file=fn, quote=F, row.names=T, sep="\t")
}
# Without VST
total.cds <- estimateSizeFactors(total.cds)
counts(total.cds, normalized=T) -> total.expression
cbind(apply(total.expression[,1:2], 1, mean),apply(total.expression[,3:4], 1, mean),apply(total.expression[,5:6], 1,
mean),apply(total.expression[,7:8], 1, mean),apply(total.expression[,9:10], 1, mean),apply(total.expression[,11:12],
1, mean),apply(total.expression[,13:14], 1, mean),apply(total.expression[,15:16], 1, mean)) -> total.expression
colnames(total.expression) <- c("00C", "2C", "4C", "8C", "16C", "32C", "CM", "BL")
cor(total.expression) -> m
pheatmap(m, color = colorRampPalette(c("navy", "white", "firebrick3"))(50), fontsize=14)
# No. of DE genes barplot
a <-
c(dim(comparison1.DE)[1],dim(comparison2.DE)[1],dim(comparison3.DE)[1],dim(comparison4.DE)[1],dim(comp
arison5.DE)[1],dim(comparison6.DE)[1],dim(comparison7.DE)[1])
b <- c("2C vs 00C", "4C vs 2C", "8C vs 4C", "16C vs 8C", "32C vs 16C", "CM vs 32C", "BL vs CM")
data.frame(a,b) -> d
```

```
ggplot(d, aes(factor(b), a)) + geom_bar(stat="identity", width=0.6, fill="skyblue") + scale_x_discrete(limits=d$b,
name="Developmental stages compared") + scale_y_continuous(name="No. of differentially expressed genes (P-
value <= 0.05)") + theme_bw()
# t <-
unique(sort(c(as.vector(comparison1.DE$id),as.vector(comparison3.DE$id),as.vector(comparison6.DE$id),as.vecto
r(comparison7.DE$id))))
# total.expression[t, ] -> total.DE
# total.DE <- total.DE[, c(1,2,4,7,8)]
# rownames(comparison1.res) <- comparison1.res$id
# rownames(comparison2.res) <- comparison2.res$id
# rownames(comparison3.res) <- comparison3.res$id
# rownames(comparison4.res) <- comparison4.res$id
# rownames(comparison5.res) <- comparison5.res$id
# rownames(comparison6.res) <- comparison6.res$id
# rownames(comparison7.res) <- comparison7.res$id
# common.tested <- intersect(intersect(intersect(comparison1.res$id, comparison3.res$id), comparison6.res$id),
comparison7.res$id)
# common.DE <- intersect(common.tested, t)
# common.foldchange <- cbind(comparison1.res[common.DE, 6], comparison3.res[common.DE, 6],
comparison6.res[common.DE, 6], comparison7.res[common.DE, 6])
# rownames(common.foldchange) <- common.DE
# colnames(common.foldchange) <- c("2C vs 00C", "8C vs 4C", "CM vs 32C", "BL vs CM")
# common.foldchange[intersect(which(is.infinite(common.foldchange)), which(common.foldchange > 0))] <- 10
# common.foldchange[intersect(which(is.infinite(common.foldchange)), which(common.foldchange < 0))] <- -10
# pheatmap(common.foldchange, cluster_cols=F)
# kmeans(total.DE, centers=16, iter.max=50) -> k
# total.DE[names(k$cluster[order(k$cluster)]),] -> reorder
# pheatmap(log10(reorder+0.1), cluster_rows=F, cluster_cols=F, show_rownames=F, color =
colorRampPalette(c("green", "black", "red"))(50))
# m <- melt(log10(total.DE+0.1))
# Cluster.ID <- factor(rep(as.vector(k$cluster), 5), levels=as.character(1:max(k$cluster)))
# m <- cbind(m, Cluster.ID)
# ggplot(m, aes(X2, value, group=X1, color=Cluster.ID)) + geom_line() +
scale_x_discrete(limits=c("00C","2C","8C","CM","BL")) + xlab("") + ylab("log10(expression)") +
scale_color_discrete(name="Cluster", breaks=as.character(1:20))
# ggplot(m, aes(X2, value, group=X1, color=Cluster.ID)) + geom_line() +
scale_x_discrete(limits=c("00C","2C","8C","CM","BL")) + xlab("") + ylab("log10(expression)") +
scale_color_discrete(name="Cluster", breaks=as.character(1:20)) + facet_wrap(~Cluster.ID, nrow=4)
Gene ontology enrichment analysis, using all GO annotated genes as background
library(topGO)
library(Rgraphviz)
# Prepare gene universe and GO annotation relationship
symbol2go <- read.table("symbol2go_cow.txt", header=F, sep="\t")
symbol2go$V1 <- as.character(symbol2go$V1)
symbol2go$V2 <- as.character(symbol2go$V2)
aggregate(symbol2go, by=list(symbol2go$V1), FUN=c) -> symbol2go.list
symbol2go.list[,c(1,3)] -> symbol2go.list
colnames(symbol2go.list) <- c("gene", "go_id")
gene2go <- as.list(symbol2go.list$go_id)
names(gene2go) <- symbol2go.list$gene

## For comparison 1 up and down regulated genes
test.genes.up <- as.vector(comparison1.DE[comparison1.DE$log2FoldChange >= 0, "gene.name"])
geneList <- factor(as.integer(names(gene2go) %in% test.genes.up))
names(geneList) <- names(gene2go)
```

```
GOdata.BP <- new("topGOdata", ontology = "BP", allGenes = geneList, annot = annFUN.gene2GO, gene2GO = gene2go)
test.stat <- new("classicCount", testStatistic = GOFisherTest, name = "Fisher test")
resultFisher <- getSigGroups(GOdata.BP, test.stat)
GenTable(GOdata.BP, classic=resultFisher, topNodes=38, numChar=500) -> BP.result
#BP.result <- BP.result[which(p.adjust(BP.result[,"classic"],method="bonferroni")<=0.05),]
ggplot(BP.result, aes(Term, -log10(as.numeric(BP.result$classic)))) + geom_bar(stat="identity") + theme(axis.text.x = element_text(angle=90, vjust=0.5)) + scale_y_continuous(name="-log10(P-value)") + scale_x_discrete(limits=BP.result$Term)
printGraph(GOdata.BP, resultFisher, firstSigNodes = 20, fn.prefix = "Comparison1_up_BP", useInfo = "def", pdfSW = TRUE)
write.table(BP.result, file="Comparison1_up_BP.txt", quote=F, row.names=F, sep="\t")
test.genes.down <- as.vector(comparison1.DE[comparison1.DE$log2FoldChange < 0, "gene.name"])
geneList <- factor(as.integer(names(gene2go) %in% test.genes.down))
names(geneList) <- names(gene2go)
GOdata.BP <- new("topGOdata", ontology = "BP", allGenes = geneList, annot = annFUN.gene2GO, gene2GO = gene2go)
test.stat <- new("classicCount", testStatistic = GOFisherTest, name = "Fisher test")
resultFisher <- getSigGroups(GOdata.BP, test.stat)
GenTable(GOdata.BP, classic=resultFisher, topNodes=46, numChar=500) -> BP.result
#BP.result <- BP.result[which(p.adjust(BP.result[,"classic"],method="bonferroni")<=0.05),]
ggplot(BP.result, aes(Term, -log10(as.numeric(BP.result$classic)))) + geom_bar(stat="identity") + theme(axis.text.x = element_text(angle=90, vjust=0.5)) + scale_y_continuous(name="-log10(P-value)") + scale_x_discrete(limits=BP.result$Term)
printGraph(GOdata.BP, resultFisher, firstSigNodes = 20, fn.prefix = "Comparison1_down_BP", useInfo = "def", pdfSW = TRUE)
write.table(BP.result, file="Comparison1_down_BP.txt", quote=F, row.names=F, sep="\t")
```

10. WGCNA

```
library(WGCNA)
library(pheatmap)
# The following setting is important, do not omit.
options(stringsAsFactors = FALSE);
# Allow multi-threading within WGCNA. This helps speed up certain calculations.
# At present this call is necessary for the code to work.
# Any error here may be ignored but you may want to update WGCNA if you see one.
enableWGCNAThreads()
rna_exp <- read.table(file = 'D:/other_people/jiang/bull_expression.txt', header = TRUE, sep = '\t')
datExpr1 <- cbind(data.frame(gene = rna_exp$gene_name), rna_exp[, 2:(ncol(rna_exp) - 1)]) # 27393
datExpr2 <- datExpr1
# remove genes that had no expression at any stage
datExpr2 <- datExpr2[apply(datExpr2[, 2:ncol(datExpr2)], 1, function(x){if(length(which(x == 0)) == length(x)) return(FALSE) else return(TRUE)}), ] # 18318 genes
# collapse same gene
row.names(datExpr2) <- 1:nrow(datExpr2)
datExpr2_colps <- collapseRows(datExpr2[, 2:ncol(datExpr2)], datExpr2$gene, 1:nrow(datExpr2))
datExpr3 <- cbind(data.frame(gene = row.names(datExpr2_colps$datETcollapsed)), datExpr2_colps$datETcollapsed)
row.names(datExpr3) <- 1:nrow(datExpr3)
# remove the first 25%
use <- (rowSums(datExpr3[, 2:ncol(datExpr3)]) > quantile(rowSums(datExpr3[, 2:ncol(datExpr3)]), probs=0.25))
datExpr3 <- datExpr3[use, ] # 13127
write.csv(datExpr3, file = 'D:/other_people/jiang/bull_expression_collapsed.txt', row.names = FALSE)
datExpr_merged <- data.frame(gene = datExpr3$gene)
datExpr_merged <- cbind(datExpr_merged, ooc = apply(datExpr3[, 2:3], 1, mean))
datExpr_merged <- cbind(datExpr_merged, c2 = apply(datExpr3[, 4:5], 1, mean))
```

```
datExpr_merged <- cbind(datExpr_merged, c4 = apply(datExpr3[, 6:7], 1, mean))
datExpr_merged <- cbind(datExpr_merged, c8 = apply(datExpr3[, 8:9], 1, mean))
datExpr_merged <- cbind(datExpr_merged, c16 = apply(datExpr3[, 10:11], 1, mean))
datExpr_merged <- cbind(datExpr_merged, c32 = apply(datExpr3[, 12:13], 1, mean))
datExpr_merged <- cbind(datExpr_merged, CM = apply(datExpr3[, 14:15], 1, mean))
datExpr_merged <- cbind(datExpr_merged, BL = apply(datExpr3[, 16:17], 1, mean))
datExpr <- t(datExpr3[, 2:ncol(datExpr3)])
tom <- TOMsimilarity((0.5 + 0.5 * cor(datExpr, use = 'p')) ^ 12)
tree <- flashClust(as.dist(1 - tom), method = 'average')
unmergedLabels = cutreeDynamic(dendro = tree, distM = 1-tom, deepSplit = 2, cutHeight = NULL, minClusterSize
= 30, pamRespectsDendro = FALSE );
unmergedColors = labels2colors(unmergedLabels)
sizeGrWindow(8,6)
plotDendroAndColors(tree, unmergedColors, "Dynamic Tree Cut", dendroLabels = FALSE, hang = 0.03, addGuide
= TRUE, guideHang = 0.05)
# compute module eigengenes
ME <- moduleEigengenes(datExpr, unmergedColors)
cor_ME <- cor(ME$eigengenes, use = 'p') # correlation among modules
# merge highly correlated modules
merge = mergeCloseModules(datExpr, unmergedLabels, MEs = ME, cutHeight = 0.3, verbose = 3)
# Numeric module labels
moduleLabels = merge$colors;
# Convert labels to colors
moduleColors = labels2colors(moduleLabels)
# label - color map
unique_lbls <- unique(moduleLabels)
lbl_clr_map <- data.frame(label = unique_lbls, color = rep('', length(unique_lbls)))
for (iLbl in 1:length(unique_lbls)) {
lbl_clr_map$color[iLbl] <- moduleColors[moduleLabels == unique_lbls[iLbl]][1]
}
sizeGrWindow(9,6)
plotDendroAndColors(tree, moduleColors, "Merged", dendroLabels = FALSE, hang = 0.03, addGuide = TRUE,
guideHang = 0.05)
bovine_module <- list()
bovine_module$unmergedLabels <- unmergedLabels
bovine_module$unmergedColors <- unmergedColors
bovine_module$tree <- tree
bovine_module$moduleLabels <- moduleLabels
bovine_module$moduleColors <- moduleColors
bovine_module$lbl_clr_map <- lbl_clr_map
bovine_module$merge <- merge
save(bovine_module, file = 'D:/other_people/jiang/rs/modules/unmerged2/module_bovine.RData')
oocyte_indicator <- c(rep(1, 2), rep(0, 14))
c2_indicator <- c(rep(0, 2), rep(1, 2), rep(0, 12))
c4_indicator <- c(rep(0, 4), rep(1, 2), rep(0, 10))
c8_indicator <- c(rep(0, 6), rep(1, 2), rep(0, 8))
c16_indicator <- c(rep(0, 8), rep(1, 2), rep(0, 6))
c32_indicator <- c(rep(0, 10), rep(1, 2), rep(0, 4))
CM_indicator <- c(rep(0, 12), rep(1, 2), rep(0, 2))
BL_indicator <- c(rep(0, 14), rep(1, 2))
bovine_module$mudule_sage_cor <- corAndPvalue(merge$newMEs, cbind(oocyte_indicator, c2_indicator,
c4_indicator, c8_indicator, c16_indicator, c32_indicator, CM_indicator, BL_indicator))
save(bovine_module, file = 'E:/project/embryo_develop/expr/rs/modules/unmerged2/module_bovine.RData')
sizeGrWindow(9,9)
# plotCor(cor_stage, new = FALSE)
```

```
colnames(bovine_module$mudule_sage_cor$cor) <- c('Oocyte', '2-cell', '4-cell', '8-cell', '16-cell', 'Early Morula',
'Compact Morula', 'Blastocyst')
pheatmap(bovine_module$mudule_sage_cor$cor[c(3, 4, 1, 2, 6, 5, 14, 17, 15, 8, 10, 9), ], cluster_rows = FALSE,
cluster_cols = FALSE,display_numbers = TRUE, number_format = "%.2f")
write.csv(bovine_module$mudule_sage_cor$cor[c(3, 4, 1, 2, 6, 5, 14, 17, 15, 8, 10, 9), ], file =
'E:/project/embryo_develop/expr/rs/modules/unmerged2/cor_r.csv')
write.csv(bovine_module$mudule_sage_cor$p[c(3, 4, 1, 2, 6, 5, 14, 17, 15, 8, 10, 9), ], file =
'E:/project/embryo_develop/expr/rs/modules/unmerged2/cor_pv.csv')
plot_modules = c(1, 15, 12, 54, 34, 6, 39, 13, 17, 29, 35, 18, 2, 4, 23, 53)
for (iM in 1:length(plot_modules)) {
module_expr <- datExpr3[bovine_module$moduleLabels == plot_modules[iM], ]
module_expr_merged <- datExpr_merged[bovine_module$moduleLabels == plot_modules[iM], ]
rowsum1 <- rowSums(module_expr[, 2:ncol(module_expr)])
use1 <- rowsum1 > quantile(rowsum1, probs = 0.1) & rowsum1 < quantile(rowsum1, probs = 0.9)
#write.csv(module1_expr, file = 'D:/other_people/jiang/rs/modules/unmerged2/module1_expr.csv', row.names =
FALSE)
jpeg(paste('Y:/other_people/jiang/rs/modules/unmerged2/module', plot_modules[iM], '_heatmap.jpeg', sep = ''))
pheatmap(log10(t(module_expr[use1, 2:ncol(module_expr)]) + 1), cluster_rows = FALSE, cluster_cols = FALSE)
dev.off()
jpeg(paste('Y:/other_people/jiang/rs/modules/unmerged2/module', plot_modules[iM], '_boxplot.jpeg', sep = ''))
boxplot(log10(module_expr_merged[use1, 2:ncol(module_expr_merged)] + 1)) dev.off()
}
# line plot
#plot_modules <- c(1, 12, 54, 34, 29, 35, 18, 2)
plot_modules <- c(1, 15, 12, 54, 34, 6, 29, 35, 18, 2, 4, 23)
plot_value <- matrix(rep(NA, length(plot_modules) * (ncol(datExpr_merged) - 1)), nrow = length(plot_modules))
for (iM in 1:length(plot_modules)) {
module_expr <- datExpr3[bovine_module$moduleLabels == plot_modules[iM], ]
module_expr_merged <- datExpr_merged[bovine_module$moduleLabels == plot_modules[iM], ]
rowsum1 <- rowSums(module_expr[, 2:ncol(module_expr)])
use1 <- rowsum1 > quantile(rowsum1, probs = 0.1) & rowsum1 < quantile(rowsum1, probs = 0.9)
plot_value[iM,] <- log10(colMeans(module_expr_merged[use1, 2:ncol(module_expr_merged)]))
}
rownames(plot_value) <- plot_modules
#write.table(plot_value, file = 'D:/other_people/jiang/rs/modules/unmerged2/line_plot_value.csv', sep = ',',
row.names = FALSE, col.names = FALSE)
write.table(plot_value, file = 'D:/other_people/jiang/rs/modules/unmerged2/line_plot_value_ex.csv', sep = ',',
col.names = FALSE)
par(col = 'black', mfrow = c(1, 1))
plot(c(1, (ncol(module_expr_merged) - 1)), range(plot_value), type = 'n',
main = 'Gene expression profiles of stage specific modules across different stages', xlab = 'Developemental stage',
xaxt = 'n', ylab = 'log10 of expression value')
colors <- rainbow(length(plot_modules))
linetype <- c(1:length(plot_modules))
plotchar <- seq(15, 15 + length(plot_modules), 1)
for (iM in 1:length(plot_modules)) {
lines(1:(ncol(module_expr_merged) - 1), plot_value[iM, ], type = 'b', lwd = 3, lty = linetype[iM], col = colors[iM],
pch = plotchar[iM], xlab = '', xaxt = 'n')
}
x_lbls <- c('Oocyte', '2-cell', '4-cell', '8-cell', '16-cell', 'Early Morula', 'Compact Morula', 'Blastocyst')
axis(1, at = 1:(ncol(module_expr_merged) - 1), labels = FALSE)
text(x = 1:(ncol(module_expr_merged) - 1), labels = x_lbls, par("usr")[3] - 0.1 , srt = 45, pos = 1, xpd = TRUE)
# add a legend
legend(1, max(plot_value), c('Oocyte', '2-cell', '4-cell', '8-cell', '16-cell', 'Early Morula', 'Compact Morula',
'Blastocyst'), cex=0.8, col=colors, font = 24, pch=plotchar, lty=linetype, title="Modules")
# module 1
```

```
module1_expr <- datExpr3[bovine_module$moduleLabels == 1, ]
module1_expr_merged <- datExpr_merged[bovine_module$moduleLabels == 1, ]
rowsum1 <- rowSums(module1_expr[, 2:ncol(module1_expr)])
use1 <- rowsum1 > quantile(rowsum1, probs = 0.1) & rowsum1 < quantile(rowsum1, probs = 0.9)
#write.csv(module1_expr, file = 'D:/other_people/jiang/rs/modules/unmerged2/module1_expr.csv', row.names =
FALSE)
pheatmap(log10(t(module1_expr[use1, 2:ncol(module1_expr)]) + 1), cluster_rows = FALSE, cluster_cols =
FALSE)
jpeg('Y:/other_people/jiang/rs/modules/unmerged2/module1_boxplot.jpeg')
boxplot(log10(module1_expr_merged[use1, 2:ncol(module1_expr_merged)] + 1))
dev.off()
# module 15
module1_expr <- datExpr3[bovine_module$moduleLabels == 1, ]
```

2.7. References

1 Schultz, R. M. The molecular foundations of the maternal to zygotic transition in the preimplantation embryo. *Human reproduction update* **8**, 323-331 (2002).

2 Schultz, R. M. Regulation of zygotic gene activation in the mouse. *BioEssays : news and reviews in molecular, cellular and developmental biology* **15**, 531-538, doi:10.1002/bies.950150806 (1993).

3 Hamatani, T., Carter, M. G., Sharov, A. A. & Ko, M. S. Dynamics of global gene expression changes during mouse preimplantation development. *Developmental cell* **6**, 117-131 (2004).

4 Wang, Q. T. *et al.* A genome-wide study of gene activity reveals developmental signaling pathways in the preimplantation mouse embryo. *Developmental cell* **6**, 133-144 (2004).

5 Braude, P., Bolton, V. & Moore, S. Human gene expression first occurs between the four- and eight-cell stages of preimplantation development. *Nature* **332**, 459-461, doi:10.1038/332459a0 (1988).

6 Telford, N. A., Watson, A. J. & Schultz, G. A. Transition from maternal to embryonic control in early mammalian development: a comparison of several species. *Molecular reproduction and development* **26**, 90-100, doi:10.1002/mrd.1080260113 (1990).

7 Memili, E. & First, N. L. Developmental changes in RNA polymerase II in bovine oocytes, early embryos, and effect of alpha-amanitin on embryo development. *Molecular reproduction and development* **51**, 381-389, doi:10.1002/(SICI)1098-2795(199812)51:4<381::AID-MRD4>3.0.CO;2-G (1998).

8 Misirlioglu, M. *et al.* Dynamics of global transcriptome in bovine matured oocytes and preimplantation embryos. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 18905-18910, doi:10.1073/pnas.0608247103 (2006).

9 Xie, D. *et al.* Rewirable gene regulatory networks in the preimplantation embryonic development of three mammalian species. *Genome research* **20**, 804-815, doi:10.1101/gr.100594.109 (2010).

10 Graf, A. *et al.* Fine mapping of genome activation in bovine embryos by RNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 4139-4144, doi:10.1073/pnas.1321569111 (2014).

11 Kues, W. A. *et al.* Genome-wide expression profiling reveals distinct clusters of transcriptional regulation during bovine preimplantation development in vivo. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 19768-19773, doi:10.1073/pnas.0805616105 (2008).

12 Driver, A. M. *et al.* RNA-Seq analysis uncovers transcriptomic variations between morphologically similar in vivo- and in vitro-derived bovine blastocysts. *BMC genomics* **13**, 118, doi:10.1186/1471-2164-13-118 (2012).

13 Chitwood, J. L., Rincon, G., Kaiser, G. G., Medrano, J. F. & Ross, P. J. RNA-seq analysis of single bovine blastocysts. *BMC genomics* **14**, 350, doi:10.1186/1471-2164-14-350 (2013).

14 Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews. Genetics* **10**, 57-63, doi:10.1038/nrg2484 (2009).

15 Ozsolak, F. & Milos, P. M. RNA sequencing: advances, challenges and opportunities. *Nature reviews. Genetics* **12**, 87-98, doi:10.1038/nrg2934 (2011).

16 Hayakawa, H., Hirai, T., Takimoto, A., Ideta, A. & Aoyagi, Y. Superovulation and embryo transfer in Holstein cattle using sexed sperm. *Theriogenology* **71**, 68-73, doi:10.1016/j.theriogenology.2008.09.016 (2009).
17 Lee, W. *et al.* Influence of Factors during Superovulation on Embryo Production in Korean Holstein Cattle. *Journal of Veterinary Medical Science* **74**, 167-174, doi:10.1292/jvms.11-0057 (2012).
18 Smith, S. L. *et al.* Global gene expression profiles reveal significant nuclear reprogramming by the blastocyst stage after cloning. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 17582-17587, doi:10.1073/pnas.0508952102 (2005).
19 Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols* **7**, 562-578, doi:10.1038/nprot.2012.016 (2012).
20 Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome biology* **11**, R106, doi:10.1186/gb-2010-11-10-r106 (2010).
21 Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics* **9**, 559, doi:10.1186/1471-2105-9-559 (2008).
22 Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Statistical applications in genetics and molecular biology* **4**, Article17, doi:10.2202/1544-6115.1128 (2005).
23 Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. (New York: Springer, 2001).
24 Langfelder, P., Zhang, B. & Horvath, S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* **24**, 719-720, doi:10.1093/bioinformatics/btm563 (2008).
25 Xue, Z. *et al.* Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* **500**, 593-597, doi:10.1038/nature12364 (2013).
26 Langfelder, P., Luo, R., Oldham, M. C. & Horvath, S. Is My Network Module Preserved and Reproducible? *Plos Comput Biol* **7**, doi:ARTN e1001057DOI 10.1371/journal.pcbi.1001057 (2011).
27 Hu, Z., Mellor, J., Wu, J. & DeLisi, C. VisANT: an online visualization and analysis tool for biological interaction data. *BMC bioinformatics* **5**, 17, doi:10.1186/1471-2105-5-17 (2004).
28 Alexa, A., Rahnenfuhrer, J. & Lengauer, T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* **22**, 1600-1607, doi:10.1093/bioinformatics/btl140 (2006).
29 Tu, Q., Cameron, R. A., Worley, K. C., Gibbs, R. A. & Davidson, E. H. Gene structure in the sea urchin Strongylocentrotus purpuratus based on transcriptome analysis. *Genome research* **22**, 2079-2087, doi:10.1101/gr.139170.112 (2012).
30 Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**, 44-57, doi:10.1038/nprot.2008.211 (2009).
31 Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 14863-14868 (1998).

32 Langfelder, P., Luo, R., Oldham, M. C. & Horvath, S. Is my network module preserved and reproducible? *PLoS computational biology* **7**, e1001057, doi:10.1371/journal.pcbi.1001057 (2011).
33 Everts-van der Wind, A. *et al.* A high-resolution whole-genome cattle-human comparative map reveals details of mammalian chromosome evolution. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 18526-18531, doi:10.1073/pnas.0509285102 (2005).
34 Elsik, C. G. *et al.* The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science* **324**, 522-528, doi:10.1126/science.1169588 (2009).
35 Van Soom, A. *et al.* Timing of compaction and inner cell allocation in bovine embryos produced in vivo after superovulation. *Biology of reproduction* **57**, 1041-1049 (1997).
36 Niakan, K. K., Han, J., Pedersen, R. A., Simon, C. & Pera, R. A. Human pre-implantation embryo development. *Development* **139**, 829-841, doi:10.1242/dev.060426 (2012).
37 Huang, S. Non-genetic heterogeneity of cells in development: more than just noise. *Development* **136**, 3853-3862, doi:10.1242/dev.035139 (2009).
38 Ito, K. & Suda, T. Metabolic requirements for the maintenance of self-renewing stem cells. *Nature reviews. Molecular cell biology* **15**, 243-256, doi:10.1038/nrm3772 (2014).
39 Li, L., Zheng, P. & Dean, J. Maternal control of early mouse development. *Development* **137**, 859-870, doi:10.1242/dev.039487 (2010).
40 Huang, W. & Khatib, H. Comparison of transcriptomic landscapes of bovine embryos using RNA-Seq. *BMC genomics* **11**, 711, doi:10.1186/1471-2164-11-711 (2010).
41 Lonergan, P. *et al.* Temporal divergence in the pattern of messenger RNA expression in bovine embryos cultured from the zygote to blastocyst stage in vitro or in vivo. *Biology of reproduction* **69**, 1424-1431, doi:10.1095/biolreprod.103.018168 (2003).
42 Gutierrez-Adan, A. *et al.* Effect of speed of development on mRNA expression pattern in early bovine embryos cultured in vivo or in vitro. *Molecular reproduction and development* **68**, 441-448, doi:10.1002/mrd.20113 (2004).
43 Wrenzycki, C. *et al.* Messenger RNA expression patterns in bovine embryos derived from in vitro procedures and their implications for development. *Reproduction, fertility, and development* **17**, 23-35 (2005).
44 Barnes, F. L. & Eyestone, W. H. Early cleavage and the maternal zygotic transition in bovine embryos. *Theriogenology* **33**, 141-152, doi:http://dx.doi.org/10.1016/0093-691X(90)90605-S (1990).
45 Farin, P. W., Piedrahita, J. A. & Farin, C. E. Errors in development of fetuses and placentas from in vitro-produced bovine embryos. *Theriogenology* **65**, 178-191, doi:10.1016/j.theriogenology.2005.09.022 (2006).
46 Warburg, O. On the origin of cancer cells. *Science* **123**, 309-314 (1956).
47 Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029-1033, doi:10.1126/science.1160809 (2009).
48 Ledgard, A. M., Lee, R. S. & Peterson, A. J. Expression of genes associated with allantois emergence in ovine and bovine conceptuses. *Molecular reproduction and development* **73**, 1084-1093, doi:10.1002/mrd.20532 (2006).
49 Moreau, J., Lebreton, S., Iouzalen, N. & Mechali, M. Characterization of Xenopus RalB and its involvement in F-actin control during early development. *Developmental biology* **209**, 268-281, doi:10.1006/dbio.1999.9254 (1999).

50 Peschard, P. *et al.* Genetic deletion of RALA and RALB small GTPases reveals redundant functions in development and tumorigenesis. *Current biology : CB* **22**, 2063-2068, doi:10.1016/j.cub.2012.09.013 (2012).
51 Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-257 (1999).
52 Li, E. Chromatin modification and epigenetic reprogramming in mammalian development. *Nature reviews. Genetics* **3**, 662-673, doi:10.1038/nrg887 (2002).
53 Beyhan, Z., Forsberg, E. J., Eilertsen, K. J., Kent-First, M. & First, N. L. Gene expression in bovine nuclear transfer embryos in relation to donor cell efficiency in producing live offspring. *Molecular reproduction and development* **74**, 18-27, doi:10.1002/mrd.20618 (2007).
54 Niemann, H. & Lucas-Hahn, A. Somatic cell nuclear transfer cloning: practical applications and current legislation. *Reproduction in domestic animals = Zuchthygiene* **47 Suppl 5**, 2-10, doi:10.1111/j.1439-0531.2012.02121.x (2012).
55 Barros, C. M. *et al.* Effect of superstimulatory treatments on the expression of genes related to ovulatory capacity, oocyte competence and embryo development in cattle. *Reproduction, fertility, and development* **25**, 17-25, doi:10.1071/RD12271 (2012).
56 Chu, T., Dufort, I. & Sirard, M. A. Effect of ovarian stimulation on oocyte gene expression in cattle. *Theriogenology* **77**, 1928-1938, doi:10.1016/j.theriogenology.2012.01.015 (2012).
57 Mundim, T. C. *et al.* Changes in gene expression profiles of bovine embryos produced in vitro, by natural ovulation, or hormonal superstimulation. *Genetics and molecular research : GMR* **8**, 1398-1407, doi:10.4238/vol8-4gmr646 (2009).
58 Urrego, R., Rodriguez-Osorio, N. & Niemann, H. Epigenetic disorders and altered in gene expression after use of Assisted Reproductive Technologies in domestic cattle. *Epigenetics : official journal of the DNA Methylation Society* **9** (2014).
59 Mapletoft, R. J., Steward, K. B. & Adams, G. P. Recent advances in the superovulation in cattle. *Reproduction, nutrition, development* **42**, 601-611 (2002).

Chapter Three

mRNA Levels of Imprinted Genes in Bovine In Vivo Oocytes, Embryos and Cross Species Comparisons in Humans, Mice and Pigs

Scientific Reports, July 2015, accepted with revision.

(Jiang Z, Dong H, Zheng X, Donovan D, Chen J &Tian XC)

3.1. Abstract

Twenty-six confirmed imprinted genes in the bovine were quantified in in vivo produced oocytes and embryos. Eighteen were detectable and their transcriptional abundance were categorized into five patterns: largely decreased (*MEST* and *PLAGL1)*; first decreased and then increased (*CDKN1C* and *IGF2R)*; peaked at a specific stage (*PHLDA2*, *SGCE, PEG10* and *PEG3*); low until peaked in blastocysts (*GNAS, MEG3, DGAT1*, *ASCL2*, *NNAT*, and *NAP1L5)*; and finally constant low (*DIRAS3*, *IGF2*, *H19* and *RTL1)*. A number of genes were expressed at surprisingly high levels. For instance, the mRNAs for the paternally expressed *MEST*, and to a lesser extent *PLAGL1,* were highly abundant in oocytes and could only be expressed from the maternal allele. On the contrary, *GNAS* and *MEG3,* both maternally expressed, were barely detectable in bovine oocytes but highly expressed in blastocysts, suggesting that their genomic imprints were not established/recognized until much later in development. Furthermore, we compared these genes to their counterparts in mice, humans and pigs. We found differences in the imprinting status, levels and dynamics of gene expression among these four species. The data presented here will serve as a reference base for expression profiles of imprinted genes by embryos produced from assisted reproductive biotechnologies.

Key words: Imprinted genes, Pre-implantation development, Bovine, Mouse, Human, Pig

3.2. Introduction

Genomic imprinting is a series of precisely regulated epigenetic processes that cause genes to be expressed in a parental-origin-specific manner in mammals [1]. Proper allelic expression of imprinted genes plays important roles in embryonic and placental development as well as in maternal behavior [2]. The exact numbers of total imprinted genes and their roles in mammalian development remain as open questions. It is estimated that approximately 150 and 100 imprinted genes are present in the mouse and human, respectively *(http://igc.otago.ac.nz/Search.html* and *http://www.mousebook.org/imprinting-gene-list)*. The identification of imprinted genes in farm species, however, lags behind, with a total of 28, 17, and 10 confirmed, imprinted genes in cattle, pigs, and sheep, respectively [3].

The specific genes imprinted in each species can also be very different. For example, 50 imprinted genes in the mouse overlap with those in humans, and numerous genes imprinted in the mouse and/or human are not imprinted in the other species or in farm animals [3]. Moreover, the timing of imprinting activation during development is also species- and developmental stage-specific. Genomic imprinting is seen in mouse embryos as early as the two-cell stage, and by the blastocyst stage, monoallelic expression of most imprinted genes is observed [4]. On the contrary, the allelic expression status of most imprinted genes is not known in human embryos [5]. To date, the onset of monoallelic expression of imprinted genes in farm species has not been examined systematically but occurs much later than the blastocyst stage in the bovine, ovine and porcine [6-10].

The exact nature of genetic imprints is still relatively uncertain. It is known that genomic imprinting is regulated through epigenetic mechanisms, including allele-specific DNA methylation at differentially methylated regions established during gametogenesis and

embryogenesis [11] and maintained in subsequent cell divisions during pre-implantation development [12]. Changes in global DNA methylation in zygotes and early embryos from several species have been studied and found to differ dramatically. For example, the male pronucleus is nearly completely demethylated in the mouse and rat, partial demethylated in cattle and goats, and minimally demethylated in sheep and pigs [13]. In addition to DNA methylation, imprinting regulation likely involves other epigenetic modifications and mechanisms such as histone modifications, chromatin architecture, and non-coding RNAs [3,12].

Because imprints are established during gametogenesis and embryo development [14], environmental factors such as in vitro culture and manipulations of oocytes and embryos could affect their expression and these changes can last through the entire life of the animals [15]. Increased incidences of imprinting disorders including large offspring syndrome (LOS) in ruminants and Beckwith-Wiedemann Syndrome (BWS) in humans have been reported in assisted reproductive technologies (ART) where in vitro culture of oocytes/embryos are conducted [16-20]. Because in vitro and in vivo embryos have step-wise differences in transcriptomes [21], qualitative and quantitative expression profiles of imprinted genes from in vivo pre-implantation embryos are essential gold standards for embryos produced from biotechnologies.

To date, the expression of only a selected few imprinted genes have been characterized and shown to be regulated in a tissue- and/or developmental stage-specific manner across species including humans [22], porcine [10] and bovine [6]. Benefited from the most comprehensive RNA sequencing (RNA-seq) profiles of pre-implantation embryos from the pig [23], cattle [24], mouse [25] and human [25,26] , we set forth to analyze the expression of 26 bovine imprinted genes in cattle and compared them across these four species during oocyte and embryo development. Our data

provide important evidence for stage- and species-specific regulation of imprinting during pre-implantation development.

3.3. Methods

The expression profiles of bovine in vivo derived oocytes and pre-implantation embryos were characterized by RNA-seq and published recently [21]. Briefly two biological replicates of in vivo produced bovine oocytes and embryos at the 2-, 4-, 8-, early morula, late morula and blastocyst stages were subjected to RNA-seq at the depth of approximately 30 million reads per sample. To analyze species differences, three other RNA-seq datasets of pre-implantation development from the human, mouse and pig were downloaded from Gene Expression Omnibus (GEO) *(www.ncbi.nlm.nih.gov/geo)* under the accession numbers GSE44183 [25], and SRA076823 [23]. All oocytes and embryos used in these studies were in vivo derived with the exception of those from humans (Table S3.1). For each embryonic stage, data were normalized among the four species by transforming uniquely mapped reads to RPKM [27]. The 26 genes that have been confirmed to be imprinted in the bovine were examined in the bovine as well as in humans, mice and pigs regardless of their imprinting status in these species [3] (Table S3.2). Expression profiles of these genes were searched against all four datasets and the RPKM values of each gene from the same developmental stage were averaged and analyzed among four species. All genes with RPKM > 0.1 were defined as detectable.

3.4. Results

The total number of imprinted genes in the bovine genome is still unknown. From the 28 confirmed imprinted genes in the bovine, *MAOA* and *XIST* were excluded because they are only imprinted in the placenta. The mRNA abundance of imprinted genes in MII oocytes and embryos were compared within and amongst four different species, bovine, humans, mice and pigs. Overall, 18 of the 26 confirmed bovine imprinted genes were expressed in bovine in vivo oocytes and/or pre-implantation embryos (Table S3.3), while only 14, 12 and 9 of these were expressed in humans, mouse and pigs, respectively. Among them, the levels of six genes, cyclin-dependent kinase inhibitor 1C *(CDKN1C),* GNAS complex locus *(GNAS),* insulin-like growth factor 2 receptor *(IGF2R),* mesoderm specific transcript *(MEST),* pleckstrin homology-like domain, family A, member 2 *(PHLDA2)* and pleiomorphic adenoma gene-like 1 *(PLAGL1),* were highly expressed (RPKM > 10) in bovine oocytes or in at least one bovine embryonic stages (Table S3.3) while others are expressed at relatively low levels.

The changes of the 18 expressed, imprinted genes were categorized into 5 different dynamic patterns according to their abundance. The first group, including the paternally expressed *MEST* (also known as the paternally expressed gene 1 (*PEG1*)) and *PLAGL1*, represents genes that underwent an overall trend of decrease in expression levels during pre-implantation development in the bovine. Incidentally, the same trend was also seen in the other three species (Figure 3.1). Specifically, *MEST* and *PLAGL1* were high in oocytes, dramatically decreased from MII oocytes to the 8-cell stage and then maintained a low but detectable level up to the blastocyst stage. While having an overall similar decrease trend, the transition to low expression occurred at different stages in the other species. Specifically, the mouse and pig had earlier decreases at the 2- or 4- cell stage while the change in humans was similar to that in cattle

(8-cell). Notably, *MEST* was the highest expressed gene among all genes studied in all four species. *PLAGL1* was also seen at relatively high levels in oocytes from all species examined. Among all the genes studied, these two genes were also consistent in their imprinting status among the four species. Also interesting to note, these genes are paternally expressed, it is reasonable to expect that the maternal alleles of these genes in the oocytes carry expression-inhibitory imprints established during gametogenesis and low levels of these genes should be seen in oocytes. Therefore their high levels in the oocytes suggest that the genomic imprints of the maternal alleles of these genes are either not established or not recognized at this stage of development.

The second dynamic expression pattern, including the maternally expressed *CDKN1C* and *IGF2R,* represents genes that first decreased and then increased their expression levels (Figure 3.2). Interestingly, the expression dynamics of these two genes were different in the other three species studied. In oocytes of the human, mouse and pig only low levels of *CDKN1C* and *IGF2R* were found. The levels then increased and then decreased in the human and mouse but continued to increase in the pig. Of further note, *CDKN1C* was no detectable in pig oocytes or embryos (Figure 3.2).

The third group represents genes that first increased, then decreased and subsequently maintained a relatively constant level from oocytes to blastocysts in the bovine (Figure 3.3A, B). For example, the maternally expressed *PHLDA2,* the paternally expressed sarcoglycan epsilon (*SGCE*) and the paternally expressed gene 10 (*PEG10*) all peaked at the 2- to 4-cell stages, while the paternally expressed gene 3 (*PEG3*) peaked at the 8-cell stage. Interestingly, members of this group were not expressed in other species. For instance *PHLDA2*, *PEG10*, and *PEG3* were not detectable in pigs, and *PEG10* was not expressed in mice. These and other drastically

different expression patterns shown in Fig. 3 suggest species variations of genetic imprinting during early embryo development.

The fourth group, including the maternally expressed *GNAS*, maternally expressed gene 3 (*MEG3*), diacylglycerol O-acyltransferase 1 (*DGAT1*), achaete-scute family bHLH transcription factor 2 (*ASCL2*), and the paternally expressed neuronatin (*NNAT*) as well as nucleosome assembly protein 1-like 5 (*NAP1L5*), contained genes that maintained relatively low expression levels until peaking at the morula or blastocyst stage in the bovine (Figure 3.4A, B). Major species differences were also seen in the expression dynamics of these genes. For example, *DGAT1* peaked at the zygotic stage in humans and 4-cell stage in pigs. The maternally expressed *MEG3* and *GNAS* accumulated in the mature mouse oocytes yet barely detectable in the bovine oocytes. This group also contained the most inconsistency in imprinting status among the four species. For example, *GNAS* was not imprinted in either the human or pig and *DGAT1* is only imprinted in the bovine.

The last group, including the paternally expressed DIRAS family, GTP-binding RAS-like 3 (*DIRAS3*), insulin-like growth factor 2 (*IGF2*) and retrotransposon-like 1 (*RTL1*), as well as the maternally expressed and non-coding *H19,* included genes that maintained relatively constant low levels of expression throughout all stages studied in bovine oocytes and embryos (Figure 3.5). Interestingly, these genes were not as silent in other species. For example, *DIRAS3* had an extremely high expression level in the pig morulae and was relatively highly expressed in multiple stages of human early embryos. High levels of the paternally expressed *IGF2* was observed in the mouse oocytes yet complete absent in pigs. Of note, neither *H19* nor *RTL1* were detected in the oocytes or embryos from the human, mouse and pig. The near undetectable

levels of *H19* in all stages of bovine embryos and absence from oocytes/embryos of other species are consistent with their absence in the gonads.

In addition to these categories based on the bovine abundance patterns, eight genes that are imprinted and were expressed in other species were not detectable in the bovine oocytes/embryos. Three, including small nuclear ribonucleoprotein polypeptide N (*SNRPN*), tumor suppressing subtransferable candidate 4 (*TSSC4*), ubiquitin specific peptidase 29 (*USP29*), and four, including *SNRPN*, *TSSC4, USP29*, antisense transcript gene of *PEG3* (*APEG3*), were imprinted and expressed in human and mouse embryos (Table 3.1 and Table S3.4), respectively.

3.5. Discussion

Imprinted genes play critical roles in normal fetal and placental development. Interestingly, gene imprinting is not only developmental stage-specific, but also species-specific. Mammalian genomic imprinting has primarily been studied in mice and humans, while only limited information is available in livestock species. Due to species variations, most information gained in the mouse/human cannot be extended to other species. In this study, we provide the first comprehensive description of total transcript levels of currently known and confirmed bovine imprinted genes during in vivo embryonic development in the bovine and in three other mammalian species. We showed that the profiles, number and identity of bovine imprinted genes that are expressed during pre-implantation may not be the same in embryos of other species.

Using reverse transcription polymerase chain reaction (RT-PCR) and uniparental embryos, selected imprinted genes such as *MEST, SGCE,* and *NNAT* were found to be bi-allelically expressed in bovine in vitro/vivo blastocysts [6,28]. These results were enhanced in our study by using a more powerful throughput technology, RNA-seq, which provides the quantitative expression of imprinted genes at multiple stages of in vivo development. We found that *MEST* and *GNAS* are the most abundant genes in early oocytes/embryos across all species studied, suggesting their conserved roles in early development. Although we were not able to distinguish the specific parental alleles from which the genes were expressed, the fact that none of the bovine imprinted gene studied to date establishes mono-allelic expression by the blastocyst stage suggest that the gene expression we quantified was a combination of accumulation from the maternal allele in the oocytes and from both parental alleles in early embryos. The counter-intuitive levels of several genes, such as *MEST* and *PLAGL1*, both

paternally expressed yet highly abundant in bovine oocytes; and *PHLDA2, GNAS, MEG3, DGAT1* and *ASCL2*, all maternally expressed yet barely detectable in bovine oocytes, are intriguing. They suggest either the lack of genomic imprints on the maternal alleles of these genes or that these imprints are not recognized. Indeed, differential methylation at the imprinting control region of several genes including those characterized here, *PLAGL1* and *PEG3*, were not established during gametogenesis in non-human primates [29]. The late onset of mono-allelic expression of imprinted genes in cattle suggests that genomic imprints may also be established post-fertilization.

All 18 expressed bovine imprinted genes had developmental stage-specific patterns of expression. Their unique dynamics may provide insights into their specific roles in the developing embryos. For example, *MEST* and *PLAGL1* were highly expressed early in development and then decreased, indicating potential roles in oocytes, fertilization, or initial cleavage events. *PHLDA2* peaked between 2- and 4-cell stages and decreased subsequently. Although the exact role of *PHLDA2* during embryo development is unclear, when siRNA specific to *PHLDA2* was injected into bovine zygotes a substantial increase in blastocyst development was resulted [30]. These observations, together with ours suggest that *PHLDA2* may inhibit embryonic development during the later pre-implantation period and is therefore selectively down-regulated. *CDKN1C* is another imprinted gene whose expression in pre-implantation development in cattle was confirmed by functional studies. Injection of *CDKN1C*-specific siRNA into one-cell zygotes resulted in a 45% reduction in blastocyst development [30], an observation consistent with our finding that it was up-regulated starting from the 16-cell stage after the initial degradation from the levels in the oocytes. The decrease in *PHLDA2* and increase in *CDKN1C* thus ensure proper blastocyst development.

A relatively large number of the bovine genes studied, 8 out of 18, either peaked or increased at the blastocyst stage. These include the maternally expressed *CDKN1C, IGF2R*, *GNAS, MEG3, DGAT1, ASCL2* and the paternally expressed *NNAT* and *NAP1L5* [31-33]. As we found previously [21] a wave of increased gene expression occurs during the morula to blastocyst transition in the bovine. These eight genes may be up-regulated to prepare the bovine embryos to undergo differentiation and further development. Interestingly, these genes have very different expression dynamics among different species. These patterns may reflect the differences in the speed of development and the timing of maternal-zygotic transition and differentiation among the species studied.

Eight bovine imprinted genes were not detectable in bovine oocytes or pre-implantation embryos while being expressed at relatively high levels in certain stages in other species. For example, *SNRPN, TSSC4* and *USP29,* were both imprinted and expressed in human and mouse pre-implantation embryos. Because monoallelic expression of imprinted genes is tissue- and developmental stage-specific [4,12,34], these genes may not play a role in bovine pre-implantation development yet are important in other species at these stages.

Lastly, we also noted differences in gene expression between in vivo and in vitro produced embryos. For example, *SNRPN* and *TSSC4* were undetectable here but detected in bovine in vitro embryos [24]. *H19*, *IGF2* and *PEG10* were undetectable in in vivo embryos in pigs, however, expression of these genes has been observed in pig in vitro blastocysts [10]. Evidences have also been shown that in vitro culture and somatic cell nuclear transfer affect the establishment of *SNRPN* imprinting [9]. These differences further demonstrate the notion that culture conditions can induce anomalies in genomic imprinting and the need that in vivo embryos should be used for establishment of gold standards of expression dynamics.

In summary, we provide here a reference base for the expression levels of imprinted genes in bovine in vivo produced oocytes and early embryos and contrasted these patterns with those in other species. The exact nature of genetic imprints and the timing of their establishment during early development are yet to be examined systematically. The connection between genomic imprints and mono-allelic expression will be of major focus for our future studies.

Table 3.1. Imprinting status in humans, mice and pigs for the eight genes that are imprinted but undetectable in bovine oocytes and embryos.

Bovine imprinted genes	Human	Mouse	Pig
APEG3	Not imprinted	Imprinted	Not imprinted
SNRPN	Imprinted	Imprinted	Not imprinted
TSSC4	Imprinted	Imprinted	Not imprinted
USP29	Imprinted	Imprinted	Not imprinted
NESP55	Not imprinted	Not imprinted	-
MAGEL2	-	-	Not imprinted
KCNQ1OT1	-	-	Not imprinted
MIMT1	-	Not imprinted	Not imprinted

" – ": Information on imprinting status is not available.

Figure 3.1. Levels of transcriptional expression of bovine imprinted genes that continuously decreased during pre-implantation development. Paternally expressed genes are labeled in blue.

Fig. 3.1

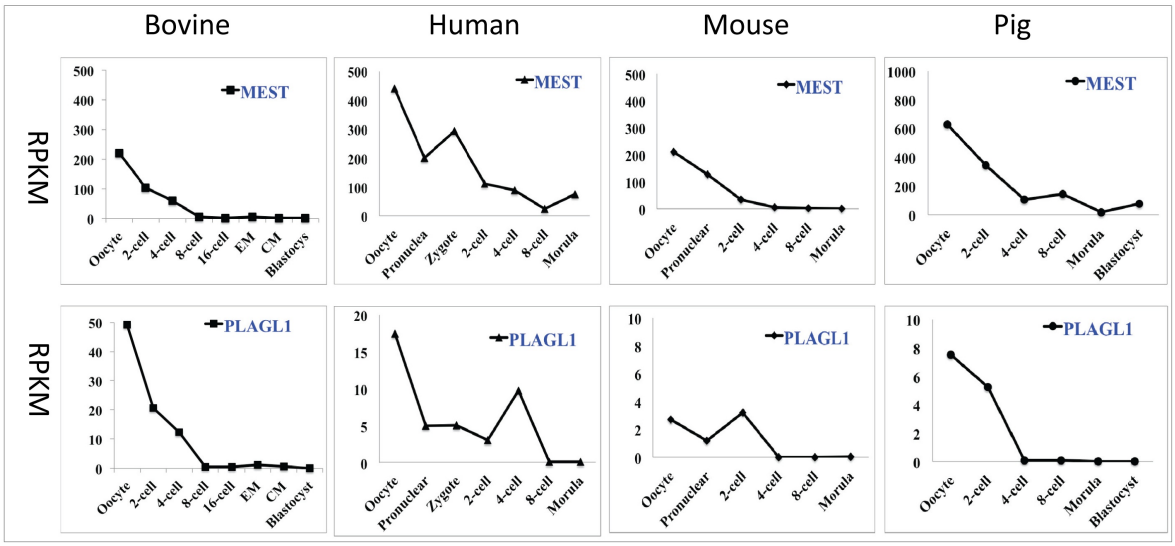


Figure 3.2. Transcriptional expression of bovine imprinted genes that were decreased first and then increased during bovine pre-implantation development. Maternally expressed genes are labeled in pink and genes that are not imprinted in a particular species are labeled in black. The lack of a graph indicates that the gene was not detected in that species.

Fig. 3.2

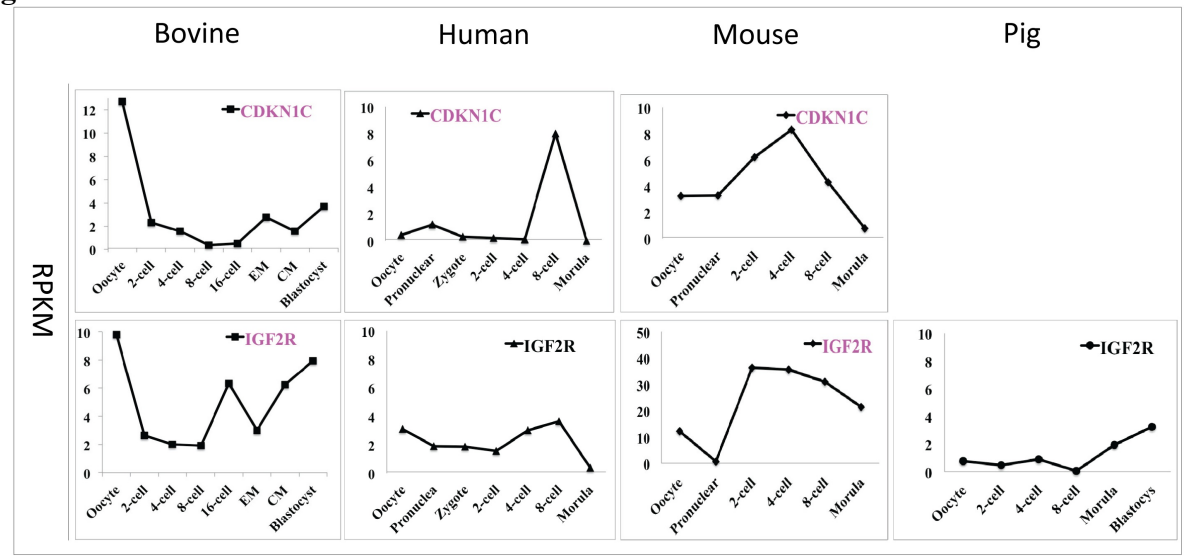


Figure 3.3. Transcriptional expression of bovine imprinted genes that increased first and then decreased at the 2- or 4-cell stage (A), or at the 8-cell stage (B). Maternally and paternally expressed genes are labeled in pink and blue, respectively. Genes that are not imprinted in a particular species are labeled in black. The lack of a graph indicates that the gene was not detected in that species.

Fig. 3.3A

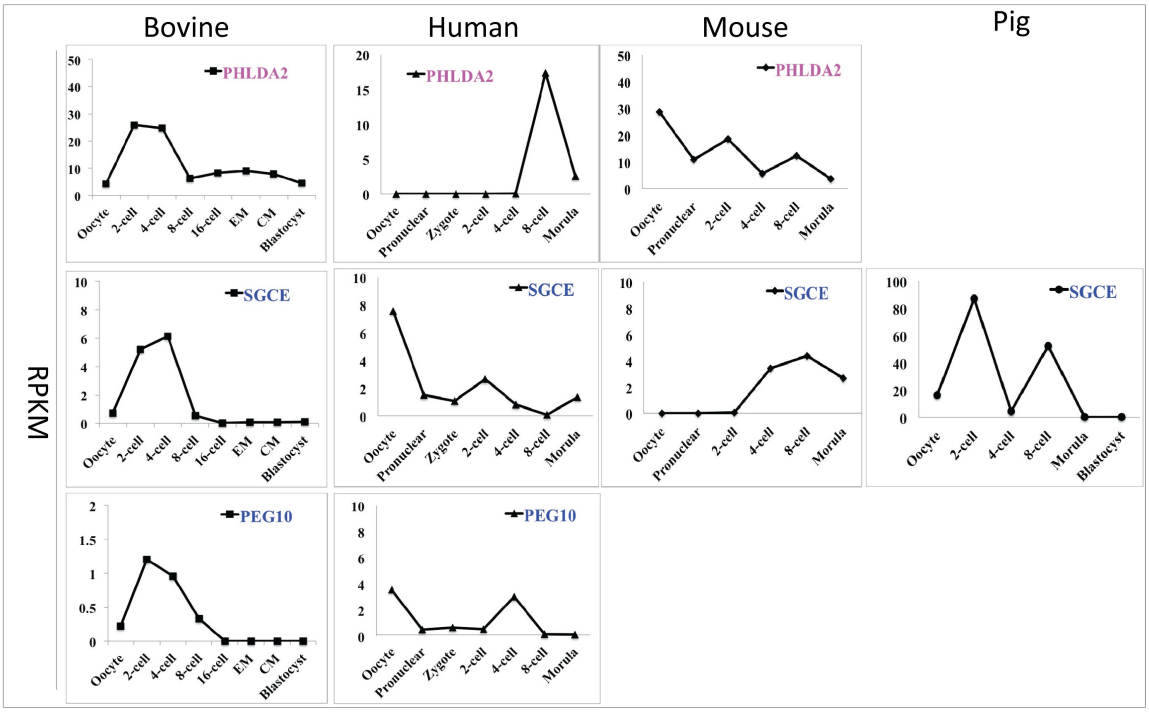


Fig. 3.3B

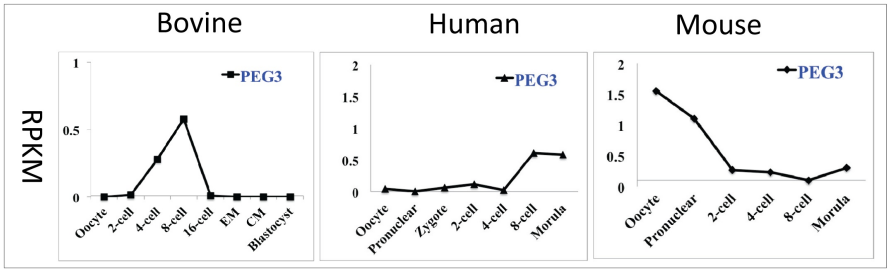


Figure 3.4. Transcriptional expression of bovine imprinted genes that maintained relatively low expression and then peaked at blastocysts to high (A) or low levels (B). Maternally and paternally expressed genes are labeled in pink and blue, respectively. Genes that are not imprinted in a particular species are labeled in black. The lack of a graph indicates that the gene was not detected in that species.

Fig. 3.4A

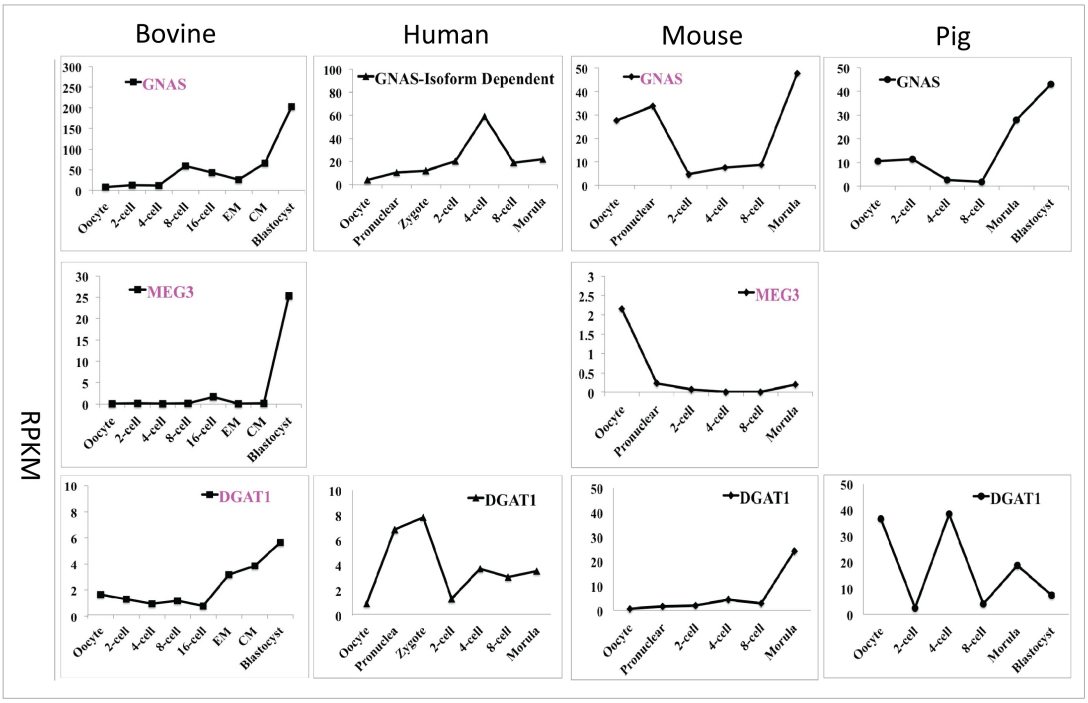


Fig. 3.4B

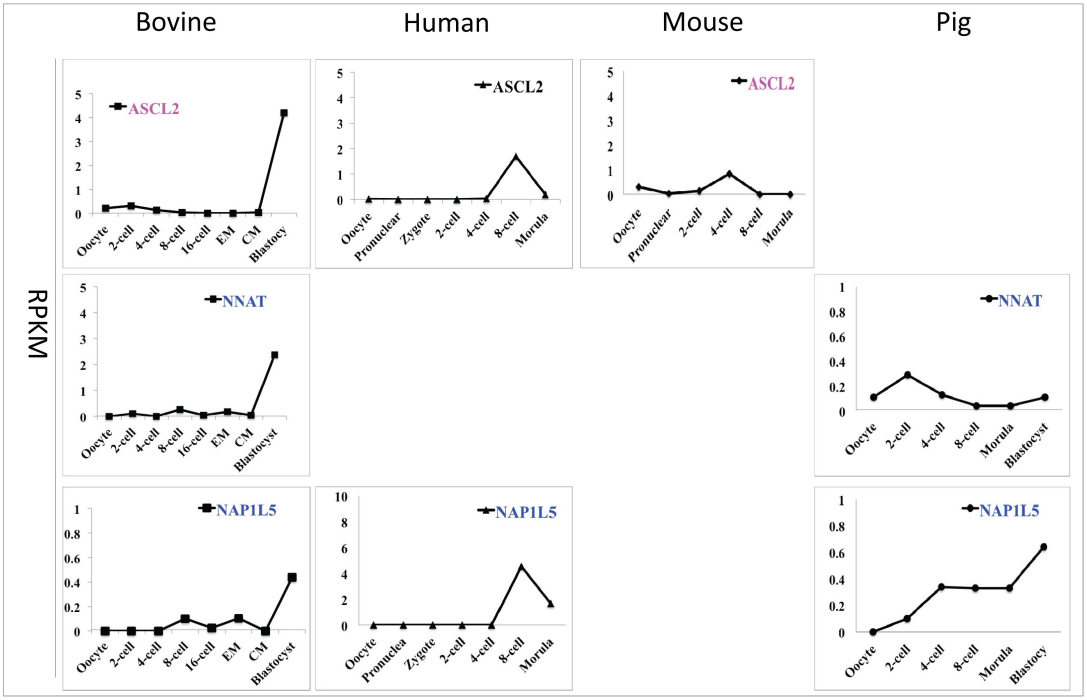
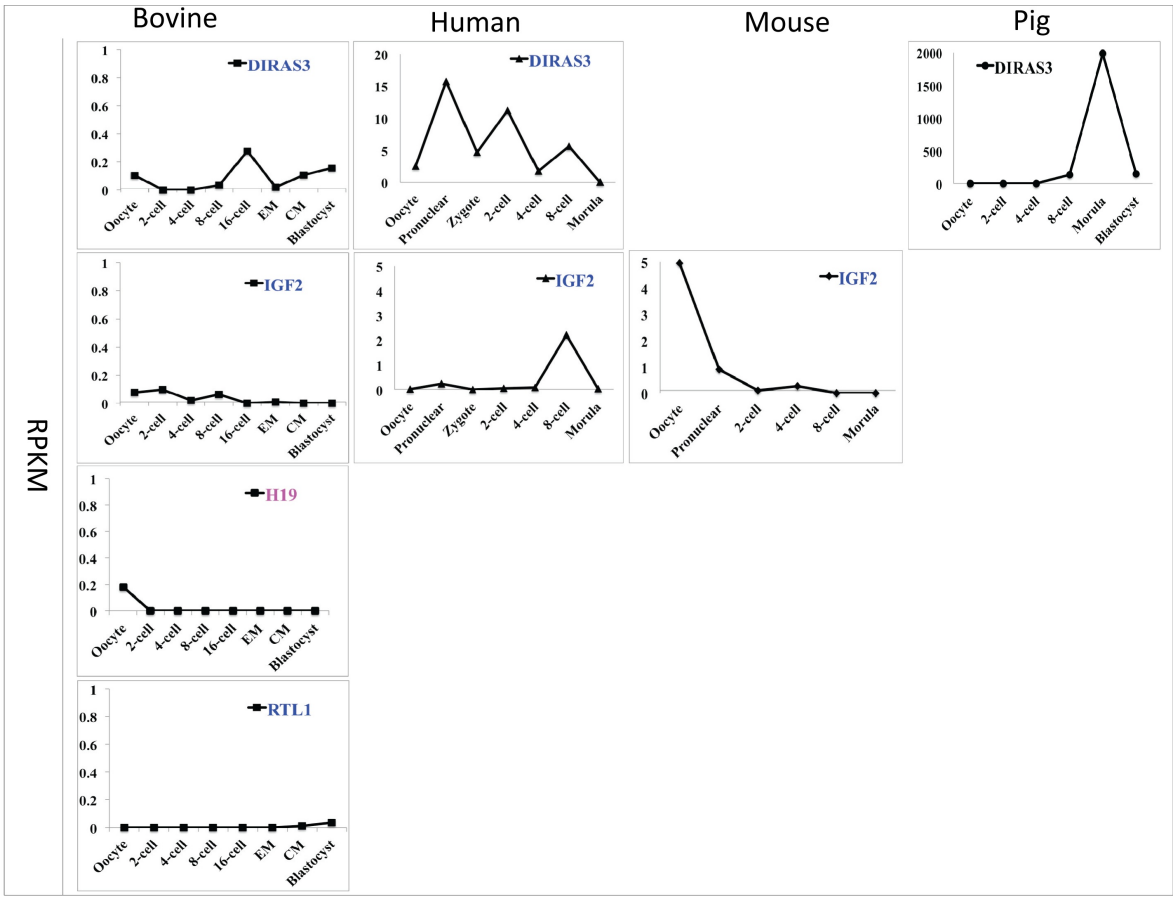


Figure 3.5. Transcriptional expression of imprinted genes that maintained low expression during pre-implantation development. Maternally and paternally expressed genes are labeled in pink and blue, respectively. Genes that are not imprinted in a particular species are labeled in black. The lack of a graph indicates that the gene was not detected in that species.

Fig. 3.5



3.6. Supplementary Information

Table S3.1. Summary of the numbers of biological replicates used in the four datasets of this study.

species / replicate / stage	Cattle*	Porcine (4-10 oocytes/embryos per group)	Mouse*	Human*
Oocyte	2	2	2	3
Pronuclei	-	-	3	3
Zygotes	-	-	-	2
2-cell	2	2	3	2
4-cell	2	2	3	4
8-cell	2	2	3	11
16-cell	2	-	-	-
Mourla	Early morula: 2 Compact morula: 2	2	3	3
Blastocyst	2	2	-	-

*Single oocyte/embryo per biological replicate was used in the studies of cattle, mice and humans. All oocytes/embryos were in vivo produced with the exception of the humans.

Table S3.2. The names, parental expression patterns and chromosomal locations of confirmed imprinted genes in cattle

Symbol	Gene Name	Expressed allele	Chromosome
APEG3	antisense transcript gene of PEG3	P	18
ASCL2	achaete-scute family bHLH transcription factor 2	M	29
CDKN1C	cyclin-dependent kinase inhibitor 1C	M	29
DGAT1	diacylglycerol O-acyltransferase 1	M	14
DIRAS3	DIRAS family, GTP-binding RAS-like 3	P	3
GNAS	GNAS complex locus	M	13
H19	H19	M	29
IGF2	insulin-like growth factor 2	P	29
IGF2R	insulin-like growth factor 2 receptor	M	9
KCNQ1OT1	KCNQ1 opposite strand/antisense transcript 1	P	29
MAGEL2	MAGE-like 2	P	21
MEG3	maternally expressed 3	M	21
MEST	mesoderm specific transcript	P	4
MIMT1	MER1 repeat containing imprinted transcript 1	P	18
NAP1L5	nucleosome assembly protein 1-like 5	P	6
NESP55	Neuroendocrine secretory protein-55	M	13
NNAT	neuronatin	P	13
PEG10	paternally expressed 10	P	4
PEG3	paternally expressed 3	P	18
PHLDA2	pleckstrin homology-like domain, family A, member 2	M	29
PLAGL1	pleiomorphic adenoma gene-like 1	P	9
RTL1	retrotransposon-like 1	P	21
SGCE	sarcoglycan, epsilon	P	4
SNRPN	small nuclear ribonucleoprotein polypeptide N	M	21
TSSC4	tumor suppressing subtransferable candidate 4	M	29
USP29	ubiquitin specific peptidase 29	P	18

Table S3.3. Spreadsheet 1. Expression levels of imprinted genes in bovine in vivo oocytes and pre-implantation embryos

Gene Name	Status	Expressed allele	Oocyte	2-cell	4-cell	8-cell	16-cell	Early Morula	Compact Morula	Blastocyst
ASCL2	**Imprinted**	**M**	0.22	0.32	0.14	0.03	0.00	0.00	0.04	4.19
CDKN1C	**Imprinted**	**M**	12.72	2.34	1.56	0.36	0.50	2.78	1.57	3.71
DGAT1	**Imprinted**	**M**	1.66	1.30	0.96	1.20	0.79	3.17	3.86	5.64
DIRAS3	**Imprinted**	**P**	0.10	0.00	0.00	0.03	0.28	0.02	0.11	0.16
GNAS	**Imprinted**	**M**	7.58	12.85	11.48	59.18	43.11	25.72	66.07	202.81
H19	**Imprinted**	**M**	0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00
IGF2	**Imprinted**	**P**	0.08	0.10	0.02	0.06	0.00	0.01	0.00	0.00
IGF2R	**Imprinted**	**M**	9.80	2.65	2.02	1.93	6.34	3.02	6.25	7.96
MEG3	**Imprinted**	**M**	0.03	0.16	0.04	0.15	1.64	0.01	0.19	25.28
MEST	**Imprinted**	**P**	220.15	102.98	59.87	4.47	1.14	4.49	1.16	0.27
NAP1L5	**Imprinted**	**P**	0.00	0.00	0.00	0.10	0.03	0.10	0.00	0.44
NNAT	**Imprinted**	**P**	0.00	0.10	0.00	0.27	0.05	0.18	0.05	2.37
PEG10	**Imprinted**	**P**	0.22	1.20	0.95	0.33	0.00	0.00	0.00	0.00
PEG3	**Imprinted**	**P**	0.00	0.02	0.28	0.58	0.01	0.00	0.00	0.00
PHLDA2	**Imprinted**	**M**	4.29	25.91	24.82	6.26	8.31	8.99	7.84	4.51
PLAGL1	**Imprinted**	**P**	49.28	20.63	12.39	0.54	0.51	1.17	0.70	0.01
RTL1	**Imprinted**	**P**	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.04
SGCE	**Imprinted**	**P**	0.71	5.20	6.15	0.55	0.01	0.07	0.09	0.12

Genes in shades have high levels of expression (RPKM>1) in at least one stage of development.

Table S3.3. Spreadsheet 2. Imprinting status and levels in human in vitro oocytes and embyros for genes that are imprinted in the bovine.

Gene Name	Status	Expressed allele	Oocyte	Pronuclear	Zygote	2-cell	4-cell	8-cell	Morula
ASCL2	**Not imprinted**	**Biallelic**	0.01	0.00	0.00	0.00	0.03	1.69	0.18
CDKN1C	**Imprinted**	**M**	0.42	1.21	0.31	0.22	0.11	7.99	0.00
DGAT1	**Not imprinted**	**Biallelic**	0.92	6.86	7.85	1.32	3.72	3.04	3.53
DIRAS3	**Imprinted**	**P**	2.49	15.72	4.68	11.21	1.77	5.60	0.03
GNAS	**Imprinted**	**Isoform dependent**	3.86	10.62	12.03	20.25	58.96	19.04	22.03
H19	**Imprinted**	**M**	ND	ND	ND	ND	ND	ND	ND
IGF2	**Imprinted**	**P**	0.01	0.24	0.00	0.05	0.09	2.21	0.03
IGF2R	**Imprinted**	**Biallelic**	3.09	1.91	1.88	1.57	3.02	3.64	0.38
MEG3	**Imprinted**	**M**	ND	ND	ND	ND	ND	ND	ND
MEST	**Imprinted**	**P**	439.80	199.68	294.27	110.77	88.72	24.19	75.57
NAP1L5	**Imprinted**	**P**	0.01	0.00	0.00	0.00	0.03	4.54	1.66
NNAT	**Imprinted**	**P**	ND	ND	ND	ND	ND	ND	ND
PEG10	**Imprinted**	**P**	3.57	0.48	0.67	0.52	3.02	0.13	0.12
PEG3	**Imprinted**	**P**	0.04	0.00	0.06	0.11	0.03	0.61	0.58
PHLDA2	**Imprinted**	**M**	0.01	0.00	0.00	0.00	0.03	17.38	2.53
PLAGL1	**Imprinted**	**P**	17.48	5.00	5.09	3.09	9.82	0.12	0.13
RTL1	**Imprinted**	**P**	ND	ND	ND	ND	ND	ND	ND
SGCE	**Imprinted**	**P**	7.57	1.49	1.06	2.65	0.81	0.07	1.35

*ND: Genes with RPKM < 0.1 are defined not detected (ND).
Genes in shades have high levels of expression (RPKM>1) in at least one stage of development.

Table S3.3. Spreadsheet 3. Imprinting status and levels in mouse in vivo oocytes and embyros for genes that are imprinted in the bovine.

Gene Name	Status	Expressed allele	Oocyte	Pronuclear	2-cell	4-cell	8-cell	Morula
ASCL2	**Imprinted**	**M**	0.30	0.03	0.13	0.83	0.00	0.00
CDKN1C	**Imprinted**	**M**	3.25	3.30	6.20	8.30	4.30	0.77
DGAT1	**Not imprinted**	**Biallelic**	0.75	1.70	2.13	4.43	2.87	24.27
DIRAS3	**Not imprinted**	**Biallelic**	ND	ND	ND	ND	ND	ND
GNAS	**Imprinted**	**M**	27.60	33.73	4.67	7.43	8.70	47.67
H19	**Imprinted**	**M**	ND	ND	ND	ND	ND	ND
IGF2	**Imprinted**	**P**	4.95	0.90	0.10	0.27	0.00	0.00
IGF2R	**Imprinted**	**M**	12.15	0.80	36.33	35.57	31.00	21.37
MEG3	**Imprinted**	**M**	2.15	0.23	0.07	0.00	0.00	0.20
MEST	**Imprinted**	**P**	210.65	127.77	34.90	5.43	2.20	0.47
NAP1L5	**Imprinted**	**P**	ND	ND	ND	ND	ND	ND
NNAT	**Imprinted**	**P**	ND	ND	ND	ND	ND	ND
PEG10	**Imprinted**	**P**	ND	ND	ND	ND	ND	ND
PEG3	**Imprinted**	**P**	1.55	1.10	0.27	0.23	0.10	0.30
PHLDA2	**Imprinted**	**M**	28.70	10.83	18.47	5.57	12.27	3.57
PLAGL1	**Imprinted**	**P**	2.70	1.20	3.20	0.00	0.00	0.03
RTL1	**Imprinted**	**P**	ND	ND	ND	ND	ND	ND
SGCE	**Imprinted**	**P**	0.00	0.00	0.07	3.43	4.37	2.67

*ND: Genes with RPKM < 0.1 are defined not detected (ND).
Genes in shades have high levels of expression (RPKM>1) in at least one stage of development.

Table S3.3. Spreadsheet 4. Imprinting status and levels in pig in vivo oocytes and embyros for genes that are imprinted in the bovine.

Gene Name	Status	Expressed allele	Oocyte	2-cell	4-cell	8-cell	morula	Blastocyst
ASCL2	**Not imprinted**	**Biallelic**	ND	ND	ND	ND	ND	ND
CDKN1C	**Not imprinted**	**Biallelic**	ND	ND	ND	ND	ND	ND
DGAT1	**Not imprinted**	**Biallelic**	36.70	2.44	38.59	3.95	18.77	7.33
DIRAS3	**Not imprinted**	**Biallelic**	0.02	0.07	0.26	140.46	1991.18	147.92
GNAS	**Not imprinted**	**Biallelic**	10.65	11.47	2.58	1.82	28.05	43.02
H19	**Imprinted**	**M**	ND	ND	ND	ND	ND	ND
IGF2	**Imprinted**	**P**	ND	ND	ND	ND	ND	ND
IGF2R	**Not imprinted**	**Biallelic**	0.75	0.46	0.89	0.05	1.95	3.24
MEG3	**Imprinted**	**M**	ND	ND	ND	ND	ND	ND
MEST	**Imprinted**	**P**	626.02	344.67	104.35	144.00	17.66	76.66
NAP1L5	**Imprinted**	**P**	0.00	0.10	0.34	0.33	0.33	0.64
NNAT	**Imprinted**	**P**	0.11	0.29	0.13	0.04	0.04	0.11
PEG10	**Imprinted**	**P**	ND	ND	ND	ND	ND	ND
PEG3	**Imprinted**	**P**	ND	ND	ND	ND	ND	ND
PHLDA2	**Imprinted**	**M**	ND	ND	ND	ND	ND	ND
PLAGL1	**Imprinted**	**P**	7.51	5.23	0.08	0.10	0.02	0.02
RTL1	**Not imprinted**	**Biallelic**	ND	ND	ND	ND	ND	ND
SGCE	**Imprinted**	**P**	16.19	86.98	4.47	52.31	0.08	0.11

*ND: Genes with RPKM < 0.1 are defined not detected (ND).

Genes in shades have high levels of expression (RPKM>1) in at least one stage of development.

Table S3.4. Spreadsheet 1. The eight confirmed bovine imprinted genes that were not detected in bovine in vivo pre-implantation embryos.

APEG3

KCNQ1OT1

MAGEL2

MIMT1

NESP55

SNRPN

TSSC4

USP29

Table S3.4. Spreadsheet 2. The expression levels in human oocytes and embryos for the genes that are undetected in bovine oocytes or embryos.

	Oocyte	**Pronuclear**	**Zygote**	**2-cell**	**4-cell**	**8-cell**	**Morula**
SNRPN	11.12	11.87	5.57	6.37	17.46	13.56	8.83
TSSC4	0.26	5.00	2.41	8.09	2.04	22.17	30.74
USP29	0.01	0.00	0.00	0.00	24.99	11.91	0.00

Table S3.4. Spreadsheet 3. Expression levels in mouse oocytes and embryos for the genes that are undetected in bovine oocytes or embryos.

	Oocyte	**Pronuclear**	**2-cell**	**4-cell**	**8-cell**	**Morula**
SNRPN	16.55	8.83	7.73	22.97	34.83	44.07
TSSC4	12.00	8.00	7.33	4.80	15.63	13.63
USP29	10.80	13.30	45.13	21.47	11.40	0.67
APEG3	3.75	5.57	1.63	2.80	0.00	1.97

3.7. References

1 Ferguson-Smith, A. C. Genomic imprinting: the emergence of an epigenetic paradigm. *Nature reviews. Genetics* **12**, 565-575, doi:10.1038/nrg3032 (2011).

2 Lawson, H. A., Cheverud, J. M. & Wolf, J. B. Genomic imprinting and parent-of-origin effects on complex traits. *Nature reviews. Genetics* **14**, 609-617, doi:10.1038/nrg3543 (2013).

3 Tian, X. C. Genomic imprinting in farm animals. *Annual review of animal biosciences* **2**, 23-40, doi:10.1146/annurev-animal-022513-114144 (2014).

4 Barlow, D. P. & Bartolomei, M. S. Genomic imprinting in mammals. *Cold Spring Harbor perspectives in biology* **6**, doi:10.1101/cshperspect.a018382 (2014).

5 Kim, K. P. *et al.* Gene-specific vulnerability to imprinting variability in human embryonic stem cell lines. *Genome research* **17**, 1731-1742, doi:10.1101/gr.6609207 (2007).

6 Cruz, N. T. *et al.* Putative imprinted gene expression in uniparental bovine embryo models. *Reproduction, fertility, and development* **20**, 589-597 (2008).

7 Thurston, A., Taylor, J., Gardner, J., Sinclair, K. D. & Young, L. E. Monoallelic expression of nine imprinted genes in the sheep embryo occurs after the blastocyst stage. *Reproduction* **135**, 29-40, doi:10.1530/REP-07-0211 (2008).

8 Tveden-Nyborg, P. Y. *et al.* Analysis of the expression of putatively imprinted genes in bovine peri-implantation embryos. *Theriogenology* **70**, 1119-1128, doi:10.1016/j.theriogenology.2008.06.033 (2008).

9 Suzuki, J., Jr. *et al.* In vitro culture and somatic cell nuclear transfer affect imprinting of SNRPN gene in pre- and post-implantation stages of development in cattle. *BMC developmental biology* **9**, 9, doi:10.1186/1471-213X-9-9 (2009).

10 Park, C. H. *et al.* Analysis of imprinted gene expression in normal fertilized and uniparental preimplantation porcine embryos. *PloS one* **6**, e22216, doi:10.1371/journal.pone.0022216 (2011).

11 Li, Y. & Sasaki, H. Genomic imprinting in mammals: its life cycle, molecular mechanisms and reprogramming. *Cell research* **21**, 466-473, doi:10.1038/cr.2011.15 (2011).

12 Bartolomei, M. S. & Ferguson-Smith, A. C. Mammalian genomic imprinting. *Cold Spring Harbor perspectives in biology* **3**, doi:10.1101/cshperspect.a002592 (2011).

13 Park, J. S., Jeong, Y. S., Shin, S. T., Lee, K. K. & Kang, Y. K. Dynamic DNA methylation reprogramming: active demethylation and immediate remethylation in the male pronucleus of bovine zygotes. *Developmental dynamics : an official publication of the American Association of Anatomists* **236**, 2523-2533, doi:10.1002/dvdy.21278 (2007).

14 Reik, W. & Walter, J. Genomic imprinting: parental influence on the genome. *Nature reviews. Genetics* **2**, 21-32, doi:10.1038/35047554 (2001).

15 Mann, M. R. *et al.* Selective loss of imprinting in the placenta following preimplantation development in culture. *Development* **131**, 3727-3735, doi:10.1242/dev.01241 (2004).

16 Denomme, M. M. & Mann, M. R. Genomic imprints as a model for the analysis of epigenetic stability during assisted reproductive technologies. *Reproduction* **144**, 393-409, doi:10.1530/REP-12-0237 (2012).

17 Young, L. E., Sinclair, K. D. & Wilmut, I. Large offspring syndrome in cattle and sheep. *Reviews of reproduction* **3**, 155-163 (1998).
18 Maher, E. R. *et al.* Beckwith-Wiedemann syndrome and assisted reproduction technology (ART). *Journal of medical genetics* **40**, 62-64 (2003).
19 Hiendleder, S. *et al.* Tissue-specific effects of in vitro fertilization procedures on genomic cytosine methylation levels in overgrown and normal sized bovine fetuses. *Biology of reproduction* **75**, 17-23, doi:10.1095/biolreprod.105.043919 (2006).
20 Sutcliffe, A. G. *et al.* Assisted reproductive therapies and imprinting disorders--a preliminary British survey. *Human reproduction* **21**, 1009-1011, doi:10.1093/humrep/dei405 (2006).
21 Jiang, Z. *et al.* Transcriptional profiles of bovine in vivo pre-implantation development. *BMC genomics* **15**, 756, doi:10.1186/1471-2164-15-756 (2014).
22 Park, S. W. *et al.* Transcriptional Profiles of Imprinted Genes in Human Embryonic Stem Cells During In vitro Differentiation. *International journal of stem cells* **7**, 108-117, doi:10.15283/ijsc.2014.7.2.108 (2014).
23 Cao, S. *et al.* Specific gene-regulation networks during the pre-implantation development of the pig embryo as revealed by deep sequencing. *BMC genomics* **15**, 4, doi:10.1186/1471-2164-15-4 (2014).
24 Graf, A. *et al.* Fine mapping of genome activation in bovine embryos by RNA sequencing. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 4139-4144, doi:10.1073/pnas.1321569111 (2014).
25 Xue, Z. *et al.* Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* **500**, 593-597, doi:10.1038/nature12364 (2013).
26 Yan, L. *et al.* Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nature structural & molecular biology* **20**, 1131-1139, doi:10.1038/nsmb.2660 (2013).
27 Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods* **5**, 621-628, doi:10.1038/nmeth.1226 (2008).
28 Ruddock, N. T. *et al.* Analysis of imprinted messenger RNA expression during bovine preimplantation development. *Biology of reproduction* **70**, 1131-1135, doi:10.1095/biolreprod.103.022236 (2004).
29 Cheong, C. Y. *et al.* Germline and somatic imprinting in the nonhuman primate highlights species differences in oocyte methylation. *Genome research* **25**, 611-623, doi:10.1101/gr.183301.114 (2015).
30 Driver, A. M., Huang, W., Kropp, J., Penagaricano, F. & Khatib, H. Knockdown of CDKN1C (p57(kip2)) and PHLDA2 results in developmental changes in bovine pre-implantation embryos. *PloS one* **8**, e69490, doi:10.1371/journal.pone.0069490 (2013).
31 Weinstein, L. S., Xie, T., Qasem, A., Wang, J. & Chen, M. The role of GNAS and other imprinted genes in the development of obesity. *Int J Obes (Lond)* **34**, 6-17, doi:10.1038/ijo.2009.222 (2010).
32 He, S. *et al.* Discovery of a Potent and Selective DGAT1 Inhibitor with a Piperidinyl-oxy-cyclohexanecarboxylic Acid Moiety. *ACS Med Chem Lett* **5**, 1082-1087, doi:10.1021/ml5003426 (2014).
33 Hoyo, C. *et al.* Erythrocyte folate concentrations, CpG methylation at genomically imprinted domains, and birth weight in a multiethnic newborn cohort. *Epigenetics :*

official journal of the DNA Methylation Society **9**, 1120-1130, doi:10.4161/epi.29332 (2014).

34 Chen, Z. *et al.* Characterization of global loss of imprinting in fetal overgrowth syndrome induced by assisted reproduction. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 4618-4623, doi:10.1073/pnas.1422088112 (2015).

Chapter Four

Effects of High Hydrostatic Pressure on Expression Profiles of In Vitro Produced Vitrified Bovine Blastocysts

(Jiang Z, Harrington P, Zhang M, Marjani S, Park J, Kuo L, Pribenszky C & Tian XC)

4.1. Abstract

High Hydrostatic Pressure (HHP) has been used to enhance stress tolerance and to promote survival of embryos before they are subjected to insults such as cryopreservation. However, the molecular mechanisms of the beneficial effects of HHP are poorly understood. Here we found re-expansion rates of bovine blastocysts after vitrification-warming were significantly ($P < 0.05$) higher in embryos treated with 40 MPa or 60 MPa than controls, demonstrating that the application of HHP promotes the in vitro developmental competence of vitrified bovine embryos. However, 80 MPa resulted in significantly reduced re-expansion rates, suggesting that this pressure started to be lethal to bovine blastocysts. Microarray analysis revealed a total of 399 differentially expressed transcripts, representing 254 unique genes, among different treatment groups. Gene ontology analysis revealed that HHP at 40 MPa and 60 MPa promoted embryo competence through down-regulation of genes involved in cell death and apoptosis, and up-regulation of RNA processing, cellular growth and proliferation. In contrast 80 MPa up-regulated expression of genes for apoptosis, and down-regulated genes for protein folding and cell cycle. These may explain why 80 MPa-treated embryos stopped developing. Moreover, gene expression was also changed by the length of the recovery time after HHP. The significantly over-represented groups are apoptosis and cell death in the 1h group, and protein folding, response to unfolded protein and cell cycle in the 2h group. Taken together, these data suggest that HHP induces specific responses in vitrified bovine blastocysts and promotes their developmental competence through modest transcriptional reprogramming.

Keywords: bovine**,** vitrified blastocysts, HHP, recovery time, microarray, transcriptional profiles

4.2. Introduction

In vitro maturation, culture and cryopreservation of gametes and embryos require meticulously adjusted conditions to avoid or minimize detrimental stress of osmotic, oxidative, cold/heat shock, nutritional and mechanic nature. Eustress can improve protein conformation, maintain cellular homeostasis as well as stabilize membrane structures, while distress will cause apoptotic cell death [1-4]. Recent studies indicate that stress from a well-defined and properly applied sub-lethal high hydrostatic pressure (HHP) may induce general adaptation and increase tolerance to various in vitro procedures [5]. For example, HHP treatment has been shown to improve the survival rates, fertilizing ability and development competence of cryopreserved oocytes [6,7], sperm [8,9], embryos [10-12] and embryonic stem cells [13]. Applying a HHP of 60 MPa for 1 hour (h) could increase the ICM cell numbers of bovine blastocysts [12]. Furthermore, re-expansion and hatching rates after vitrification-warming were also found to be affected by the duration of the recovery time, the period between the termination of HHP treatment and the initiation of vitrification. HHP followed by 1h of recovery proved to be superior with regards to both re-expansion and hatching rates [11,14]. Upon further fine-tuning, this invention can be applied to improve in vitro embryo biotechnologies in various species including humans.

A transcriptional effect is believed to be involved in the elevated stress tolerance induced by HHP. Previous studies on mouse embryos evaluated candidate genes from pressure-related functional groups. The expression of the antizyme inhibitor 1 (*AZIN1*), the mitochondrial superoxide dismutase 2 (*SOD2*) and the gamma growth arrest and DNA-damage-inducible (*GADD45G*) are found to be significantly up-regulated by HHP treatment [10]. Further investigations revealed that pressure can change protein structures and enhance the production of heat shock proteins (HSP) [15,16] such as HSP70 [12]. In bovine blastocysts, several candidate stress

genes have also been examined. These include stress related genes *SOD2,* glutathione peroxidase 4 (*GPX4*) and heat shock 70kDa protein 1A (*HSPA1A*). None of the evaluated genes, however, were significantly changed by the 60 MPa HHP treatment [11]. Therefore, the molecular mechanisms for the beneficial effects of HHP are yet to be ascertained.

Without many known definitive candidate genes, genome-wide expression profiling by DNA microarray is highly effective in throughput examinations of transcriptomes. Accordingly, the aim of the present study is to evaluate the effects of HHP treatments at three different levels with different recovery times on the transcriptomes of bovine in vitro produced (IVP) vitrified blastocysts. To our knowledge, this is the first transcriptional profiling of bovine blastocysts treated by HHP. Pathways such as apoptosis, protein folding, cell cycle regulation, RNA processing and translation were found to be affected by HHP.

4.3. Materials and Methods

4.3.1. Microarray Design and Annotation

The Cattle Array-Ready oligonucleotides were designed at the University of Illinois at Urbana-Champaign and described in detail by Everts et al [17]. The microarray contained 13,254 70-mer oligonucleotide probes that were synthesized at Illumina (www.illumina.com, San Diego, CA, USA). All probes were printed in duplicates on glass slides at Microarrays Inc. (Nashville, TN, USA). In total, these oligonucleotide probes represent 10,991 unique genes.

4.3.2. Collection of IVP blastocysts

In vitro bovine blastocysts were produced as described previously [11,18]. Briefly, in vitro fertilization (Day 0) was performed using abattoir bovine oocytes and embryos were immediately placed in CR1aa medium supplemented with BSA for Days 1 and 2 of culture. Cleaved embryos were transferred to CR1aa + 10 % FBS and cultured at 38.5 °C in 5 % CO_2 in humidified air until Day 7/expanded blastocyst stage. Embryos were examined and staged under light microscopy and only morphologically intact embryos meeting the standards of Grade 1 by the International Embryo Transfer Society (IETS) were used in the following experiments.

4.3.3. High Hydrostatic Pressure Treatment of Bovine IVP Blastocysts

Blastocysts were randomly distributed into the control and HHP treated groups as shown in Fig. 4.1. To apply HHP, groups of embryos were transferred to 0.25 ml straws in embryo holding medium (TCM-199; Gibco, Grand Island, NY, USA) without air bubbles. Straws were sealed by plastic plugs and were then placed into a pre-warmed stainless steel pressure machine (HHP machine 100; Cryo-Innovation, Budapest, Hungary) filled with water as the pressure medium. The following treatments were included (Fig. 4.1): (1) Control embryos were left

unaffected in the incubator (one atmospheric pressure or 0.1 MPa); (2) treatment groups were assigned to 40, 60 and 80 MPa HHP for 1h at either 25 °C (room temperature) or 37 °C (body temperature), followed by three different recovery time periods (0, 1 and 2h) post-HHP in the holding medium before the embryos were vitrified using the Solid Surface Vitrification (SSV) method [19]. The cryopreserved blastocysts were then thawed by immersing the straws into 0.5 M sucrose solution for 5 min at 39 °C, after which the blastocysts were transferred into TCM-199 medium and cultured in an incubator at 39 °C, 5% CO2 and humidified air. The re-expansion rates of embryos given 2h of recovery were evaluated for morphological survival 24h after warming. Then pools of 5 embryos from each treatment were washed twice in D-PBS and stored in RNAlater (Ambion, Grand Island, NY, USA) in liquid nitrogen. All treatments were repeated three times (n = 3).

The re-expansion rates were analyzed using One Way ANOVA with Tukey's multiple comparisons, or the Student's t-test. Figures were presented as mean ± standard deviation (SD). A P-value < 0.05 (*) or < 0.01 (**) was considered statistically significant.

4.3.4. RNA Isolation, Linear Amplification, Labeling and Microarray Hybridization

Following the reproducible procedures of RNA extraction and linear amplification from our previous study [18], we isolated total RNA from each pool of 5 blastocysts using TRIzol reagent (Invitrogen, Grand Island, NY) and precipitated the RNA with linear acrylamide (Ambion, Grand Island, NY). The quality of the total RNA was examined with the Aglient RNA 6000 Pico kit (Aglient Technologies, Santa Clara, CA) using the Aglient Bioanalyzer 2100. RNA was then amplified twice using the TargetAmp 2-round aminoallyl-aRNA amplification kit 1.0 (Epicentre, Madison, WI) according to the manufacturer's instructions. From 5 blastocysts,

we were able to generate an average of 60 µg of amplified RNA after two rounds of amplification. Amplified RNA was stored at -80 °C until utilization on the microarray.

The reference microarray design, in which the embryonic expression profiles were compared to a standard reference RNA, was used. The reference RNA was isolated from brain, kidney, liver and lung tissues of a naturally reproduced heifer and pooled at an equal proportion. More than 90 % of the genes on the microarray were lit up by the standard reference. Two microgram of amplified RNA from each sample and the reference were reverse transcribed, labeled, and hybridized to each microarray as previously described for single embryos [18]. In total, 144 microarrays were used including dye-swap hybridizations.

4.3.5. Microarray Data Analysis

The microarrays were scanned with GenePix 400B (Molecular Devices, Union City, CA, USA) and normalization of fluorescence intensities was accomplished by using the GenePix Pro 6.0 scanning software (Axon Instruments, Union City, CA, USA). Each scanned image was examined thoroughly and dust particles and spots with high background were flagged and removed from analysis. The background and standard deviation were calculated for each raw data file after scanning, and only those spots with intensities three standard deviations above background were considered "expressed" and loaded into Genespring 12.1 (Agilent Technologies Palo Alto, CA, USA). Loess normalization was applied to all microarrays before statistical comparisons. In the analysis, each probe was considered individually.

In the evaluations of post-normalization of the probes in the microarrays, 12,274 transcripts present in either the standard reference or sample on 90% of the microarrays underwent further analysis. A SAS model with covariates being pressure, recovery time and temperature were considered. We looked at two metrics, firstly the number of genes for which

we find that factor significant using a significance level of 0.01 and secondly the sum of the P-values for all probes. Based on these two measures, we ranked HHP as the most significant factor, followed by recovery time, and lastly temperature, which confirmed our observation that no significant differences were found between 25 °C and 37 °C on the re-expansion rates by HHP. We therefore combined the data from the two temperatures to increase the statistical power. The SAS model was then evaluated to identify the genes differentially expressed between treated and control blastocysts. Genes were deemed differentially expressed if they showed a Benjamini Hochberg procedure controlling for false discovery rate (FDR) at P-value < 0.05. Hierarchical clusters were generated by using Genespring GX 12.1 with the K-means clustering algorithm. Heatmaps and Venn diagrams of the summary of differentially expressed genes were developed with R.

4.3.6. Gene Ontology Analysis

Functional annotation enrichment analysis for Gene Ontology (GO) was conducted using DAVID [20]. GO terms shown in this study summarized all similar sub-terms into an overarching term, and Benjamani-Hochberg adjusted P-values are shown for the representative term.

4.3.7. Quantitative Real Time-Reverse Transcription Polymerase Chain Reaction (qRT-PCR) Analysis

qRT-PCR was performed to validate differential expression of eight selected genes using amplified RNA. Amplified RNA was reverse transcribed to cDNA by SuperScript III Reverse Transcriptase (Invitrogen) and amplified with specific primers designed by using Primer 3.0 (Table S4.1). The qRT-PCR was performed using SYBR Green PCR Master Mix (ABI) and an ABI 7500 Fast instrument. The data were analyzed using the 7500 software version 2.0.2 provided with the instrument. All values were normalized to the internal control, *β-ACTIN*. The

efficiency of each primer pair was calculated over a 3.5 log dilution range and the relative gene expression values were calculated using the $2^{-\triangle\triangle Ct}$ method. Expression levels that were relative to those in the standard reference were calculated and the mean for each group was determined and compared for an overall fold change. Data from qRT-PCR were analyzed as described above for re-expansion rates.

4.4. Results

4.4.1. The Effect of HHP on Re-expansion Rates of Cryopreserved Bovine IVP Blastocysts

Blastocyst re-expansion is a predictive value for implantation of frozen-thawed blastocysts. We first investigated the post-thaw survival of vitrified IVP bovine blastocysts. Because comparisons of 0 and 1h recovery time had been studied previously [11], only 2h recovery was included here. Re-expansion rates were significantly affected by HHP and recovery but not by temperature. Therefore data for the two temperatures were combined. Application of appropriate HHP (40 or 60 MPa) followed by 2h of recovery proved to be superior with regard to re-expansion. Specifically, re-expansion rates were significantly ($P < 0.05$) higher in the 40-MPa (90 %) and 60-MPa (87 %) groups than the controls (63.5 %) (Fig. 4.2). These data demonstrated that application of HHP stress promoted the in vitro developmental competence of vitrified bovine embryos. However, HHP treatment at 80 MPa resulted in significantly reduced re-expansion rates (43.5 %) compared to controls (63.5 %; Fig. 4.2), suggesting that 80 MPa is starting to be lethal to bovine blastocysts.

4.4.2. Hierarchical Clustering of Expression Profiles of Pressure-Treated and Cryopreserved Bovine IVP blastocysts

Building upon the previous notion and our observation that vitrified embryos re-expanded better when appropriate HHP treatment is combined with a short recovery period, a comprehensive genome-wide transcriptomic investigation was conducted. The "ranking" of the most significant factors was pressure, then recovery time and lastly temperature, which is consistent with our observation that re-expansion rates were not significantly affected by temperature. Expression data were also combined for the two temperatures. Hierarchical

clustering of all treatments based on 12,274 analyzed transcripts clearly showed a tendency for different expression profiles (Fig. 4.3A). Specifically, the 40- and 60-MPa groups were separated from the 80-MPa and control groups (Fig. 4.3A). Within each HHP level, 1h and 2h clustered together and separated from 0h with the exception in 80-MPa group (Fig. 4.3A). This overall clustering pattern in gene expression is consistent with the re-expansion results.

A total of 399 transcripts (254 unique genes) were identified as differentially expressed among different treatment groups ($P < 0.05$) (Table S4.2). The hierarchical clustering of all differentials (Fig. 4.3B) was again based on pressure, further suggesting that pressure played a significant role in gene expression changes.

4.4.3. Effects of HHP on Gene Expression

Among the 399 total differentially expressed transcripts, 340 were caused by HHP while 59 were not related to pressure changes. Of the 340, 83 and 182, 84 and 44, were down- and up-regulated in cryopreserved embryos treated with 40 MPa or 60 Mpa HHP, respectively (Fig. 4.4A, Table S4.3). The common down-regulated transcripts in both the 40 MPa and 60 MPa HHP-treated embryos were involved in the cell death and apoptosis (Table 4.1), among these were heat shock 22kDa protein 8 (*HSPB8*), death inducer-obliterator 1 (*DIDO1*), meuroepithelial cell-transforming gene 1 (*NET1*), coagulation factor III (*F3*) and caspase 7 (*CASP7*). Interestingly, CASP7, a protein of the caspase family and considered to be an important executioner protein of apoptosis, and HSPB8, a common heat shock protein involved in regulation of cell proliferation and apoptosis, were both down-regulated upon HHP treatment (Fig. 4.5). Moreover, DIDO1 and NET1, both activated early in apoptosis (pro-apoptotic) were also down-regulated by HHP (Fig. 4.5). The expression changes of these genes upon 40 and 60 MPa HHP are in agreement with the higher survival rates of these groups. Conversely, the up-

regulated transcripts, such as serine/arginine-rich splicing factor 7 (*SFRS7*), *SFRS9,* DNA-directed RNA polymerase II subunit G (*POLR2G*), *POLR2F, POLR2L,* small nuclear ribonucleoprotein D3 (*SNRPD3*), *SNRPD2,* eukaryotic translation initiation factor 4B (*EIF4B*), ribosomal protein L38 (*RPL38*) and mitochondrial ribosomal protein L43 (*MRPL43*) (Table 4.1), are involved in RNA processing and transcription, as well as regulations of protein synthesis, and likely promoted embryo survival. Collectively, these results suggested that stress caused by elevated HHP induced embryos to degrade apoptotic transcripts and increase RNA transcription and translation. These effects, while not necessarily specific for pressure resistance or cryo-tolerance, prepared the embryos to resist insults and survive better. However, when pressure was further increased to 80 MPa, 25 and 92 transcripts were down- and up-regulated, respectively (Fig. 4.4A, Table S4.3). The biological processes significantly represented among down-regulated transcripts included protein folding and cell cycle, which including BAG family molecular chaperone regulator 4 (*BAG4*), DBF4 zinc finger A (*DBF4*), mitogen-activated protein kinase phosphatase 1 (*DUSP1*), and serine/threonine-protein kinase SNK (*PLK2;* Table 4.1), while the up-regulated transcripts significantly over-represented cell death, apoptosis, and chromatin assembly/disassembly, such as *EIF4B,* hexokinase 1 (*HK1*), histone cluster 1, H1e (*HIST1H1E*) and histone deacetylase 8 (*HDAC8*; Table 4.1). Some of these changes are opposite to those seen at 40 or 60 MPa. These results suggested that over exerting HHP would disturb cell structure and proliferation and therefore be detrimental to bovine embryos.

Among genes up- and down-regulated at each pressure level, 136, 21 and 39 transcripts were uniquely differentially expressed in the 40-, 60- or 80-MPa treated groups compared to the controls, respectively (Fig. 4.4B, Table S4.4). Additionally, 66 transcripts uniquely overlapped in the comparisons between the 40- or 60-MPa treated embryos and controls (Fig. 4.4B, Table

S4.5). These include down-regulated transcripts, *NET1,* which is involved in cell death and apoptosis, and up-regulated transcripts, *SFRS7, SFRS9, POLR2G, POLR2F, POLR2L, SNRPD2* and amyloid beta (A4) precursor-like protein 1 (*APLP1*), which are involved in RNA processing and translation. Twenty-six transcripts were differentially expressed in all pressure treated groups compared to un-pressured controls. These include *DIDO1, HSPH1, HSPB8, HK1* and *EIF4B* (Fig. 4.4B, Table S4.6) and may represent essential genes responding to pressure stress.

4.4.4. Effects of Different Recovery Time on Gene Expression

The duration of the recovery period is particularly interesting because HHP induces gene changes in cellular metabolism and functions and time is needed for the synthesis of special RNA and proteins. In a previous study, allowing bovine embryos to recover for 1h after HHP was found to further increase embryo survival compared to 2h or HHP alone without recovery [11]. Here, we compared gene expression at these three time points. Among the total of 399 differentially expressed transcripts, 167 were caused by recovery. We identified 49 and 98 down-regulated, and 16 and 33 up-regulated transcripts at 1h and 2h compared to controls (0h), respectively (Fig. 4.6A, Table S4.7). Gene ontology analysis of the down-regulated transcripts revealed apoptosis, proteolysis and phosphate metabolic process in the 1h group, and protein folding, cell cycle and cell death in the 2h group as significantly overrepresented (Table 4.2). Up-regulated transcripts were involved in cellular growth and proliferation, cell morphology, and cellular function and maintenance in the 1h group, cellular growth and proliferation, DNA replication and G1/S transition of mitotic cell cycle in the 2h group (Table 4.2). Of special interest was the dramatically higher number of differentially expressed transcripts in the 0 vs. 2h comparison than that of the 0 vs. 1h comparison, suggesting that 2h of recovery allowed more changes of gene expression to occur. Among the differentials unique to the 0 vs. 1h and 0 vs. 2h

comparisons, 20 and 41 transcripts were revealed, respectively (Fig. 4.6B, Table S4.8). A total of 42 transcripts were common differentials in both the 0 vs. 1h and 0 vs. 2h comparisons (Fig. 4.6B, Table S4.9), including *CASP7, DUSP1* and *F3,* which are involved in apoptosis and cell death.

It is noteworthy that although a 2h recovery induced more gene expression changes, it did not promote better embryo re-expansion than 1h. The additional changes in gene expression during the second hour of recovery may have corrected changes already taken place during the first hour, thereby canceling some of the changes needed to resist insults from the subsequent cryopreservation.

4.4.5. Confirmation of Microarray Data by Real Time qRT-PCR

To confirm the throughput results of microarray, we performed qRT-PCR on eight genes, namely *CASP7, NET1, APLP1, EIF4B, HSPH1, HSPB8, DIDO1* and *F3*, which were significantly affected by HHP treatments and play crucial roles in cell death and apoptosis. In nearly all cases, the qRT-PCR detected bigger fold changes and substantiated results from microarray (Fig. 4.7A, B).

4.5. Discussion

In the bovine embryo transfer industry, vitrification is the most common method to cryopreserve IVP embryos [21]. Sublethal HHP was reported to enhance stress tolerance and increase post-thaw survival of sperm, embryos or stem cells after cryopreservation in murine, porcine and bovine [10]. Different HHP conditions (pressure level, recovery time and temperature) have been explored on gametes and embryos subjected to various assisted reproductive technology procedures [22]. HHP at 60 MPa for 1h have been shown to increase the in vitro development of bovine blastocysts [11,12]. In the present study, we extended previous findings by testing 40 and 80 MPa and found that 80 MPa is not well-tolerated. Additionally we also found that HHP at 40 prior to vitrification gave a higher re-expansion rate than the previously tested 60 MPa and non-treated group. The results in the bovine are also in accordance with previous reports that in the mouse and ovine [13].

Despite advances in morphological studies, limited information on the molecular mechanisms of the positive effect of HHP is available. Without clear candidates, comparing the entire transcriptomes of the treated and controls was the approach of choice. Interestingly, changes of transcriptomes were reflective of the re-expansion data that the best treatment condition, 40 MPa, elicited the most changes in gene expression compared to controls. Most of the down-regulated genes in the beneficial levels of HHP, 40 and 60 MPa, belonged to cell death and apoptosis, up-regulated genes were involved in RNA processing, cellular growth and proliferation, while some of these changed in the opposite direction by the harmful level of HHP at 80 MPa. The majority of the genes reported here are newly identified for HHP-treated, vitrified embryos. Collectively, it appears that while the embryos responded to HHP stress by changing gene expression, these events prepared them for the upcoming insult of vitrification.

However, when too much stress was given (80 MPa for bovine blastocysts), embryo lethality occurred. By the candidate approach it was previously reported that sublethal stress such as heat, affects the embryos through apoptosis [2,3,23] by influencing the expression of development- [24] and stress-related genes [25,26]. In this study, the same pathways were revealed for HHP stress. It appears that cells have limited pathways to resist stress and use the same mechanisms for different external insults.

Because a short recovery period after HHP was shown to be beneficial for cell/embryo survival [6,11,27], we also determined genes that were affected by different lengths of recovery. Similarly to pressure, genes involved in regulation of cell death, apoptosis and protein folding were down-regulated, while up-regulated genes belonged to cell morphology, DNA replication and cellular growth and proliferation after either a 1h or 2h recovery. HHP and recovery seem to affect the same developmental pathways because alleviating cell death is essential for embryonic development.

In addition to the well-known apoptotic events induced by stress, we also identified many new pathways involved in the protection mechanisms. These include RNA processing, translation, cell cycle, oxidative phosphorylation and cellular growth and proliferation. A closer look at the gene lists revealed members of the above-mentioned pathways such as *CASP7, DIDO1, NET1, HSPH1,* and *HSPB8,* which may be responsible for the possible protection mechanisms induced by HHP. Particularly, *CASP7*, a member of caspase family, *HSPH1* and *HSPB8*, members from the heat shock gene family, *DIDO1* and *NET1* were down-regulated by HHP (40 and 60 MPa). CASP7 has been shown to be an important executioner protein of apoptosis [28,29]. *HSPH1* has been shown to prevent the aggregation of denatured proteins in cells under severe stress [30,31]. *HSPB8* belongs to the superfamily of small heat-shock proteins and

involved in regulation of cell proliferation and apoptosis through the activation of transforming growth factor- β activated kinase 1 (*TAK1*) [32]. In addition, *DIDO1* and *NET1* were activated early in apoptosis through regulation of BCL-2 [33,34]. Sequential activation of these essential genes to respond to pressure stress plays a central role in improving the stress tolerance of vitrified bovine embryos. Furthermore, *EIF4B* was up-regulated in bovine blastocysts by HHP. *EIF4B* is required for cell proliferation and survival through regulation of protein synthesis [35-37]. Its up-regulation suggests that 40- and 60-MPa requires the formation of new proteins to promote embryo survival. However, the well-known *HSP70* (*HSPA1A* in our microarray) which was induced by sublethal pressures in microorganism [38] was not significantly regulated by HHP in our study, an observation consistent with a previous report employing the candidate approach [11]. It is possible that, instead of the *HSP70,* the bovine embryos activate alternative proteins such as *HSPH1* and *HSPB8*, to respond to pressure stress.

This is the first throughput study on HHP-treated and cryopreserved blastocysts. It should be noted that our microarray represents 10,991 genes, about half of the bovine expressed genome [39]. Of the approximately 12,000 genes expressed by the bovine blastocysts, 6,086 were represented in the microarray used. Therefore, half of the blastocyst's transcriptomes is not studied here. However, we were able to identify multiple HHP-induced pathways and some were common to those cells used to resist other stresses such as heat. These data suggest that the cells are limited to only a number of pathways to counteract external stress and it is therefore likely that this throughput study identified all pathways, albeit not all genes, involved in HHP stress resistance. Application of the more powerful RNA-seq technology may help to identify additional differentially expressed genes, but it is unlikely that pathways additional to those found here will be revealed.

Previous studies mainly focused on the development of embryos from HHP-treated oocytes as well as gene expression of treated oocytes [5-7,10-12,27]. These studies together with ours revealed that only a moderate number of genes was changed by HHP. Changes at the protein level such as folding, post-translation modifications [15] and protein levels likely represent major responses induced by HHP. To date, proteomic analysis has only been applied to HHP-treated microorganisms [40]. Such study on embryos, however, is currently unfeasible due to small sample sizes and high expenses.

In summary, our results showed 1) bovine cryopreserved embryos exhibit higher developmental competance after treatment of HHP at 40 or 60 MPa, 80 MPa, however, is not well-tolerated; 2) HHP treatments induced a modest transcriptional reprogramming in bovine embryos; and 3) HHP affected expression of genes involved in cell death and survival, RNA processing, as well as cell cycle and cell proliferation.

Table 4.1. Representative functional groups of differentially expressed genes affected by HHP treatments.

Treatment	Change	GO terms	P-value	Genes
0 vs. 40	Down-regulated	Cell death	1.27E-04	*DOCK1, HSPB8, F3, CASP7, RFFL, DIDO1, TOP2A, NET1*
		Apoptosis	2.55E-04	*DOCK1, F3, CASP7, RFFL, DIDO1, TOP2A, NET1*
		Transition metal ion transport	3.04E-02	*ATP2C1, SLC39A9*
	Up-regulated	RNA processing	1.16E-06	*POLR2G, POLR2F, POLR2L, SNRPD2, SART3, SF3B5, APLP1, SMNDC1, SFRS7, RBPMS, PPP1R8, SFRS9, U2AF1, DDX51*
		Translation	2.82E-02	*EIF4B, MRPL10, RPL15, RPL36, RPL38, MRPL43*
		Oxidative phosphorylation	1.42E-02	*NDUFA3, NDUFB10, NDUFB8, NDUFC2*
0 vs. 60	Down-regulated	Cell Death	5.47E-05	*BAG4, CASP7, DUSP6, F3, GADD45B, HK1, HSPB8, IGFBP3, DIDO1*
		Apoptosis	4.35E-02	*BAG4, GADD45B, DIDO1, NET1*
		Transition metal ion transport	3.04E-02	*ATP2C1, SLC39A9*
	Up-regulated	RNA processing	7.45E-06	*SFRS7, POLR2G, POLR2F, PRPF4B, POLR2L, SNRPD3, SFRS9, TFB2M, SNRPD2, APLP1*
		Transcription, DNA-dependent	6.58E-03	*POLR2G, BGLAP, POLR2F, POLR2L, TFB2M*
		Translation	4.34E-02	*EIF4B, RPL36, RPL38, MRPL43*
0 vs. 80	Down-regulated	Protein folding	1.53E-02	*BAG4, DNAJC10, DNAJA1, DNAJB6*
		Response to unfolded protein	1.86E-02	*HSPH1, DNAJA1, DNAJB6*
		Cell cycle	2.53E-02	*PLK2, DUSP1, DBF4, NSL1, ANXA1, BANP, RAD54B*
	Up-regulated	Chromatin assembly or disassembly	1.38E-05	*HIST1H1E, HDAC8*
		Apoptosis and Cell Death	3.38E-05	*AP2A2, EIF4B, HK1*
		Cellular Assembly and Organization	1.17E-02	*HDAC8*

Table 4.2. Representative functional groups of differentially expressed genes affected by HHP recovery time.

Treatment	Change	GO terms	P-value	Genes
0 vs. 1h	Down-regulated	Apoptosis	1.27E-02	*F3, CASP7, MDM4, GADD45B, DIDO1*
		Proteolysis	2.02E-02	*WSB1, UHRF2, F3, CASP7, YOD1, CNOT4*
		Phosphate metabolic process	5.97E-02	*PRPF4B, DUSP1, GADD45B, IGFBP3, DUSP6*
	Up-regulated	Cellular growth and proliferation	2.64E-03	*DBF4*
		Cell Morphology	1.74E-03	*CNTNAP1*
		Cellular Function and Maintenance	1.39E-02	*NDUFA3*
0 vs. 2h	Down-regulated	Protein folding	1.53E-02	*BAG4, DNAJC10, DNAJA1, DNAJB6, HSPH1*
		Cell cycle	2.53E-02	*FAM83D, UHRF2, PLK2, DUSP1, BANP, BIRC5, CABLES1*
		Cell death	5.92E-02	*BAG4, HSPB8, F3, CASP7, BIRC5, APLP1*
	Up-regulated	Cellular Growth and Proliferation	6.24E-05	*APLN*
		DNA replication	1.18E-02	*DBF4, WRNIP1, TOP2A*
		G1/S transition of mitotic cell cycle	4.86E-02	*BCAT1, DBF4*

Figure 4.1. Experiment design of high hydrostatic pressure (HHP) treatment on cryopreserved bovine IVP blastocysts.

Fig. 4.1

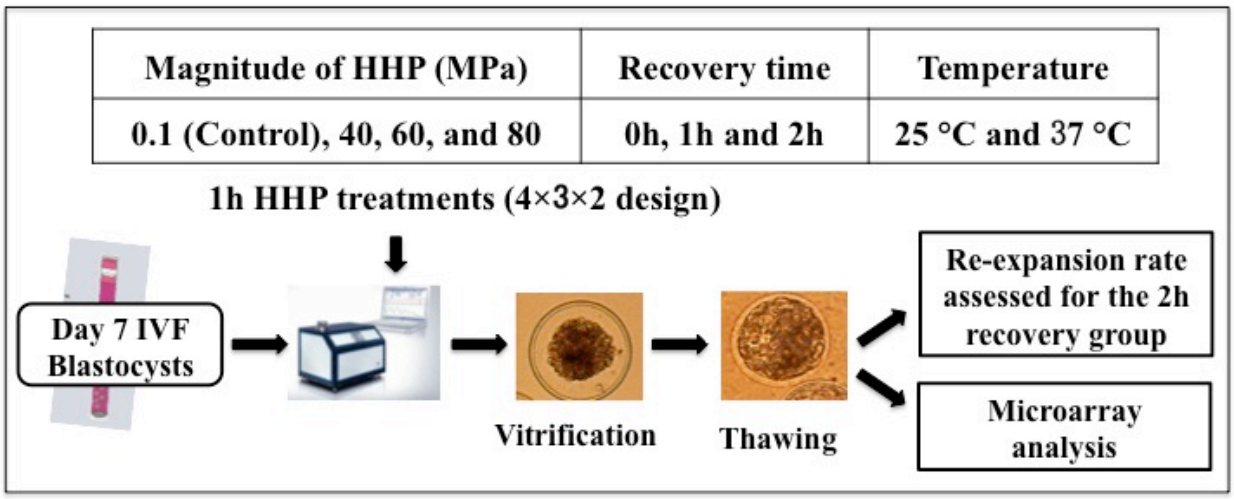


Figure 4.2. Re-expansion rates (mean ± SD) of vitrified and thawed bovine IVP blastocysts upon different HHP treatments with a 2h recovery time. (**: P-value < 0.01, *: P-value <0.05; n = 3).

Fig. 4.2

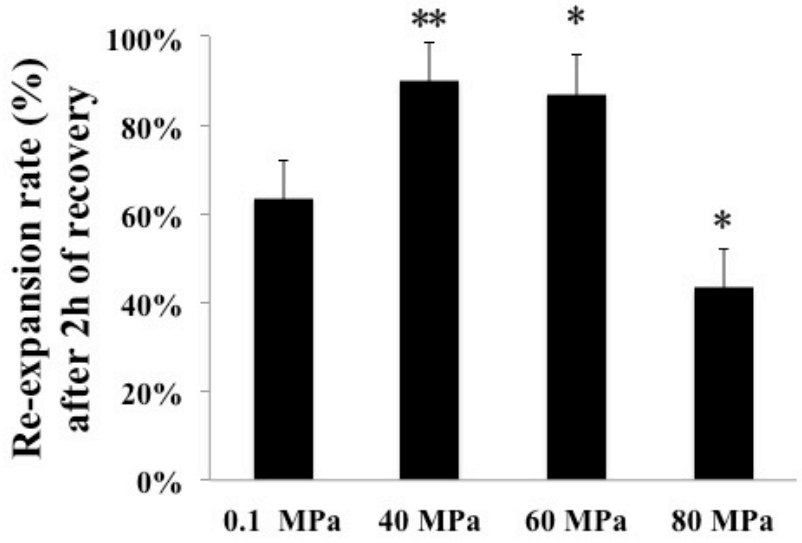


Figure 4.3. (A). Hierarchical clustering of all 12,274 analyzed transcripts among different HHP treatments and recovery times. Clear separations by HHP levels (0, 40, 60 and 80 MPa) were seen, demonstrating that pressure is the more significant factor than recovery time. (B). Heatmap of differentially expressed genes among different HHP treatments. The color spectrum, ranging from red to green, indicates normalized levels of gene expression from .

Fig. 4.3

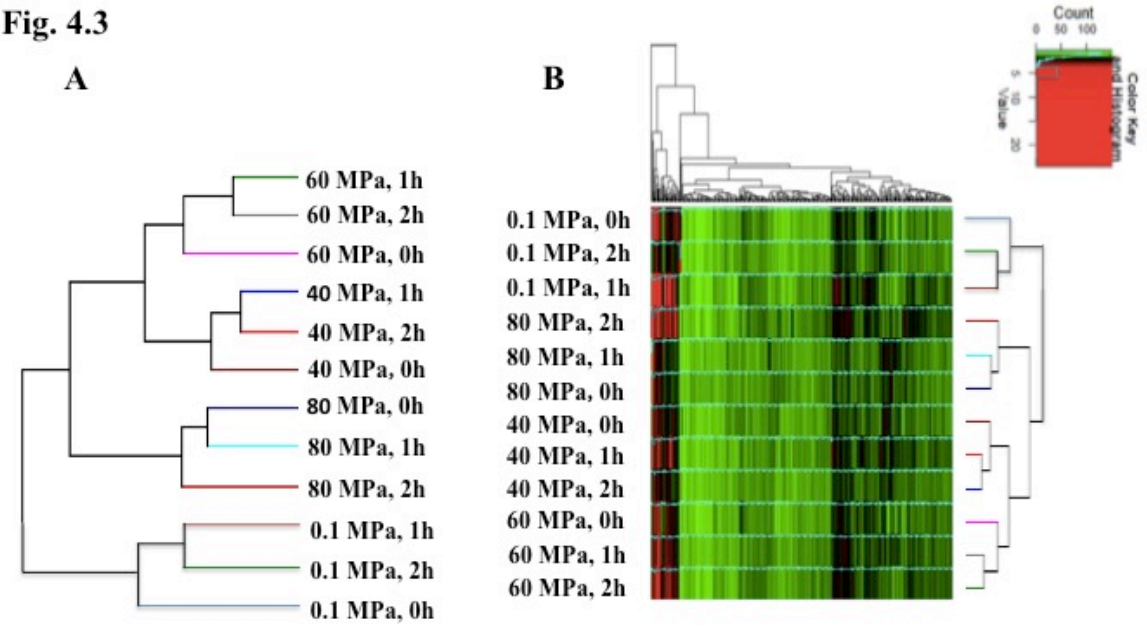


Figure 4.4. (A). The numbers of differentially expressed genes between HHP-treated embryos and controls. (B) Venn diagram shows the number of differentially genes specific to each comparison.

Fig. 4.4

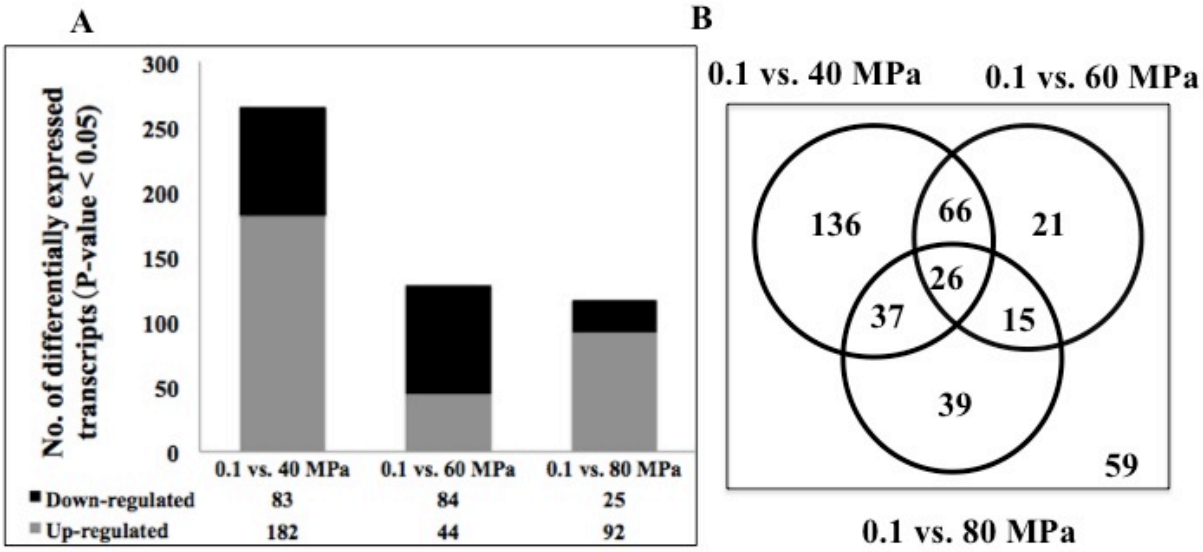


Figure 4.5. Modified apoptotic pathways in HHP-treated embryos. Genes in the green and red boxes were down- and up-regulated in both the 40- and 60-MPa treated groups (P-value < 0.05), respectively.

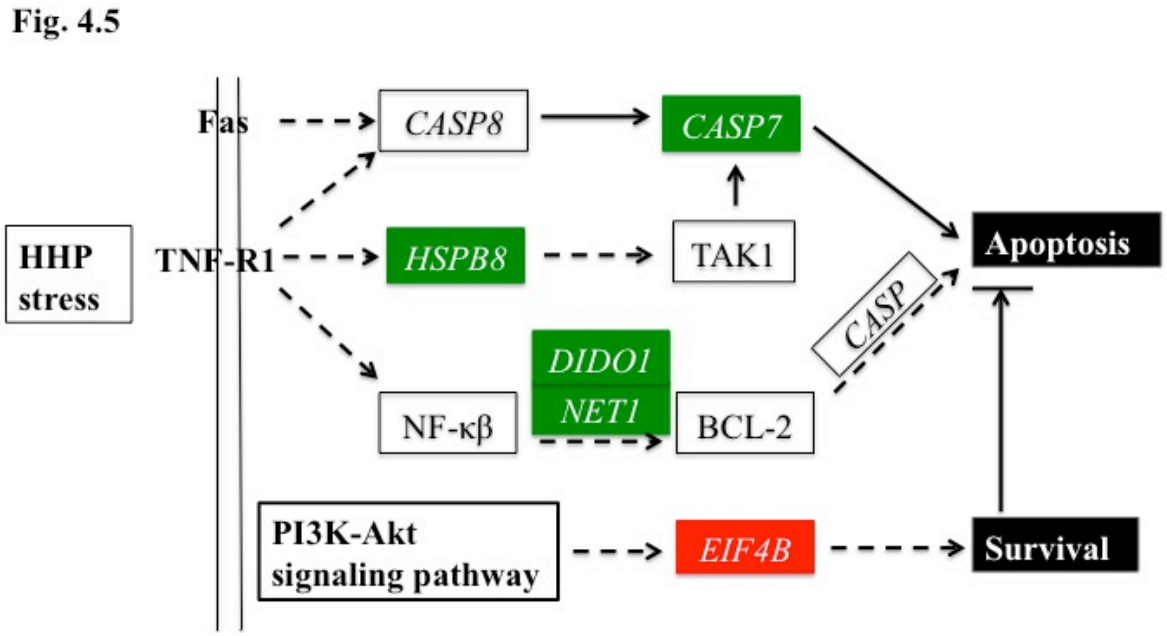


Figure 4.6. (A). The numbers of differentially expressed genes between embryos allowed recovery time and controls. (B) Venn diagram shows the number of differentially genes specific to each comparison.

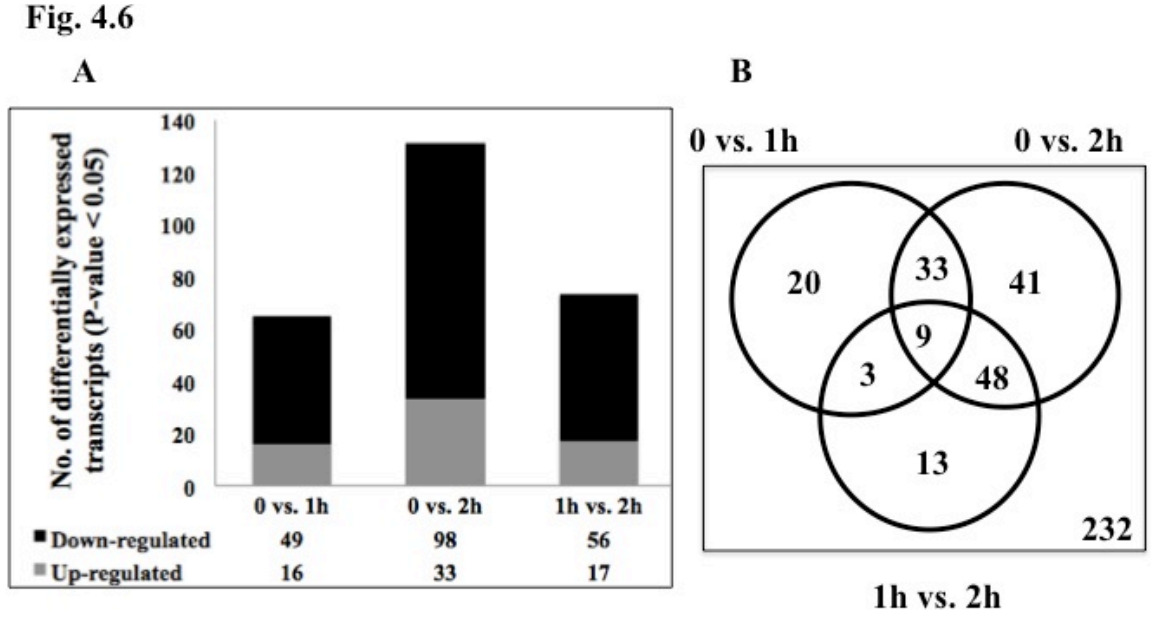
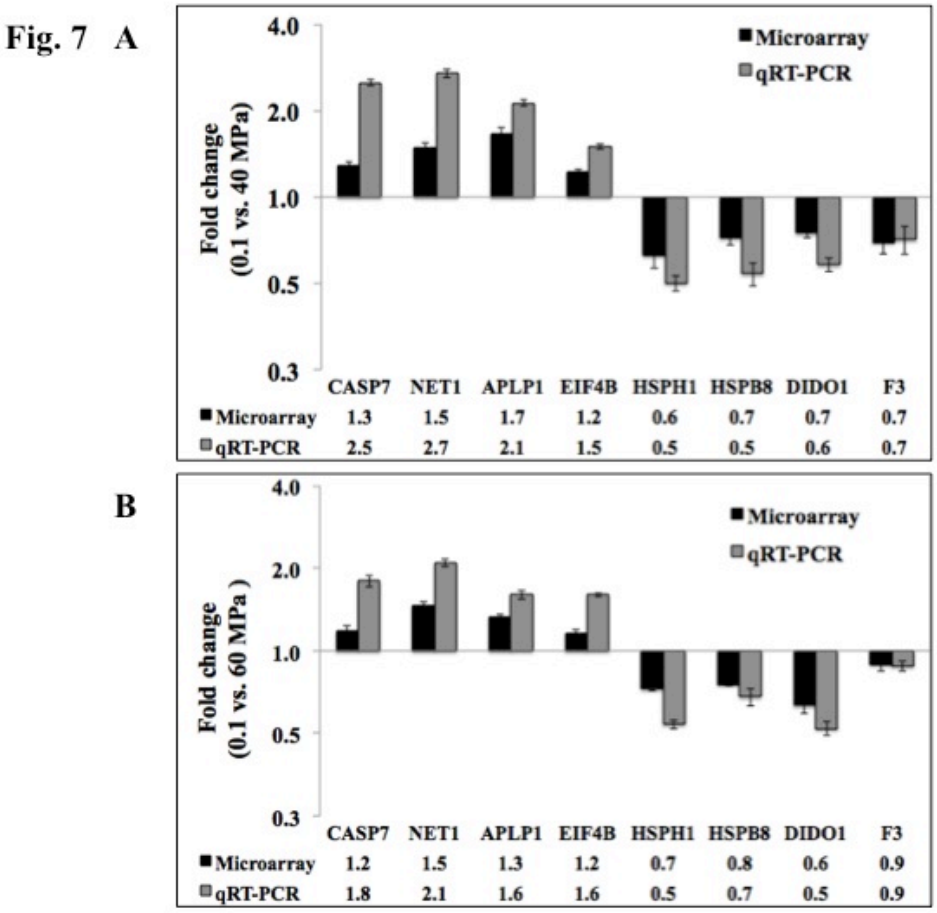


Figure 4.7. Comparisons of microarray and quantitative real-time RT-PCR (qRT-PCR) results of 8 selected genes from 40 and 60 MPa treated embryos. Fold change (mean ± SD) was expressed as the ratios of the levels of the 40 or 60 MPa treated embryos (n=3) to those of the controls (0.1 MPa) (n=3). In all cases, real time RT-PCR results substantiated the differential gene expression patterns from microarray.



4.6. Supplementary Information

Table S4.1: Primers for real time qRT-PCR.

Genes	Primers sequences (5' - 3')	Annealing temperature (°C)	Fragment size (bp)	Accession number
HSPB8	ACGACTTGACTGCCTCTTGG GCTTAAAGCTGTGCACGTTG	60	195	NM_001014955.1
HSPH1	ACAGCCCCAGGTACAAACTG TCCCTAACTGCCAGACCAAG	60	205	NM_001075302.1
CASP7	TGGAACAGATGGCAAGACAG CGGCTTCAACTGGGATCTTA	60	200	XM_002698509.3
NET1	CAAGCAAAGTGTTGCTCTGC CATCCTCACGTCTCCATCCT	60	197	XM_010811281.1
DIDO1	TCTCGACTCAGCACCATTTG AGCTCCTTGGACACTGAGGA	60	200	XM_010811602.1
EIF4B	TACAGGCGGGATGATAGAGG GCTGCAGTTTCTCCTGTTCC	60	205	NM_001035028.2
F3	GGTGATGCTGGTGCTGAGTA AGGAGGCATTGCTGACAATC	60	199	NM_182651.2
APLP1	GAACCAGACCCACAGCTCTC GTGGAAAGGCACAACTTGGT	60	195	NM_001038095.2
β-ACTIN	GACATCCGCAAGGACCTCTA ACATCTGCTGGAAGGTGGAC	60	205	NM_173979.3

Table S4.2: List of all differentially expressed genes in bovine IVP embryos affected by HHP treatments.

Probe	UniGene ID	Gene Name	Product	P-value
4927	Bt.55759	*DKFZp434K1815*	hypothetical protein DKFZp434K1815	6.86E-08
3757	Bt.2005	*LSM1*	LSM1 homolog, U6 small nuclear RNA associated (S. cerevisiae)	1.50E-07
6288	Bt.13965	*CNTNAP1*	contactin associated protein 1	2.48E-07
8731	Bt.27784	*PPP1R10*	protein phosphatase 1, regulatory (inhibitor) subunit 10	4.21E-07
2287	Bt.27376	*Bag4*	BCL2-associated athanogene 4	4.25E-07
2719	Bt.75092	*DIDO1*	death inducer-obliterator 1	4.49E-07
6779	Bt.90151	*LOC100008589*	28S ribosomal RNA	8.45E-07
2380	Bt.23564	*C6orf120*	chromosome 6 open reading frame 120	1.02E-06
8859	Bt.27784	*PPP1R10*	protein phosphatase 1, regulatory (inhibitor) subunit 10	1.13E-06
3495	Bt.8279	*C3orf60*	chromosome 3 open reading frame 60	1.30E-06
9586	Bt.88378	*KIAA1303*	raptor	1.52E-06
6419	Bt.13965	*CNTNAP1*	contactin associated protein 1	1.56E-06
288	Bt.28094	*CASP7*	caspase 7, apoptosis-related cysteine peptidase	1.68E-06
3623	Bt.8279	*C3orf60*	chromosome 3 open reading frame 60	1.85E-06
9896	Bt.58080	*ZNF3*	zinc finger protein 3	1.85E-06
10009	Bt.48940	*PPAP2B*	phosphatidic acid phosphatase type 2B	1.96E-06
2357	Bt.5184	*Ankrd40*	ankyrin repeat domain 40	2.39E-06
2263	Bt.23564	*C6orf120*	chromosome 6 open reading frame 120	2.66E-06
6641	Bt.90151	*LOC100008589*	28S ribosomal RNA	2.96E-06
1090	Bt.7391	*APLP1*	amyloid beta (A4) precursor-like protein 1	3.84E-06
1536	Bt.66354	*POLR2L*	polymerase (RNA) II (DNA directed) polypeptide L, 7.6kDa	6.22E-06
9438	Bt.88378	*KIAA1303*	raptor	6.94E-06
2402	Bt.7873	*BCAM*	basal cell adhesion molecule (Lutheran blood group)	6.99E-06
9762	Bt.58080	*ZNF3*	zinc finger protein 3	7.52E-06
6978	Bt.22374	*MGC137396*	similar to CG9339-PF	9.39E-06
3888	Bt.2005	*LSM1*	LSM1 homolog, U6 small nuclear RNA associated (S. cerevisiae)	9.51E-06
1895	Bt.42858	*ABCC5*	ATP-binding cassette, sub-family C (CFTR/MRP), member 5	1.02E-05
2285	Bt.7873	*BCAM*	basal cell adhesion molecule (Lutheran blood group)	1.10E-05
3093	Bt.53211	*LOC615039*	hypothetical LOC615039	1.14E-05
1662	Bt.66354	*POLR2L*	polymerase (RNA) II (DNA directed) polypeptide L, 7.6kDa	1.19E-05
3177	Bt.48951	*ASB11*	ankyrin repeat and SOCS box-containing 11	1.21E-05
747	Bt.28249	*SEC24A*	SEC24 related gene family, member A (S. cerevisiae)	1.53E-05
2575	Bt.2041	*AP2A2*	adaptor-related protein complex 2, alpha 2 subunit	1.71E-05
9020	Bt.4586	*GNG2*	guanine nucleotide binding protein (G protein), gamma 2	1.81E-05
1477	Bt.12399	*NEU1*	sialidase 1 (lysosomal sialidase)	1.84E-05
6849	Bt.22374	*MGC137396*	similar to CG9339-PF	1.89E-05

8159	Bt.89557	*NA*	Transcribed locus	1.96E-05
4237	Bt.87845	*COX5B*	cytochrome c oxidase subunit Vb	2.00E-05
4305	Bt.89659	*FKBP9L*	FK506 binding protein 9-like	2.02E-05
8416	Bt.21965	*ZNF358*	zinc finger protein 358	2.10E-05
1767	Bt.42858	*ABCC5*	ATP-binding cassette, sub-family C (CFTR/MRP), member 5	2.23E-05
5625	Bt.45080	*LRRC45*	leucine rich repeat containing 45	2.36E-05
8432	Bt.5633	*TMEM129*	transmembrane protein 129	2.38E-05
7909	Bt.64563	*RFFL*	ring finger and FYVE-like domain containing 1	2.54E-05
5766	Bt.65363	*C16orf69*	chromosome 16 open reading frame 69	2.56E-05
5328	Bt.55867	*HSPC152*	hypothetical protein HSPC152	2.62E-05
320		*TMEM121*	transmembrane protein 121	2.64E-05
8747	Bt.23521	*DUSP26*	dual specificity phosphatase 26 (putative)	2.74E-05
2441	Bt.2041	*AP2A2*	adaptor-related protein complex 2, alpha 2 subunit	3.05E-05
8883	Bt.4586	*GNG2*	guanine nucleotide binding protein (G protein), gamma 2	3.28E-05
11924	Bt.37906	*YOD1*	YOD1 OTU deubiquinating enzyme 1 homolog (S. cerevisiae)	3.39E-05
893	Bt.59331	*UAP1*	UDP-N-acteylglucosamine pyrophosphorylase 1	3.55E-05
1897	Bt.12732	*NYREN18*	NEDD8 ultimate buster-1	3.57E-05
2569	Bt.7562	*FLJ41352*	FLJ41352 protein	3.66E-05
4162	Bt.89659	*FKBP9L*	FK506 binding protein 9-like	3.78E-05
5332	Bt.12314	*PFKFB3*	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3	3.97E-05
904	Bt.7418	*ANKH*	ankylosis, progressive homolog (mouse)	4.40E-05
4243	Bt.6710	*LOC440330*	hypothetical LOC440330	4.50E-05
4349	Bt.66386	*WRNIP1*	Werner helicase interacting protein 1	4.59E-05
4380	Bt.87845	*COX5B*	cytochrome c oxidase subunit Vb	4.62E-05
8808	Bt.53659	*ABHD13*	abhydrolase domain containing 13	4.64E-05
3287	Bt.48951	*ASB11*	ankyrin repeat and SOCS box-containing 11	4.73E-05
1436	Bt.6451	*HIST1H1E*	histone cluster 1, H1e	4.92E-05
1603	Bt.12399	*NEU1*	sialidase 1 (lysosomal sialidase)	5.03E-05
8644	Bt.8738	*HSPH1*	heat shock 105kDa/110kDa protein 1	5.51E-05
10144	Bt.48940	*PPAP2B*	phosphatidic acid phosphatase type 2B	5.70E-05
5355	Bt.269	*ATP2C1*	ATPase, Ca++ transporting, type 2C, member 1	5.96E-05
2339	Bt.52556	*RCP9*	calcitonin gene-related peptide-receptor component protein	6.21E-05
781	Bt.7418	*ANKH*	ankylosis, progressive homolog (mouse)	6.37E-05
1841	Bt.12138	*RP11-413M3.2*	similar to CG12379-PA	6.48E-05
4205	Bt.66386	*WRNIP1*	Werner helicase interacting protein 1	6.48E-05
586	Bt.12578	*ORC5L*	origin recognition complex, subunit 5-like (yeast)	6.51E-05
5212	Bt.12314	*PFKFB3*	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3	6.60E-05
770	Bt.59331	*UAP1*	UDP-N-acteylglucosamine pyrophosphorylase 1	6.66E-05
8983	Bt.45485	*KLF9*	Kruppel-like factor 9	6.68E-05
1685	Bt.77469	*KCTD20*	potassium channel tetramerisation domain containing 20	6.84E-05

967	Bt.73261	*LOC513955*	hypothetical LOC513955	6.85E-05
8617	Bt.23521	*DUSP26*	dual specificity phosphatase 26 (putative)	6.93E-05
9505	Bt.46839	*SNRPD2*	small nuclear ribonucleoprotein D2 polypeptide 16.5kDa	6.96E-05
4808	Bt.55759	*DKFZp434K1815*	hypothetical protein DKFZp434K1815	7.42E-05
4386	Bt.6710	*LOC440330*	hypothetical LOC440330	7.53E-05
10996	Bt.22526	*HSPB8*	heat shock 22kDa protein 8	7.83E-05
8250	Bt.22175	*KCNT1*	potassium channel, subfamily T, member 1	7.92E-05
9907	Bt.59214	*FGD5*	FYVE, RhoGEF and PH domain containing 5	8.01E-05
9654	Bt.46839	*SNRPD2*	small nuclear ribonucleoprotein D2 polypeptide 16.5kDa	8.10E-05
4144	Bt.49595	*TMEM111*	transmembrane protein 111	8.44E-05
8544	Bt.21965	*ZNF358*	zinc finger protein 358	8.87E-05
4834	Bt.8545	*MRPL43*	mitochondrial ribosomal protein L43	9.09E-05
5208	Bt.55867	*HSPC152*	hypothetical protein HSPC152	9.14E-05
9121	Bt.45485	*KLF9*	Kruppel-like factor 9	9.23E-05
8774	Bt.8738	*HSPH1*	heat shock 105kDa/110kDa protein 1	9.29E-05
3089	Bt.22538	*FZD8*	frizzled homolog 8 (Drosophila)	9.71E-05
2133	Bt.4185	*NHP2L1*	NHP2 non-histone chromosome protein 2-like 1 (S. cerevisiae)	1.01E-04
5945	Bt.26976	*COX11*	COX11 homolog, cytochrome c oxidase assembly protein (yeast)	1.02E-04
5154	Bt.52786	*LOC100008588*	18S ribosomal RNA	1.03E-04
2970	Bt.53211	*LOC615039*	hypothetical LOC615039	1.03E-04
1968	Bt.53760	*LOC728216*	similar to ubiquitin specific peptidase 18	1.05E-04
5262	Bt.8206	*SFRS7*	splicing factor, arginine/serine-rich 7, 35kDa	1.05E-04
7305	Bt.89284	*LOC100008589*	28S ribosomal RNA	1.06E-04
2404	Bt.27376	*Bag4*	BCL2-associated athanogene 4	1.06E-04
7223	Bt.500	*HK1*	hexokinase 1	1.09E-04
3865	Bt.52937	*hCG_20417*	mitochondrial ribosomal protein L20	1.09E-04
403	Bt.28094	*CASP7*	caspase 7, apoptosis-related cysteine peptidase	1.13E-04
1769	Bt.12732	*NYREN18*	NEDD8 ultimate buster-1	1.13E-04
5768	Bt.49597	*LOC528919*	similar to F-box and WD-40 domain protein 12	1.13E-04
3414	Bt.61384	*UNC84B*	unc-84 homolog B (C. elegans)	1.16E-04
2079	Bt.53760	*LOC728216*	similar to ubiquitin specific peptidase 18	1.17E-04
7060	Bt.11280	*BANP*	BTG3 associated nuclear protein	1.22E-04
6385	Bt.9092	*POLR2G*	polymerase (RNA) II (DNA directed) polypeptide G	1.23E-04
3885	Bt.51520	*C14orf94*	chromosome 14 open reading frame 94	1.24E-04
4656	Bt.32544	*NVL*	nuclear VCP-like	1.28E-04
7185	Bt.11280	*BANP*	BTG3 associated nuclear protein	1.32E-04
9205	Bt.21568	*FAM118B*	family with sequence similarity 118, member B	1.32E-04
1402	Bt.4279	*SHC1*	SHC (Src homology 2 domain containing) transforming protein 1	1.35E-04
8583	Bt.52379	*QRICH1*	glutamine-rich 1	1.37E-04

7986	Bt.66	*NDUFA3*	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 3, 9kDa	1.38E-04
3733	Bt.52937	*hCG_20417*	mitochondrial ribosomal protein L20	1.40E-04
2027	Bt.10259	*PLK2*	polo-like kinase 2 (Drosophila)	1.47E-04
4986	Bt.89574	*KIAA0182*	KIAA0182	1.50E-04
9829	Bt.27267	*TMEM4*	transmembrane protein 4	1.51E-04
6659	Bt.3196	*STRA6*	stimulated by retinoic acid gene 6 homolog (mouse)	1.52E-04
5633	Bt.2568	*ADFP*	adipose differentiation-related protein	1.57E-04
123	Bt.64291	*H3F3A*	H3 histone, family 3A	1.58E-04
5025	Bt.52786	*LOC100008588*	18S ribosomal RNA	1.61E-04
393	Bt.58700	*C9orf58*	chromosome 9 open reading frame 58	1.67E-04
9501				1.69E-04
11807	Bt.37906	*YOD1*	YOD1 OTU deubiquinating enzyme 1 homolog (S. cerevisiae)	1.70E-04
870	Bt.28249	*SEC24A*	SEC24 related gene family, member A (S. cerevisiae)	1.74E-04
10034	Bt.7363	*LOC789389*	similar to Wolf-Hirschhorn syndrome candidate 2 protein	1.75E-04
7097	Bt.500	*HK1*	hexokinase 1	1.87E-04
3739	Bt.28081	*POGZ*	pogo transposable element with ZNF domain	1.89E-04
89	Bt.32543	*IPO8*	importin 8	1.91E-04
3754	Bt.51520	*C14orf94*	chromosome 14 open reading frame 94	1.92E-04
2654	Bt.4298	*GPR172B*	G protein-coupled receptor 172B	1.94E-04
626	Bt.38424	*DST*	dystonin	1.99E-04
5818	Bt.26976	*COX11*	COX11 homolog, cytochrome c oxidase assembly protein (yeast)	2.01E-04
5495	Bt.2568	*ADFP*	adipose differentiation-related protein	2.03E-04
1640	Bt.49587	*GPI*	glucose phosphate isomerase	2.07E-04
7811	Bt.39540	*IL15RA*	interleukin 15 receptor, alpha	2.07E-04
7448	Bt.89284	*LOC100008589*	28S ribosomal RNA	2.08E-04
3263	Bt.37779	*BPHL*	biphenyl hydrolase-like (serine hydrolase; breast epithelial mucin-associated antigen)	2.11E-04
9186	Bt.65788		Transcribed locus	2.14E-04
2167	Bt.20080	*WSB1*	WD repeat and SOCS box-containing 1	2.15E-04
9208	Bt.26958	*LOC788996*	similar to Potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 2 (Brain cyclic nucleotide-gated channel 2) (BCNG-2)	2.16E-04
6252	Bt.9092	*POLR2G*	polymerase (RNA) II (DNA directed) polypeptide G	2.21E-04
1401	Bt.3199	*RUSC1*	RUN and SH3 domain containing 1	2.21E-04
2193	Bt.52796	*COPS6*	COP9 constitutive photomorphogenic homolog subunit 6 (Arabidopsis)	2.25E-04
7856	Bt.66	*NDUFA3*	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 3, 9kDa	2.26E-04
5236	Bt.269	*ATP2C1*	ATPase, Ca++ transporting, type 2C, member 1	2.29E-04
8035	Bt.38851	*RHPN2*	rhophilin, Rho GTPase binding protein 2	2.30E-04
2807	Bt.64366	*NDUFC2*	NADH dehydrogenase (ubiquinone) 1, subcomplex unknown, 2, 14.5kDa	2.41E-04
2521	Bt.4298	*GPR172B*	G protein-coupled receptor 172B	2.52E-04

11261	Bt.9525	*RNF113A*	ring finger protein 113A	2.58E-04
5727	Bt.70	*NDUFB10*	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 10, 22kDa	2.60E-04
5271	Bt.4908	*RBBP7*	retinoblastoma binding protein 7	2.66E-04
5894	Bt.65363	*C16orf69*	chromosome 16 open reading frame 69	2.71E-04
5840	Bt.1658	*DUSP1*	dual specificity phosphatase 1	2.76E-04
2986	Bt.27527	*STXBP3*	syntaxin binding protein 3	2.78E-04
9337	Bt.26958	*LOC788996*	similar to Potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 2 (Brain cyclic nucleotide-gated channel 2) (BCNG-2)	2.80E-04
2223	Bt.52556	*RCP9*	calcitonin gene-related peptide-receptor component protein	2.81E-04
12100	Bt.30953	*DOCK1*	dedicator of cytokinesis 1	3.01E-04
6155	Bt.7804	*GYG1*	glycogenin 1	3.01E-04
7816	Bt.6438	*TGFB2*	transforming growth factor, beta 2	3.05E-04
1560	Bt.77469	*KCTD20*	potassium channel tetramerisation domain containing 20	3.11E-04
6408	Bt.52209	*RAD54B*	RAD54 homolog B (S. cerevisiae)	3.16E-04
2274	Bt.44330	*LRP10*	low density lipoprotein receptor-related protein 10	3.17E-04
4802	Bt.22766	*SART3*	squamous cell carcinoma antigen recognized by T cells 3	3.23E-04
5966	Bt.5519	*TncRNA*	trophoblast-derived noncoding RNA	3.28E-04
1689	Bt.39564	*LOC506268*	similar to C2orf29 protein	3.33E-04
8404	Bt.76983	*BIRC5*	Baculoviral IAP repeat-containing 5 (survivin)	3.43E-04
8640	Bt.44264	*LTF*	lactotransferrin	3.51E-04
5165	Bt.2056	*APEH*	N-acylaminoacyl-peptide hydrolase	3.57E-04
3375	Bt.37779	*BPHL*	biphenyl hydrolase-like (serine hydrolase; breast epithelial mucin-associated antigen)	3.58E-04
2830	Bt.75092	*DIDO1*	death inducer-obliterator 1	3.64E-04
11731	Bt.22575	*TMEM167*	transmembrane protein 167	3.68E-04
506	Bt.6024	*POLR3K*	polymerase (RNA) III (DNA directed) polypeptide K, 12.3 kDa	3.70E-04
8460	Bt.5530	*DHRS3*	dehydrogenase/reductase (SDR family) member 3	3.77E-04
8533	Bt.76983	*BIRC5*	Baculoviral IAP repeat-containing 5 (survivin)	3.79E-04
8128	Bt.33726	*NPY*	neuropeptide Y	3.89E-04
385	Bt.36587	*APLN*	apelin, AGTRL1 ligand	3.91E-04
4066	Bt.73096	*CTH*	cystathionase (cystathionine gamma-lyase)	4.04E-04
7061	Bt.42818	*POLR3B*	polymerase (RNA) III (DNA directed) polypeptide B	4.06E-04
1273	Bt.4279	*SHC1*	SHC (Src homology 2 domain containing) transforming protein 1	4.09E-04
758	Bt.21023	*AS3MT*	arsenic (+3 oxidation state) methyltransferase	4.10E-04
3541	Bt.61384	*UNC84B*	unc-84 homolog B (C. elegans)	4.17E-04
3518	Bt.28502	*FAM83D*	family with sequence similarity 83, member D	4.18E-04
281	Bt.9783	*DDX3X*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked	4.21E-04
2111	Bt.9186	*FBXL3*	F-box and leucine-rich repeat protein 3	4.28E-04
8679	Bt.53659	*ABHD13*	abhydrolase domain containing 13	4.32E-04

2240	Bt.5184	*Ankrd40*	ankyrin repeat domain 40	4.36E-04
9071	Bt.66625	*CNOT4*	CCR4-NOT transcription complex, subunit 4	4.38E-04
2887	Bt.61617	*UHRF2*	ubiquitin-like, containing PHD and RING finger domains, 2	4.40E-04
2559	Bt.64693	*Dnaja1*	DnaJ (Hsp40) homolog, subfamily A, member 1	4.42E-04
42	Bt.34521	*RBM12*	RNA binding motif protein 12	4.53E-04
6830	Bt.52335	*IFT57*	intraflagellar transport 57 homolog (Chlamydomonas)	4.55E-04
8292	Bt.89557		Transcribed locus	4.63E-04
11028	Bt.30639	*CNOT2*	CCR4-NOT transcription complex, subunit 2	4.67E-04
7556	Bt.63484	*SURF2*	surfeit 2	4.67E-04
1307	Bt.6451	*HIST1H1E*	histone cluster 1, H1e	4.71E-04
13	Bt.64291	*H3F3A*	H3 histone, family 3A	4.71E-04
708	Bt.732		Transcribed locus	4.72E-04
3393	Bt.28502	*FAM83D*	family with sequence similarity 83, member D	4.74E-04
11811	Bt.625	*C6orf106*	hypothetical protein LOC617655	4.75E-04
1858	Bt.62530	*ITPK1*	inositol 1,3,4-triphosphate 5/6 kinase	4.83E-04
4505	Bt.15915	*IPO11*	importin 11	4.89E-04
437		*TMEM121*	transmembrane protein 121	4.92E-04
1019	Bt.9195	*SIVA1*	SIVA1, apoptosis-inducing factor	4.93E-04
4043				5.09E-04
4008	Bt.26835	*RANGNRF*	RAN guanine nucleotide release factor	5.11E-04
5896	Bt.49597	*LOC528919*	similar to F-box and WD-40 domain protein 12	5.15E-04
10194	Bt.62845	*ASAH1*	N-acylsphingosine amidohydrolase (acid ceramidase) 1	5.18E-04
6897	Bt.21510	*GPR107*	G protein-coupled receptor 107	5.28E-04
3939	Bt.73096	*CTH*	cystathionase (cystathionine gamma-lyase)	5.28E-04
3585	Bt.26997	*Ptpmt1*	protein tyrosine phosphatase, mitochondrial 1	5.31E-04
1794	Bt.64539	*QSCN6*	quiescin Q6	5.35E-04
538	Bt.13737	*GALK2*	galactokinase 2	5.44E-04
1290	Bt.63116	*AKR1B1*	aldo-keto reductase family 1, member B1 (aldose reductase)	5.50E-04
4532	Bt.4949	*SF3B5*	splicing factor 3b, subunit 5, 10kDa	5.53E-04
4001	Bt.26739	*TOMM40L*	translocase of outer mitochondrial membrane 40 homolog (yeast)-like	5.57E-04
1109	Bt.17683	*DBF4*	DBF4 homolog (S. cerevisiae)	5.60E-04
6960	Bt.52335	*IFT57*	intraflagellar transport 57 homolog (Chlamydomonas)	5.75E-04
7496	Bt.10332	*Cramp1l*	Crm, cramped-like (Drosophila)	5.81E-04
4952	Bt.2642	*ZNF706*	zinc finger protein 706	5.94E-04
2391	Bt.44330	*LRP10*	low density lipoprotein receptor-related protein 10	6.03E-04
779	Bt.90218	*TES*	testis derived transcript (3 LIM domains)	6.11E-04
5487	Bt.45080	*LRRC45*	leucine rich repeat containing 45	6.11E-04
10232	Bt.1858	*C20orf108*	chromosome 20 open reading frame 108	6.16E-04
971	Bt.7391	*APLP1*	amyloid beta (A4) precursor-like protein 1	6.19E-04
8117	Bt.22175	*KCNT1*	potassium channel, subfamily T, member 1	6.19E-04
5040	Bt.1786	*YME1L1*	YME1-like 1 (S. cerevisiae)	6.20E-04

11140	Bt.9525	*RNF113A*	ring finger protein 113A	6.29E-04
9841	Bt.87389	*GAPDH*	glyceraldehyde-3-phosphate dehydrogenase	6.29E-04
8877	Bt.48519	*RBPMS*	RNA binding protein with multiple splicing	6.31E-04
1902	Bt.12743	*ABCD4*	ATP-binding cassette, sub-family D (ALD), member 4	6.35E-04
10681	Bt.59095	*C9orf97*	chromosome 9 open reading frame 97	6.48E-04
279	Bt.58700	*C9orf58*	chromosome 9 open reading frame 58	6.48E-04
1514	Bt.49587	*GPI*	glucose phosphate isomerase	6.52E-04
5186	Bt.57711	*NDUFB8*	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 8, 19kDa	6.55E-04
4400	Bt.49192	*RPL15*	ribosomal protein L15	6.58E-04
2020	Bt.4185	*NHP2L1*	NHP2 non-histone chromosome protein 2-like 1 (S. cerevisiae)	6.65E-04
1419	Bt.63116	*AKR1B1*	aldo-keto reductase family 1, member B1 (aldose reductase)	6.78E-04
2471	Bt.49501	*C15orf40*	chromosome 15 open reading frame 40	6.81E-04
8041	Bt.64563	*RFFL*	ring finger and FYVE-like domain containing 1	6.96E-04
4274	Bt.11909	*DUSP6*	dual specificity phosphatase 6	7.01E-04
5773	Bt.38862	*RALA*	v-ral simian leukemia viral oncogene homolog A (ras related)	7.03E-04
4027	Bt.11096	*KPNA4*	karyopherin alpha 4 (importin alpha 3)	7.04E-04
2112	Bt.44470	*SATB1*	SATB homeobox 1	7.05E-04
4931	Bt.27370	*HEATR3*	HEAT repeat containing 3	7.07E-04
5036	Bt.2056	*APEH*	N-acylaminoacyl-peptide hydrolase	7.16E-04
6501	Bt.48788	*RAB30*	RAB30, member RAS oncogene family	7.18E-04
5382	Bt.8206	*SFRS7*	splicing factor, arginine/serine-rich 7, 35kDa	7.25E-04
10097	Bt.1858	*C20orf108*	chromosome 20 open reading frame 108	7.25E-04
519	Bt.49151	*LOC729046*	similar to 60S ribosomal protein L17 (L23)	7.50E-04
11849	Bt.9469	*DNAJC10*	DnaJ (Hsp40) homolog, subfamily C, member 10	7.59E-04
6124	Bt.43820	*RPL36*	ribosomal protein L36	7.60E-04
1818	Bt.62762	*LOC124446*	hypothetical protein BC017488	7.64E-04
2435	Bt.7562	*FLJ41352*	FLJ41352 protein	7.64E-04
4263	Bt.23076	*C1orf93*	chromosome 1 open reading frame 93	7.65E-04
636	Bt.6024	*POLR3K*	polymerase (RNA) III (DNA directed) polypeptide K, 12.3 kDa	7.67E-04
2709	Bt.21070	*MRPL10*	mitochondrial ribosomal protein L10	7.69E-04
5854	Bt.70	*NDUFB10*	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 10, 22kDa	7.71E-04
4257	Bt.49192	*RPL15*	ribosomal protein L15	7.74E-04
4914	Bt.64707	*LOC146346*	hypothetical protein LOC146346	7.76E-04
4406	Bt.23076	*C1orf93*	chromosome 1 open reading frame 93	7.84E-04
7684	Bt.39540	*IL15RA*	interleukin 15 receptor, alpha	7.85E-04
265	Bt.20277	*TOP2A*	topoisomerase (DNA) II alpha 170kDa	7.86E-04
4018	Bt.49595	*TMEM111*	transmembrane protein 111	7.88E-04
4665	Bt.4949	*SF3B5*	splicing factor 3b, subunit 5, 10kDa	7.98E-04
8266	Bt.16440	*USP6NL*	USP6 N-terminal like	8.00E-04

5987	Bt.13810	*U2AF1*	U2 small nuclear RNA auxiliary factor 1	8.16E-04
8588	Bt.5530	*DHRS3*	dehydrogenase/reductase (SDR family) member 3	8.16E-04
5	Bt.29924	*MDM4*	Mdm4, transformed 3T3 cell double minute 4, p53 binding protein (mouse)	8.25E-04
6339	Bt.44021	*WDR8*	WD repeat domain 8	8.30E-04
2053	Bt.20080	*WSB1*	WD repeat and SOCS box-containing 1	8.32E-04
1886	Bt.76825	*LOC789371*	hypothetical protein LOC789371	8.34E-04
395	Bt.9783	*DDX3X*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked	8.36E-04
942	Bt.6650	*GGH*	gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)	8.42E-04
1782	Bt.53344	*ARHGAP21*	Rho GTPase activating protein 21	8.51E-04
3806	Bt.25997	*CABLES1*	Cdk5 and Abl enzyme substrate 1	8.53E-04
10961	Bt.20683	*TP53INP2*	tumor protein p53 inducible nuclear protein 2	8.75E-04
4345	Bt.17589	*DMTF1*	cyclin D binding myb-like transcription factor 1	8.80E-04
2992	Bt.18512	*ASF1A*	ASF1 anti-silencing function 1 homolog A (S. cerevisiae)	8.81E-04
289	Bt.21926	*LOC508142*	similar to translocase of outer mitochondrial membrane 34	8.82E-04
4418	Bt.11909	*DUSP6*	dual specificity phosphatase 6	8.90E-04
2776	Bt.61617	*UHRF2*	ubiquitin-like, containing PHD and RING finger domains, 2	9.04E-04
6018	Bt.7804	*GYG1*	glycogenin 1	9.06E-04
3472	Bt.49273	*POLR2F*	polymerase (RNA) II (DNA directed) polypeptide F	9.07E-04
7594	Bt.11282		Transcribed locus	9.17E-04
2877	Bt.30434	*SFRS9*	splicing factor, arginine/serine-rich 9	9.18E-04
884	Bt.49509	*LZTFL1*	leucine zipper transcription factor-like 1	9.19E-04
10764				9.47E-04
3054	Bt.43655	*MGC152057*	hypothetical LOC507084	9.47E-04
11275	Bt.2988	*GNG11*	guanine nucleotide binding protein (G protein), gamma 11	9.47E-04
11761	Bt.17953	*FOXP1*	forkhead box P1	9.56E-04
564	Bt.28460	*SNRPD3*	small nuclear ribonucleoprotein D3 polypeptide 18kDa	9.71E-04
7968	Bt.49157	*APOA1*	apolipoprotein A-I	9.73E-04
3701	Bt.24435	*SYNCRIP*	synaptotagmin binding, cytoplasmic RNA interacting protein	9.75E-04
3833	Bt.24435	*SYNCRIP*	synaptotagmin binding, cytoplasmic RNA interacting protein	9.81E-04
2634	Bt.90254	*HDAC8*	histone deacetylase 8	9.81E-04
5114	Bt.89574	*KIAA0182*	KIAA0182	9.88E-04
10059	Bt.62845	*ASAH1*	N-acylsphingosine amidohydrolase (acid ceramidase) 1	9.96E-04
2555	Bt.5528	*SLC7A5*	solute carrier family 7 (cationic amino acid transporter, y+ system), member 5	1.01E-03
8529	Bt.7930	*NUCB1*	nucleobindin 1	1.01E-03
2422	Bt.5528	*SLC7A5*	solute carrier family 7 (cationic amino acid transporter, y+ system), member 5	1.01E-03
1116	Bt.25186	*BCAT1*	branched chain aminotransferase 1, cytosolic	1.01E-03
10864	Bt.22526	*HSPB8*	heat shock 22kDa protein 8	1.02E-03

7027	Bt.21510	*GPR107*	G protein-coupled receptor 107	1.03E-03
8133	Bt.16440	*USP6NL*	USP6 N-terminal like	1.05E-03
6908	Bt.2453	*POLR2H*	polymerase (RNA) II (DNA directed) polypeptide H	1.05E-03
4183	Bt.89506	*NSL1*	NSL1, MIND kinetochore complex component, homolog (S. cerevisiae)	1.05E-03
2934	Bt.17132	*KRIT1*	KRIT1, ankyrin repeat containing	1.06E-03
3818	Bt.22310	*ARIH2*	ariadne homolog 2 (Drosophila)	1.07E-03
10834	Bt.18408	*CCDC117*	coiled-coil domain containing 117	1.07E-03
1731	Bt.62530	*ITPK1*	inositol 1,3,4-triphosphate 5/6 kinase	1.07E-03
4884		*SEMA3G*	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3G	1.08E-03
3736	Bt.49705	*PRDX6*	peroxiredoxin 6	1.09E-03
3794	Bt.44982	*CHRNA10*	cholinergic receptor, nicotinic, alpha 10	1.09E-03
3045	Bt.36226	*SMNDC1*	survival motor neuron domain containing 1	1.10E-03
9773	Bt.59214	*FGD5*	FYVE, RhoGEF and PH domain containing 5	1.10E-03
3140	Bt.56130	*SIAHBP1*	fuse-binding protein-interacting repressor	1.11E-03
11668	Bt.49238	*HPRT1*	hypoxanthine phosphoribosyltransferase 1 (Lesch-Nyhan syndrome)	1.12E-03
6217	Bt.11128	*MGC140717*	similar to uracil-DNA glycosylase isoform UNG2	1.13E-03
2865	Bt.89629	*EIF4B*	eukaryotic translation initiation factor 4B	1.14E-03
4134	Bt.26835	*RANGNRF*	RAN guanine nucleotide release factor	1.17E-03
3109	Bt.27527	*STXBP3*	syntaxin binding protein 3	1.18E-03
3671	Bt.27129	*SLC39A9*	solute carrier family 39 (zinc transporter), member 9	1.18E-03
11544	Bt.49238	*HPRT1*	hypoxanthine phosphoribosyltransferase 1 (Lesch-Nyhan syndrome)	1.21E-03
5839	Bt.5519	*TncRNA*	trophoblast-derived noncoding RNA	1.21E-03
9706	Bt.87389	*GAPDH*	glyceraldehyde-3-phosphate dehydrogenase	1.22E-03
3017	Bt.56130	*SIAHBP1*	fuse-binding protein-interacting repressor	1.22E-03
989	Bt.17683	*DBF4*	DBF4 homolog (S. cerevisiae)	1.23E-03
11879	Bt.17953	*FOXP1*	forkhead box P1	1.23E-03
7513	Bt.63820	*LOC517063*	similar to adenylate kinase isoenzyme 4, mitochondrial	1.23E-03
3868	Bt.49705	*PRDX6*	peroxiredoxin 6	1.24E-03
4707	Bt.46116	*MXRA7*	matrix-remodelling associated 7	1.25E-03
2672	Bt.6612	*DDX51*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 51	1.26E-03
6997	Bt.15697	*PER2*	period homolog 2 (Drosophila)	1.26E-03
4711	Bt.8545	*MRPL43*	mitochondrial ribosomal protein L43	1.26E-03
8090	Bt.14027	*DDX23*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 23	1.27E-03
10966	Bt.18408	*CCDC117*	coiled-coil domain containing 117	1.27E-03
874	Bt.22035	*C5orf14*	chromosome 5 open reading frame 14	1.28E-03
7621	Bt.9482	*LOC504245*	similar to Rho GTPase activating protein 17	1.29E-03
1806	Bt.51523	*TBC1D14*	TBC1 domain family, member 14	1.31E-03
1607	Bt.4371	*GADD45B*	growth arrest and DNA-damage-inducible, beta	1.31E-03
7122	Bt.22	*PPP1R8*	protein phosphatase 1, regulatory (inhibitor) subunit 8	1.32E-03
3543	Bt.4055	*BGLAP*	bone gamma-carboxyglutamate (gla) protein (osteocalcin)	1.32E-03

2425	Bt.64693	*Dnaja1*	DnaJ (Hsp40) homolog, subfamily A, member 1	1.32E-03
5988	Bt.43820	*RPL36*	ribosomal protein L36	1.33E-03
9334	Bt.21568	*FAM118B*	family with sequence similarity 118, member B	1.33E-03
5079	Bt.2642	*ZNF706*	zinc finger protein 706	1.34E-03
5169	Bt.1786	*YME1L1*	YME1-like 1 (S. cerevisiae)	1.34E-03
5967	Bt.1658	*DUSP1*	dual specificity phosphatase 1	1.34E-03
3900	Bt.11096	*KPNA4*	karyopherin alpha 4 (importin alpha 3)	1.35E-03
1923	Bt.64539	*QSCN6*	quiescin Q6	1.35E-03
10650	Bt.16032	*ANXA1*	annexin A1	1.36E-03
4566	Bt.6431	*PCAF*	p300/CBP-associated factor	1.38E-03
9151	Bt.77424	*Fubp1*	far upstream element (FUSE) binding protein 1	1.39E-03
7717	Bt.11282		Transcribed locus	1.39E-03
4322	Bt.13666	*ATP5S*	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit s (factor B)	1.40E-03
7815	Bt.6432	*FAM44B*	family with sequence similarity 44, member B	1.40E-03
11164	Bt.7658	*RPL38*	ribosomal protein L38	1.40E-03
1947	Bt.62762	*LOC124446*	hypothetical protein BC017488	1.41E-03
10742	Bt.89587	*USP37*	ubiquitin specific peptidase 37	1.42E-03
5381	Bt.6330		Transcribed locus	1.42E-03
7352	Bt.10332	*Cramp1l*	Crm, cramped-like (Drosophila)	1.43E-03
11069	Bt.15722	*FBL*	fibrillarin	1.43E-03
3598	Bt.49273	*POLR2F*	polymerase (RNA) II (DNA directed) polypeptide F	1.43E-03
5344	Bt.39475		Transcribed locus	1.44E-03
8654	Bt.13784	*TIMM10*	translocase of inner mitochondrial membrane 10 homolog (yeast)	1.44E-03
8474	Bt.89593	*PRPF4B*	PRP4 pre-mRNA processing factor 4 homolog B (yeast)	1.44E-03
7456	Bt.46948	*LRRFIP2*	leucine rich repeat (in FLII) interacting protein 2	1.45E-03
3189	Bt.44977	*SLC4A1AP*	solute carrier family 4 (anion exchanger), member 1, adaptor protein	1.45E-03
7204	Bt.65398		Transcribed locus	1.45E-03
3871	Bt.28081	*POGZ*	pogo transposable element with ZNF domain	1.46E-03
2415	Bt.46615	*MGC138967*	similar to CG10809-PA	1.48E-03
7833	Bt.9192	*MGAT4B*	mannosyl (alpha-1,3-)-glycoprotein beta-1,4-N-acetylglucosaminyltransferase, isozyme B	1.48E-03
7199	Bt.54456	*TFB2M*	transcription factor B2, mitochondrial	1.49E-03
595	Bt.6522	*DNAJB6*	DnaJ (Hsp40) homolog, subfamily B, member 6	1.50E-03
1686	Bt.33132	*TRIM24*	tripartite motif-containing 24	1.51E-03
5143	Bt.4106	*F3*	coagulation factor III (thromboplastin, tissue factor)	1.51E-03
2908	Bt.7221	*PPID*	peptidylprolyl isomerase D (cyclophilin D)	1.51E-03
11008	Bt.28567	*NFE2L1*	nuclear factor (erythroid-derived 2)-like 1	1.51E-03
5836	Bt.64649	*Znf313*	zinc finger protein 313	1.52E-03
4638	Bt.15915	*IPO11*	importin 11	1.53E-03
11499	Bt.64697	*LMNB2*	lamin B2	1.54E-03
9280	Bt.77424	*Fubp1*	far upstream element (FUSE) binding protein 1	1.54E-03

8933	Bt.30026		Full-length cDNA clone CS0DC015YK09 of Neuroblastoma Cot 25-normalized of Homo sapiens (human)	1.54E-03
2606	Bt.63071	*ARRDC4*	arrestin domain containing 4	1.54E-03
8455	Bt.52379	*QRICH1*	glutamine-rich 1	1.56E-03
10674	Bt.48127	*Tia1*	cytotoxic granule-associated RNA binding protein 1	1.57E-03
337	Bt.17282	*SNIP1*	Smad nuclear interacting protein 1	1.58E-03
761	Bt.49509	*LZTFL1*	leucine zipper transcription factor-like 1	1.59E-03
8867	Bt.64897	*IGFBP3*	insulin-like growth factor binding protein 3	1.60E-03
7528	Bt.49586	*NET1*	neuroepithelial cell transforming gene 1	1.60E-03
6758	Bt.26855	*LOC506746*	hypothetical LOC506746	1.61E-03
7193	Bt.3871	*LOC506318*	similar to VKORC1-like protein 1	1.61E-03
6140	Bt.27014	*FEZ2*	fasciculation and elongation protein zeta 2 (zygin II)	1.62E-03
6081	Bt.11128	*MGC140717*	similar to uracil-DNA glycosylase isoform UNG2	1.62E-03
10830	Bt.20683	*TP53INP2*	tumor protein p53 inducible nuclear protein 2	1.62E-03

Table S4.3: Differentially expressed genes between HHP-treated embryos and controls.
Spreadsheet 1: Up and down-regulated genes on 40 MPa treatment compared to control;
Spreadsheet 2: Up and down-regulated genes on 60 MPa treatment compared to control;
Spreadsheet 3: Up and down-regulated genes on 80 MPa treatment compared to control.

Table S4.3 Sheet 1. Genes significantly changed upon 40 MPa treatment

Probe	UniGene.ID	Gene name	P-value	Fold change (0.1 vs. 40 MPa)	Comment
1436	Bt.6451	*HIST1H1E*	4.92E-05	2.411047524	Up-regulated
1436	Bt.6451	*HIST1H1E*	0.000470673	2.411047524	Up-regulated
9020	Bt.4586	*GNG2*	1.81E-05	1.932158299	Up-regulated
9020	Bt.4586	*GNG2*	3.28E-05	1.932158299	Up-regulated
6659	Bt.3196	*STRA6*	0.000152416	1.868115792	Up-regulated
6978	Bt.22374	*MGC137396*	9.39E-06	1.746985206	Up-regulated
6978	Bt.22374	*MGC137396*	1.89E-05	1.746985206	Up-regulated
4927	Bt.55759	*DKFZp434K1815*	7.42E-05	1.739437184	Up-regulated
4927	Bt.55759	*DKFZp434K1815*	6.86E-08	1.739437184	Up-regulated
6288	Bt.13965	*CNTNAP1*	2.48E-07	1.722992416	Up-regulated
6288	Bt.13965	*CNTNAP1*	1.56E-06	1.722992416	Up-regulated
8747	Bt.23521	*DUSP26*	2.74E-05	1.680996282	Up-regulated
8747	Bt.23521	*DUSP26*	6.93E-05	1.680996282	Up-regulated
1090	Bt.7391	*APLP1*	3.84E-06	1.666490366	Up-regulated
1090	Bt.7391	*APLP1*	0.000618699	1.666490366	Up-regulated
5625	Bt.45080	*LRRC45*	2.36E-05	1.622618529	Up-regulated
5625	Bt.45080	*LRRC45*	0.000610941	1.622618529	Up-regulated
5262	Bt.8206	*SFRS7*	0.000724622	1.605936136	Up-regulated
5262	Bt.8206	*SFRS7*	0.000104799	1.605936136	Up-regulated
4884		*SEMA3G*	0.001075152	1.601008337	Up-regulated
2112	Bt.44470	*SATB1*	0.000705402	1.578782385	Up-regulated
10034	Bt.7363	*LOC789389*	0.000174731	1.565901939	Up-regulated
6830	Bt.52335	*IFT57*	0.000455475	1.562051072	Up-regulated
6830	Bt.52335	*IFT57*	0.000574792	1.562051072	Up-regulated
8416	Bt.21965	*ZNF358*	2.10E-05	1.549758396	Up-regulated
8416	Bt.21965	*ZNF358*	8.87E-05	1.549758396	Up-regulated
626	Bt.38424	*DST*	0.000199418	1.536786643	Up-regulated
8983	Bt.45485	*KLF9*	6.68E-05	1.533083305	Up-regulated
8983	Bt.45485	*KLF9*	9.23E-05	1.533083305	Up-regulated
3739	Bt.28081	*POGZ*	0.000189317	1.514137509	Up-regulated
3739	Bt.28081	*POGZ*	0.001460923	1.514137509	Up-regulated
1841	Bt.12138	*RP11-413M3.2*	6.48E-05	1.509031839	Up-regulated
3495	Bt.8279	*C3orf60*	1.30E-06	1.508682322	Up-regulated
3495	Bt.8279	*C3orf60*	1.85E-06	1.508682322	Up-regulated
4802	Bt.22766	*SART3*	0.000322672	1.50529845	Up-regulated

7986	Bt.66	*NDUFA3*	0.000138015	1.503232005	Up-regulated
7986	Bt.66	*NDUFA3*	0.000226394	1.503232005	Up-regulated
3885	Bt.51520	*C14orf94*	0.000123869	1.499872772	Up-regulated
3885	Bt.51520	*C14orf94*	0.000191948	1.499872772	Up-regulated
6997	Bt.15697	*PER2*	0.001261833	1.496187981	Up-regulated
2575	Bt.2041	*AP2A2*	3.05E-05	1.480948526	Up-regulated
2575	Bt.2041	*AP2A2*	1.71E-05	1.480948526	Up-regulated
5633	Bt.2568	*ADFP*	0.000157168	1.46820925	Up-regulated
5633	Bt.2568	*ADFP*	0.000202974	1.46820925	Up-regulated
393	Bt.58700	*C9orf58*	0.000166772	1.46778073	Up-regulated
393	Bt.58700	*C9orf58*	0.000647975	1.46778073	Up-regulated
2380	Bt.23564	*C6orf120*	1.02E-06	1.466207518	Up-regulated
2380	Bt.23564	*C6orf120*	2.66E-06	1.466207518	Up-regulated
9505	Bt.46839	*SNRPD2*	6.96E-05	1.458717013	Up-regulated
9505	Bt.46839	*SNRPD2*	8.10E-05	1.458717013	Up-regulated
9205	Bt.21568	*FAM118B*	0.001325776	1.450485835	Up-regulated
9205	Bt.21568	*FAM118B*	0.000132473	1.450485835	Up-regulated
8250	Bt.22175	*KCNT1*	7.92E-05	1.447039962	Up-regulated
8250	Bt.22175	*KCNT1*	0.000618956	1.447039962	Up-regulated
8877	Bt.48519	*RBPMS*	0.00063129	1.438466172	Up-regulated
8266	Bt.16440	*USP6NL*	0.000800332	1.433541215	Up-regulated
8266	Bt.16440	*USP6NL*	0.001046373	1.433541215	Up-regulated
5768	Bt.49597	*LOC528919*	0.000113481	1.431943194	Up-regulated
5768	Bt.49597	*LOC528919*	0.00051465	1.431943194	Up-regulated
4707	Bt.46116	*MXRA7*	0.001246439	1.429316172	Up-regulated
4263	Bt.23076	*C1orf93*	0.000764781	1.429229215	Up-regulated
4263	Bt.23076	*C1orf93*	0.000783807	1.429229215	Up-regulated
4237	Bt.87845	*COX5B*	2.00E-05	1.428369941	Up-regulated
4237	Bt.87845	*COX5B*	4.62E-05	1.428369941	Up-regulated
3263	Bt.37779	*BPHL*	0.000210985	1.415263686	Up-regulated
3263	Bt.37779	*BPHL*	0.000358399	1.415263686	Up-regulated
11275	Bt.2988	*GNG11*	0.000947206	1.413028655	Up-regulated
8432	Bt.5633	*TMEM129*	2.38E-05	1.410996343	Up-regulated
4243	Bt.6710	*LOC440330*	4.50E-05	1.410510514	Up-regulated
4243	Bt.6710	*LOC440330*	7.53E-05	1.410510514	Up-regulated
3585	Bt.26997	*Ptpmt1*	0.000530902	1.410002052	Up-regulated
3045	Bt.36226	*SMNDC1*	0.001098454	1.409950628	Up-regulated
11761	Bt.17953	*FOXP1*	0.000956226	1.406683227	Up-regulated
11761	Bt.17953	*FOXP1*	0.001230279	1.406683227	Up-regulated
3414	Bt.61384	*UNC84B*	0.000416827	1.405923185	Up-regulated
3414	Bt.61384	*UNC84B*	0.00011637	1.405923185	Up-regulated
3543	Bt.4055	*BGLAP*	0.001321025	1.401462282	Up-regulated
1886	Bt.76825	*LOC789371*	0.000833992	1.400577385	Up-regulated

884	Bt.49509	*LZTFL1*	0.000919334	1.399418351	Up-regulated
884	Bt.49509	*LZTFL1*	0.001590158	1.399418351	Up-regulated
10674	Bt.48127	*Tia1*	0.001569778	1.393954831	Up-regulated
4986	Bt.89574	*KIAA0182*	0.00015045	1.390348374	Up-regulated
4986	Bt.89574	*KIAA0182*	0.000987792	1.390348374	Up-regulated
1902	Bt.12743	*ABCD4*	0.00063545	1.387069515	Up-regulated
11164	Bt.7658	*RPL38*	0.001398942	1.382080773	Up-regulated
967	Bt.73261	*LOC513955*	6.85E-05	1.381148907	Up-regulated
2887	Bt.61617	*UHRF2*	0.000904125	1.377190598	Up-regulated
2887	Bt.61617	*UHRF2*	0.000439983	1.377190598	Up-regulated
6124	Bt.43820	*RPL36*	0.000760297	1.376458819	Up-regulated
6124	Bt.43820	*RPL36*	0.001325246	1.376458819	Up-regulated
6385	Bt.9092	*POLR2G*	0.000122794	1.373601453	Up-regulated
6385	Bt.9092	*POLR2G*	0.000220579	1.373601453	Up-regulated
5945	Bt.26976	*COX11*	0.000101741	1.370818345	Up-regulated
5945	Bt.26976	*COX11*	0.00020081	1.370818345	Up-regulated
6155	Bt.7804	*GYG1*	0.000301114	1.368438996	Up-regulated
6155	Bt.7804	*GYG1*	0.000905921	1.368438996	Up-regulated
7193	Bt.3871	*LOC506318*	0.00160686	1.367978974	Up-regulated
4345	Bt.17589	*DMTF1*	0.000879507	1.367584324	Up-regulated
9896	Bt.58080	*ZNF3*	1.85E-06	1.36433034	Up-regulated
9896	Bt.58080	*ZNF3*	7.52E-06	1.36433034	Up-regulated
2807	Bt.64366	*NDUFC2*	0.00024111	1.360154624	Up-regulated
6501	Bt.48788	*RAB30*	0.000717957	1.359494169	Up-regulated
7223	Bt.500	*HK1*	0.000108524	1.352377959	Up-regulated
7223	Bt.500	*HK1*	0.000187069	1.352377959	Up-regulated
4532	Bt.4949	*SF3B5*	0.000553053	1.342608753	Up-regulated
4532	Bt.4949	*SF3B5*	0.000798069	1.342608753	Up-regulated
1536	Bt.66354	*POLR2L*	6.22E-06	1.342343172	Up-regulated
1536	Bt.66354	*POLR2L*	1.19E-05	1.342343172	Up-regulated
1019	Bt.9195	*SIVA1*	0.000493087	1.339920698	Up-regulated
8159			0.000169244	1.337835553	Up-regulated
8159			0.000509444	1.337835553	Up-regulated
8159	Bt.30026		0.001543259	1.337835553	Up-regulated
8159	Bt.65788		0.00021414	1.337835553	Up-regulated
8159	Bt.732		0.000471878	1.337835553	Up-regulated
8159	Bt.6330		0.001422256	1.337835553	Up-regulated
9208	Bt.26958	*LOC788996*	0.000215607	1.33553659	Up-regulated
8654	Bt.13784	*TIMM10*	0.001437679	1.330900944	Up-regulated
8583	Bt.52379	*QRICH1*	0.001559509	1.330801741	Up-regulated
8583	Bt.52379	*QRICH1*	0.000137079	1.330801741	Up-regulated
4008	Bt.26835	*RANGNRF*	0.001169618	1.329837495	Up-regulated
4008	Bt.26835	*RANGNRF*	0.000511388	1.329837495	Up-regulated

4656	Bt.32544	*NVL*	0.000128076	1.326912959	Up-regulated
4400	Bt.49192	*RPL15*	0.000657603	1.321596299	Up-regulated
4400	Bt.49192	*RPL15*	0.000774218	1.321596299	Up-regulated
7122	Bt.22	*PPP1R8*	0.00131595	1.32005256	Up-regulated
2877	Bt.30434	*SFRS9*	0.00091765	1.318511686	Up-regulated
2569	Bt.7562	*FLJ41352*	3.66E-05	1.314844992	Up-regulated
2569	Bt.7562	*FLJ41352*	0.00076392	1.314844992	Up-regulated
4001	Bt.26739	*TOMM40L*	0.000557284	1.299224292	Up-regulated
2357	Bt.5184	*Ankrd40*	2.39E-06	1.298987503	Up-regulated
2357	Bt.5184	*Ankrd40*	0.000435507	1.298987503	Up-regulated
5328	Bt.55867	*HSPC152*	2.62E-05	1.298459765	Up-regulated
5328	Bt.55867	*HSPC152*	9.14E-05	1.298459765	Up-regulated
11668	Bt.49238	*HPRT1*	0.001115084	1.297719755	Up-regulated
11668	Bt.49238	*HPRT1*	0.00120866	1.297719755	Up-regulated
4505	Bt.15915	*IPO11*	0.000488539	1.297146056	Up-regulated
4505	Bt.15915	*IPO11*	0.001528633	1.297146056	Up-regulated
11499	Bt.64697	*LMNB2*	0.001540602	1.294782438	Up-regulated
6339	Bt.44021	*WDR8*	0.000830074	1.2883317	Up-regulated
4349	Bt.66386	*WRNIP1*	4.59E-05	1.285299393	Up-regulated
4349	Bt.66386	*WRNIP1*	6.48E-05	1.285299393	Up-regulated
5987	Bt.13810	*U2AF1*	0.000815782	1.285213607	Up-regulated
2471	Bt.49501	*C15orf40*	0.000680521	1.27986987	Up-regulated
4931	Bt.27370	*HEATR3*	0.000707341	1.278728054	Up-regulated
3472	Bt.49273	*POLR2F*	0.001433416	1.278467132	Up-regulated
3472	Bt.49273	*POLR2F*	0.00090677	1.278467132	Up-regulated
11028	Bt.30639	*CNOT2*	0.000466637	1.2781415	Up-regulated
5836	Bt.64649	*Znf313*	0.001516718	1.276434672	Up-regulated
2339	Bt.52556	*RCP9*	6.21E-05	1.263776461	Up-regulated
2339	Bt.52556	*RCP9*	0.000280575	1.263776461	Up-regulated
2654	Bt.4298	*GPR172B*	0.000194199	1.263467859	Up-regulated
2654	Bt.4298	*GPR172B*	0.000251915	1.263467859	Up-regulated
7621	Bt.9482	*LOC504245*	0.001291534	1.262970177	Up-regulated
1858	Bt.62530	*ITPK1*	0.000482942	1.262815915	Up-regulated
1858	Bt.62530	*ITPK1*	0.001071091	1.262815915	Up-regulated
2415	Bt.46615	*MGC138967*	0.001483978	1.255239148	Up-regulated
9907	Bt.59214	*FGD5*	8.01E-05	1.251449378	Up-regulated
2193	Bt.52796	*COPS6*	0.000225488	1.244506582	Up-regulated
10009	Bt.48940	*PPAP2B*	1.96E-06	1.243042178	Up-regulated
6140	Bt.27014	*FEZ2*	0.001617183	1.241905343	Up-regulated
2709	Bt.21070	*MRPL10*	0.000769143	1.241342612	Up-regulated
5727	Bt.70	*NDUFB10*	0.000260029	1.239802419	Up-regulated
5727	Bt.70	*NDUFB10*	0.000771141	1.239802419	Up-regulated
7556	Bt.63484	*SURF2*	0.000466669	1.23945914	Up-regulated

10232	Bt.1858	*C20orf108*	0.000616414	1.238627466	Up-regulated
10232	Bt.1858	*C20orf108*	0.000725454	1.238627466	Up-regulated
2992	Bt.18512	*ASF1A*	0.000880779	1.237140595	Up-regulated
1640	Bt.49587	*GPI*	0.000206942	1.231214321	Up-regulated
1640	Bt.49587	*GPI*	0.000651686	1.231214321	Up-regulated
2555	Bt.5528	*SLC7A5*	0.001005994	1.224638311	Up-regulated
2555	Bt.5528	*SLC7A5*	0.001009624	1.224638311	Up-regulated
2865	Bt.89629	*EIF4B*	0.001144637	1.223356364	Up-regulated
2672	Bt.6612	*DDX51*	0.001257178	1.217956704	Up-regulated
5773	Bt.38862	*RALA*	0.000703113	1.215908003	Up-regulated
4914	Bt.64707	*LOC146346*	0.000776269	1.196658362	Up-regulated
4144	Bt.49595	*TMEM111*	8.44E-05	1.185066377	Up-regulated
3794	Bt.44982	*CHRNA10*	0.001086009	1.178015959	Up-regulated
8460	Bt.5530	*DHRS3*	0.000816115	1.175738577	Up-regulated
4834	Bt.8545	*MRPL43*	0.001262114	1.173020122	Up-regulated
4834	Bt.8545	*MRPL43*	9.09E-05	1.173020122	Up-regulated
5186	Bt.57711	*NDUFB8*	0.000654609	1.150863212	Up-regulated
1689	Bt.39564	*LOC506268*	0.000332691	1.131055131	Up-regulated
5271	Bt.4908	*RBBP7*	0.000265911	0.857163239	Down-regulated
289	Bt.21926	*LOC508142*	0.000881912	0.828800813	Down-regulated
6897	Bt.21510	*GPR107*	0.001032206	0.820087446	Down-regulated
11849	Bt.9469	*DNAJC10*	0.000759213	0.818915235	Down-regulated
5040	Bt.1786	*YME1L1*	0.00133787	0.778151768	Down-regulated
5040	Bt.1786	*YME1L1*	0.000619741	0.778151768	Down-regulated
288	Bt.28094	*CASP7*	1.68E-06	0.773399708	Down-regulated
288	Bt.28094	*CASP7*	0.000112935	0.773399708	Down-regulated
123	Bt.64291	*H3F3A*	0.000158185	0.768681681	Down-regulated
123	Bt.64291	*H3F3A*	0.000471416	0.768681681	Down-regulated
10194	Bt.62845	*ASAH1*	0.000518233	0.766830961	Down-regulated
10194	Bt.62845	*ASAH1*	0.000996226	0.766830961	Down-regulated
5766	Bt.65363	*C16orf69*	2.56E-05	0.765511665	Down-regulated
5766	Bt.65363	*C16orf69*	0.000270629	0.765511665	Down-regulated
3518	Bt.28502	*FAM83D*	0.000417802	0.763218429	Down-regulated
3518	Bt.28502	*FAM83D*	0.000474399	0.763218429	Down-regulated
3189	Bt.44977	*SLC4A1AP*	0.00145282	0.761790522	Down-regulated
265	Bt.20277	*TOP2A*	0.00078607	0.758350021	Down-regulated
779	Bt.90218	*TES*	0.000610689	0.755475675	Down-regulated
10650	Bt.16032	*ANXA1*	0.001356986	0.751616221	Down-regulated
874	Bt.22035	*C5orf14*	0.001283155	0.749024192	Down-regulated
9841	Bt.87389	*GAPDH*	0.000629315	0.74867001	Down-regulated
9841	Bt.87389	*GAPDH*	0.001215282	0.74867001	Down-regulated
2719	Bt.75092	*DIDO1*	4.49E-07	0.747525798	Down-regulated
8159	Bt.89557		0.000463052	0.747476024	Down-regulated

8159			0.000946699	0.747476024	Down-regulated
3140	Bt.56130	*SIAHBP1*	0.001110559	0.739135751	Down-regulated
3140	Bt.56130	*SIAHBP1*	0.001220173	0.739135751	Down-regulated
1897	Bt.12732	*NYREN18*	3.57E-05	0.738596777	Down-regulated
1897	Bt.12732	*NYREN18*	0.00011322	0.738596777	Down-regulated
893	Bt.59331	*UAP1*	3.55E-05	0.730191033	Down-regulated
893	Bt.59331	*UAP1*	6.66E-05	0.730191033	Down-regulated
1782	Bt.53344	*ARHGAP21*	0.000850676	0.724781482	Down-regulated
4305	Bt.89659	*FKBP9L*	3.78E-05	0.721420708	Down-regulated
4305	Bt.89659	*FKBP9L*	2.02E-05	0.721420708	Down-regulated
10996	Bt.22526	*HSPB8*	7.83E-05	0.719942034	Down-regulated
10996	Bt.22526	*HSPB8*	0.001019679	0.719942034	Down-regulated
7456	Bt.46948	*LRRFIP2*	0.001449248	0.715436627	Down-regulated
11731	Bt.22575	*TMEM167*	0.000368237	0.709895119	Down-regulated
1116	Bt.25186	*BCAT1*	0.001013464	0.70421284	Down-regulated
10961	Bt.20683	*TP53INP2*	0.000874769	0.703614152	Down-regulated
10961	Bt.20683	*TP53INP2*	0.001621616	0.703614152	Down-regulated
7909	Bt.64563	*RFFL*	2.54E-05	0.701882852	Down-regulated
7909	Bt.64563	*RFFL*	0.00069605	0.701882852	Down-regulated
12100	Bt.30953	*DOCK1*	0.000300917	0.701335885	Down-regulated
5143	Bt.4106	*F3*	0.00150891	0.693874794	Down-regulated
595	Bt.6522	*DNAJB6*	0.001496388	0.687365516	Down-regulated
3177	Bt.48951	*ASB11*	1.21E-05	0.682168948	Down-regulated
3177	Bt.48951	*ASB11*	4.73E-05	0.682168948	Down-regulated
8529	Bt.7930	*NUCB1*	0.001008492	0.680744745	Down-regulated
7528	Bt.49586	*NET1*	0.001601412	0.674008596	Down-regulated
538	Bt.13737	*GALK2*	0.0005438	0.673522651	Down-regulated
5355	Bt.269	*ATP2C1*	5.96E-05	0.672029196	Down-regulated
5355	Bt.269	*ATP2C1*	0.000228997	0.672029196	Down-regulated
904	Bt.7418	*ANKH*	4.40E-05	0.656166025	Down-regulated
904	Bt.7418	*ANKH*	6.37E-05	0.656166025	Down-regulated
942	Bt.6650	*GGH*	0.000841897	0.655009257	Down-regulated
4274	Bt.11909	*DUSP6*	0.000701201	0.652235376	Down-regulated
4274	Bt.11909	*DUSP6*	0.000889526	0.652235376	Down-regulated
5332	Bt.12314	*PFKFB3*	3.97E-05	0.637898738	Down-regulated
5332	Bt.12314	*PFKFB3*	6.60E-05	0.637898738	Down-regulated
747	Bt.28249	*SEC24A*	0.000173964	0.632782086	Down-regulated
1290	Bt.63116	*AKR1B1*	0.000550475	0.627096985	Down-regulated
1290	Bt.63116	*AKR1B1*	0.000678361	0.627096985	Down-regulated
1968	Bt.53760	*LOC728216*	0.000104525	0.625218368	Down-regulated
1968	Bt.53760	*LOC728216*	0.000116805	0.625218368	Down-regulated
8644	Bt.8738	*HSPH1*	5.51E-05	0.624614079	Down-regulated
8644	Bt.8738	*HSPH1*	9.29E-05	0.624614079	Down-regulated

281	Bt.9783	*DDX3X*	0.000421378	0.618222185	Down-regulated
281	Bt.9783	*DDX3X*	0.000835839	0.618222185	Down-regulated
6779	Bt.90151	*LOC100008589*	8.45E-07	0.601998463	Down-regulated
6779	Bt.90151	*LOC100008589*	2.96E-06	0.601998463	Down-regulated
2986	Bt.27527	*STXBP3*	0.000277868	0.597526182	Down-regulated
2986	Bt.27527	*STXBP3*	0.001176051	0.597526182	Down-regulated
1477	Bt.12399	*NEU1*	1.84E-05	0.582769853	Down-regulated
1477	Bt.12399	*NEU1*	5.03E-05	0.582769853	Down-regulated
7968	Bt.49157	*APOA1*	0.000973083	0.565725636	Down-regulated
1794	Bt.64539	*QSCN6*	0.000535468	0.556775247	Down-regulated
1794	Bt.64539	*QSCN6*	0.001353419	0.556775247	Down-regulated
2402	Bt.7873	*BCAM*	6.99E-06	0.545012548	Down-regulated
2402	Bt.7873	*BCAM*	1.10E-05	0.545012548	Down-regulated
9586	Bt.88378	*KIAA1303*	1.52E-06	0.534767689	Down-regulated
9586	Bt.88378	*KIAA1303*	6.94E-06	0.534767689	Down-regulated

Table S4.3 Sheet 2. Genes significantly changed upon 60 MPa treatment

Probe	UniGene.ID	Gene name	P-value	Fold change (0.1 vs. 60 MPa)	comment
7199	Bt.54456	*TFB2M*	0.001489348	2.689638309	Up-regulated
4884		*SEMA3G*	0.001075152	1.838760276	Up-regulated
4927	Bt.55759	*DKFZp434K1815*	7.42E-05	1.755178583	Up-regulated
4927	Bt.55759	*DKFZp434K1815*	6.86E-08	1.755178583	Up-regulated
9020	Bt.4586	*GNG2*	1.81E-05	1.655132083	Up-regulated
9020	Bt.4586	*GNG2*	3.28E-05	1.655132083	Up-regulated
5966	Bt.5519	*TncRNA*	0.000328087	1.543041709	Up-regulated
3543	Bt.4055	*BGLAP*	0.001321025	1.523434989	Up-regulated
5262	Bt.8206	*SFRS7*	0.000104799	1.509690386	Up-regulated
8128	Bt.33726	*NPY*	0.000388758	1.498029214	Up-regulated
4243	Bt.6710	*LOC440330*	4.50E-05	1.48970907	Up-regulated
4243	Bt.6710	*LOC440330*	7.53E-05	1.48970907	Up-regulated
3585	Bt.26997	*Ptpmt1*	0.000530902	1.483292753	Up-regulated
6997	Bt.15697	*PER2*	0.001261833	1.469871756	Up-regulated
11164	Bt.7658	*RPL38*	0.001398942	1.447273881	Up-regulated
2167	Bt.20080	*WSB1*	0.000214607	1.426672599	Up-regulated
2167	Bt.20080	*WSB1*	0.000831809	1.426672599	Up-regulated
1401	Bt.3199	*RUSC1*	0.000221096	1.424924532	Up-regulated
9208	Bt.26958	*LOC788996*	0.000280312	1.410670624	Up-regulated
9208	Bt.26958	*LOC788996*	0.000215607	1.410670624	Up-regulated
4237	Bt.87845	*COX5B*	2.00E-05	1.406990258	Up-regulated
4237	Bt.87845	*COX5B*	4.62E-05	1.406990258	Up-regulated
4345	Bt.17589	*DMTF1*	0.000879507	1.402102734	Up-regulated

9205	Bt.21568	*FAM118B*	0.001325776	1.380979557	Up-regulated
9205	Bt.21568	*FAM118B*	0.000132473	1.380979557	Up-regulated
2654	Bt.4298	*GPR172B*	0.000194199	1.355529622	Up-regulated
2654	Bt.4298	*GPR172B*	0.000251915	1.355529622	Up-regulated
626	Bt.38424	*DST*	0.000199418	1.355261898	Up-regulated
6288	Bt.13965	*CNTNAP1*	2.48E-07	1.345266634	Up-regulated
6288	Bt.13965	*CNTNAP1*	1.56E-06	1.345266634	Up-regulated
8654	Bt.13784	*TIMM10*	0.001437679	1.342819566	Up-regulated
5	Bt.29924	*MDM4*	0.000824537	1.33956028	Up-regulated
2877	Bt.30434	*SFRS9*	0.00091765	1.339324809	Up-regulated
1895	Bt.42858	*ABCC5*	1.02E-05	1.336991574	Up-regulated
1895	Bt.42858	*ABCC5*	2.23E-05	1.336991574	Up-regulated
9505	Bt.46839	*SNRPD2*	6.96E-05	1.335226279	Up-regulated
9505	Bt.46839	*SNRPD2*	8.10E-05	1.335226279	Up-regulated
1090	Bt.7391	*APLP1*	3.84E-06	1.328202593	Up-regulated
1090	Bt.7391	*APLP1*	0.000618699	1.328202593	Up-regulated
8266	Bt.16440	*USP6NL*	0.000800332	1.327893163	Up-regulated
8266	Bt.16440	*USP6NL*	0.001046373	1.327893163	Up-regulated
7986	Bt.66	*NDUFA3*	0.000138015	1.306412347	Up-regulated
7986	Bt.66	*NDUFA3*	0.000226394	1.306412347	Up-regulated
1536	Bt.66354	*POLR2L*	6.22E-06	1.305545697	Up-regulated
1536	Bt.66354	*POLR2L*	1.19E-05	1.305545697	Up-regulated
7061	Bt.42818	*POLR3B*	0.000405958	1.304098893	Up-regulated
7223	Bt.500	*HK1*	0.000108524	1.296405031	Up-regulated
7223	Bt.500	*HK1*	0.000187069	1.296405031	Up-regulated
5328	Bt.55867	*HSPC152*	2.62E-05	1.294876918	Up-regulated
5328	Bt.55867	*HSPC152*	9.14E-05	1.294876918	Up-regulated
564	Bt.28460	*SNRPD3*	0.000971154	1.294689883	Up-regulated
8474	Bt.89593	*PRPF4B*	0.001438541	1.294309816	Up-regulated
6217	Bt.11128	*MGC140717*	0.001127116	1.292296325	Up-regulated
6217	Bt.11128	*MGC140717*	0.00162156	1.292296325	Up-regulated
4263	Bt.23076	*C1orf93*	0.000764781	1.286025986	Up-regulated
4263	Bt.23076	*C1orf93*	0.000783807	1.286025986	Up-regulated
2380	Bt.23564	*C6orf120*	1.02E-06	1.286013992	Up-regulated
2380	Bt.23564	*C6orf120*	2.66E-06	1.286013992	Up-regulated
2887	Bt.61617	*UHRF2*	0.000439983	1.279022933	Up-regulated
6124	Bt.43820	*RPL36*	0.000760297	1.275187577	Up-regulated
6124	Bt.43820	*RPL36*	0.001325246	1.275187577	Up-regulated
2471	Bt.49501	*C15orf40*	0.000680521	1.268427709	Up-regulated
89	Bt.32543	*IPO8*	0.000191059	1.26532392	Up-regulated
5768	Bt.49597	*LOC528919*	0.000113481	1.255821679	Up-regulated
5768	Bt.49597	*LOC528919*	0.00051465	1.255821679	Up-regulated
6385	Bt.9092	*POLR2G*	0.000122794	1.248232894	Up-regulated

6385	Bt.9092	*POLR2G*	0.000220579	1.248232894	Up-regulated
2569	Bt.7562	*FLJ41352*	3.66E-05	1.240282952	Up-regulated
2569	Bt.7562	*FLJ41352*	0.00076392	1.240282952	Up-regulated
3495	Bt.8279	*C3orf60*	1.30E-06	1.233969501	Up-regulated
3495	Bt.8279	*C3orf60*	1.85E-06	1.233969501	Up-regulated
4008	Bt.26835	*RANGNRF*	0.000511388	1.206818212	Up-regulated
4914	Bt.64707	*LOC146346*	0.000776269	1.202220304	Up-regulated
3472	Bt.49273	*POLR2F*	0.00090677	1.195577264	Up-regulated
4952	Bt.2642	*ZNF706*	0.000593589	1.184466128	Up-regulated
6140	Bt.27014	*FEZ2*	0.001617183	1.182520808	Up-regulated
8583	Bt.52379	*QRICH1*	0.000137079	1.180039699	Up-regulated
5186	Bt.57711	*NDUFB8*	0.000654609	1.162510081	Up-regulated
2865	Bt.89629	*EIF4B*	0.001144637	1.160682533	Up-regulated
4834	Bt.8545	*MRPL43*	9.09E-05	1.157716265	Up-regulated
5945	Bt.26976	*COX11*	0.00020081	1.155350621	Up-regulated
8159	Bt.65788		0.00021414	0.977496298	Up-regulated
8159	Bt.732		0.000471878	0.977496298	Up-regulated
8159	Bt.6330		0.001422256	0.977496298	Up-regulated
4305	Bt.89659	*FKBP9L*	2.02E-05	0.858161966	Down-regulated
5271	Bt.4908	*RBBP7*	0.000265911	0.84164316	Down-regulated
5040	Bt.1786	*YME1L1*	0.00133787	0.826213324	Down-regulated
5040	Bt.1786	*YME1L1*	0.000619741	0.826213324	Down-regulated
3189	Bt.44977	*SLC4A1AP*	0.00145282	0.811971933	Down-regulated
7456	Bt.46948	*LRRFIP2*	0.001449248	0.799930229	Down-regulated
9071	Bt.66625	*CNOT4*	0.000437896	0.799324154	Down-regulated
1607	Bt.4371	*GADD45B*	0.001314898	0.795554642	Down-regulated
3140	Bt.56130	*SIAHBP1*	0.001110559	0.783701753	Down-regulated
3140	Bt.56130	*SIAHBP1*	0.001220173	0.783701753	Down-regulated
3701	Bt.24435	*SYNCRIP*	0.000975088	0.77774974	Down-regulated
3701	Bt.24435	*SYNCRIP*	0.000980781	0.77774974	Down-regulated
2287	Bt.27376	*Bag4*	4.25E-07	0.774820777	Down-regulated
2287	Bt.27376	*Bag4*	0.000106309	0.774820777	Down-regulated
10681	Bt.59095	*C9orf97*	0.000647743	0.773006175	Down-regulated
8867	Bt.64897	*IGFBP3*	0.001597096	0.764593659	Down-regulated
10996	Bt.22526	*HSPB8*	7.83E-05	0.755003559	Down-regulated
10996	Bt.22526	*HSPB8*	0.001019679	0.755003559	Down-regulated
3177	Bt.48951	*ASB11*	1.21E-05	0.749433209	Down-regulated
3177	Bt.48951	*ASB11*	4.73E-05	0.749433209	Down-regulated
11849	Bt.9469	*DNAJC10*	0.000759213	0.744953014	Down-regulated
5355	Bt.269	*ATP2C1*	5.96E-05	0.742417945	Down-regulated
5355	Bt.269	*ATP2C1*	0.000228997	0.742417945	Down-regulated
2027	Bt.10259	*PLK2*	0.00014694	0.732118669	Down-regulated
8644	Bt.8738	*HSPH1*	5.51E-05	0.725798249	Down-regulated

8644	Bt.8738	*HSPH1*	9.29E-05	0.725798249	Down-regulated
2402	Bt.7873	*BCAM*	6.99E-06	0.686258068	Down-regulated
2402	Bt.7873	*BCAM*	1.10E-05	0.686258068	Down-regulated
7528	Bt.49586	*NET1*	0.001601412	0.685970046	Down-regulated
11261	Bt.9525	*RNF113A*	0.000258404	0.67344634	Down-regulated
11261	Bt.9525	*RNF113A*	0.000628905	0.67344634	Down-regulated
2719	Bt.75092	*DIDO1*	0.000364325	0.63187093	Down-regulated
2719	Bt.75092	*DIDO1*	4.49E-07	0.63187093	Down-regulated
9586	Bt.88378	*KIAA1303*	1.52E-06	0.616588017	Down-regulated
9586	Bt.88378	*KIAA1303*	6.94E-06	0.616588017	Down-regulated
5154	Bt.52786	*LOC100008588*	0.000102515	0.600036918	Down-regulated
5154	Bt.52786	*LOC100008588*	0.000160742	0.600036918	Down-regulated
3671	Bt.27129	*SLC39A9*	0.001181192	0.588360292	Down-regulated
8731	Bt.27784	*PPP1R10*	4.21E-07	0.588026867	Down-regulated
8731	Bt.27784	*PPP1R10*	1.13E-06	0.588026867	Down-regulated
6779	Bt.90151	*LOC100008589*	8.45E-07	0.486632109	Down-regulated
6779	Bt.90151	*LOC100008589*	2.96E-06	0.486632109	Down-regulated
6779	Bt.89284	*LOC100008589*	0.000106092	0.486632109	Down-regulated
6779	Bt.89284	*LOC100008589*	0.000208105	0.486632109	Down-regulated

Table S4.3 Sheet 3. Genes significantly changed upon 80 MPa treatment

Probe	UniGene ID	Gene name	P-value	Fold change (0.1 vs. 80 MPa)	Comment
1436	Bt.6451	*HIST1H1E*	4.92E-05	12.74953624	Up-regulated
1436	Bt.6451	*HIST1H1E*	0.000470673	12.74953624	Up-regulated
42	Bt.34521	*RBM12*	0.000452853	2.703349	Up-regulated
8416	Bt.21965	*ZNF358*	2.10E-05	1.55765483	Up-regulated
8416	Bt.21965	*ZNF358*	8.87E-05	1.55765483	Up-regulated
4001	Bt.26739	*TOMM40L*	0.000557284	1.496010949	Up-regulated
8159	Bt.89557		1.96E-05	1.477117365	Up-regulated
7986	Bt.66	*NDUFA3*	0.000138015	1.406306395	Up-regulated
7986	Bt.66	*NDUFA3*	0.000226394	1.406306395	Up-regulated
6288	Bt.13965	*CNTNAP1*	2.48E-07	1.398862236	Up-regulated
6288	Bt.13965	*CNTNAP1*	1.56E-06	1.398862236	Up-regulated
4927	Bt.55759	*DKFZp434K1815*	6.86E-08	1.374371341	Up-regulated
8877	Bt.48519	*RBPMS*	0.00063129	1.373950549	Up-regulated
2634	Bt.90254	*HDAC8*	0.000981392	1.355704704	Up-regulated
9205	Bt.21568	*FAM118B*	0.000132473	1.29067183	Up-regulated
2415	Bt.46615	*MGC138967*	0.001483978	1.288080872	Up-regulated
7223	Bt.500	*HK1*	0.000108524	1.275324741	Up-regulated
7223	Bt.500	*HK1*	0.000187069	1.275324741	Up-regulated
1841	Bt.12138	*RP11-413M3.2*	6.48E-05	1.268377263	Up-regulated

884	Bt.49509	*LZTFL1*	0.000919334	1.267012976	Up-regulated
884	Bt.49509	*LZTFL1*	0.001590158	1.267012976	Up-regulated
2575	Bt.2041	*AP2A2*	1.71E-05	1.259958213	Up-regulated
3414	Bt.61384	*UNC84B*	0.00011637	1.214418026	Up-regulated
2865	Bt.89629	*EIF4B*	0.001144637	1.190531948	Up-regulated
4914	Bt.64707	*LOC146346*	0.000776269	1.168484456	Up-regulated
5040	Bt.1786	*YME1L1*	0.000619741	0.81416973	Down-regulated
3865	Bt.52937	*hCG_20417*	0.00010856	0.810226455	Down-regulated
3865	Bt.52937	*hCG_20417*	0.000139517	0.810226455	Down-regulated
4305	Bt.89659	*FKBP9L*	3.78E-05	0.801611975	Down-regulated
4305	Bt.89659	*FKBP9L*	2.02E-05	0.801611975	Down-regulated
6779	Bt.89284	*LOC100008589*	0.000106092	0.793721372	Down-regulated
6779	Bt.89284	*LOC100008589*	0.000208105	0.793721372	Down-regulated
9829	Bt.27267	*TMEM4*	0.000150732	0.791384829	Down-regulated
6758	Bt.26855	*LOC506746*	0.001605013	0.789528482	Down-regulated
1685	Bt.77469	*KCTD20*	6.84E-05	0.786367218	Down-regulated
1685	Bt.77469	*KCTD20*	0.000311099	0.786367218	Down-regulated
1402	Bt.4279	*SHC1*	0.000135301	0.784658922	Down-regulated
1402	Bt.4279	*SHC1*	0.00040895	0.784658922	Down-regulated
1782	Bt.53344	*ARHGAP21*	0.000850676	0.783994648	Down-regulated
9071	Bt.66625	*CNOT4*	0.000437896	0.769761713	Down-regulated
3140	Bt.56130	*SIAHBP1*	0.001110559	0.769103488	Down-regulated
3140	Bt.56130	*SIAHBP1*	0.001220173	0.769103488	Down-regulated
5840	Bt.1658	*DUSP1*	0.000276438	0.769080319	Down-regulated
5840	Bt.1658	*DUSP1*	0.001339091	0.769080319	Down-regulated
3701	Bt.24435	*SYNCRIP*	0.000975088	0.766606894	Down-regulated
3701	Bt.24435	*SYNCRIP*	0.000980781	0.766606894	Down-regulated
5271	Bt.4908	*RBBP7*	0.000265911	0.765349306	Down-regulated
3177	Bt.48951	*ASB11*	1.21E-05	0.762018911	Down-regulated
3177	Bt.48951	*ASB11*	4.73E-05	0.762018911	Down-regulated
337	Bt.17282	*SNIP1*	0.001581269	0.761885147	Down-regulated
7456	Bt.46948	*LRRFIP2*	0.001449248	0.757525629	Down-regulated
11849	Bt.9469	*DNAJC10*	0.000759213	0.754942748	Down-regulated
3089	Bt.22538	*FZD8*	9.71E-05	0.739475424	Down-regulated
10194	Bt.62845	*ASAH1*	0.000518233	0.733242501	Down-regulated
10194	Bt.62845	*ASAH1*	0.000996226	0.733242501	Down-regulated
3757	Bt.2005	*LSM1*	1.50E-07	0.733206298	Down-regulated
3757	Bt.2005	*LSM1*	9.51E-06	0.733206298	Down-regulated
10834	Bt.18408	*CCDC117*	0.001065327	0.725922458	Down-regulated
10834	Bt.18408	*CCDC117*	0.001272419	0.725922458	Down-regulated
7060	Bt.11280	*BANP*	0.00012219	0.72338105	Down-regulated
7060	Bt.11280	*BANP*	0.00013219	0.72338105	Down-regulated
10650	Bt.16032	*ANXA1*	0.001356986	0.722304465	Down-regulated

6408	Bt.52209	*RAD54B*	0.000316422	0.714724398	Down-regulated
9841	Bt.87389	*GAPDH*	0.000629315	0.714170291	Down-regulated
9841	Bt.87389	*GAPDH*	0.001215282	0.714170291	Down-regulated
4027	Bt.11096	*KPNA4*	0.00070383	0.710854719	Down-regulated
4027	Bt.11096	*KPNA4*	0.001347272	0.710854719	Down-regulated
1897	Bt.12732	*NYREN18*	3.57E-05	0.708285051	Down-regulated
1897	Bt.12732	*NYREN18*	0.00011322	0.708285051	Down-regulated
1607	Bt.4371	*GADD45B*	0.001314898	0.704830784	Down-regulated
4274	Bt.11909	*DUSP6*	0.000701201	0.704226522	Down-regulated
4274	Bt.11909	*DUSP6*	0.000889526	0.704226522	Down-regulated
1290	Bt.63116	*AKR1B1*	0.000550475	0.697691491	Down-regulated
1290	Bt.63116	*AKR1B1*	0.000678361	0.697691491	Down-regulated
893	Bt.59331	*UAP1*	3.55E-05	0.693202226	Down-regulated
893	Bt.59331	*UAP1*	6.66E-05	0.693202226	Down-regulated
1109	Bt.17683	*DBF4*	0.000560078	0.693128763	Down-regulated
3671	Bt.27129	*SLC39A9*	0.001181192	0.692301399	Down-regulated
4183	Bt.89506	*NSL1*	0.001051208	0.690574993	Down-regulated
5143	Bt.4106	*F3*	0.00150891	0.680025782	Down-regulated
8644	Bt.8738	*HSPH1*	5.51E-05	0.679349341	Down-regulated
8644	Bt.8738	*HSPH1*	9.29E-05	0.679349341	Down-regulated
4322	Bt.13666	*ATP5S*	0.001395012	0.678999792	Down-regulated
8159	Bt.65398		0.001454744	0.676994275	Down-regulated
4066	Bt.73096	*CTH*	0.000403511	0.676064746	Down-regulated
4066	Bt.73096	*CTH*	0.000528372	0.676064746	Down-regulated
10681	Bt.59095	*C9orf97*	0.000647743	0.670492728	Down-regulated
1477	Bt.12399	*NEU1*	1.84E-05	0.669743822	Down-regulated
1477	Bt.12399	*NEU1*	5.03E-05	0.669743822	Down-regulated
3093	Bt.53211	*LOC615039*	1.14E-05	0.668364363	Down-regulated
3093	Bt.53211	*LOC615039*	0.000103313	0.668364363	Down-regulated
9586	Bt.88378	*KIAA1303*	1.52E-06	0.663116964	Down-regulated
9586	Bt.88378	*KIAA1303*	6.94E-06	0.663116964	Down-regulated
5766	Bt.65363	*C16orf69*	2.56E-05	0.66155183	Down-regulated
5766	Bt.65363	*C16orf69*	0.000270629	0.66155183	Down-regulated
2559	Bt.64693	*Dnaja1*	0.000442293	0.659012958	Down-regulated
2559	Bt.64693	*Dnaja1*	0.001322657	0.659012958	Down-regulated
288	Bt.28094	*CASP7*	1.68E-06	0.656324763	Down-regulated
288	Bt.28094	*CASP7*	0.000112935	0.656324763	Down-regulated
904	Bt.7418	*ANKH*	4.40E-05	0.644975795	Down-regulated
904	Bt.7418	*ANKH*	6.37E-05	0.644975795	Down-regulated
8867	Bt.64897	*IGFBP3*	0.001597096	0.642292452	Down-regulated
595	Bt.6522	*DNAJB6*	0.001496388	0.632396231	Down-regulated
11924	Bt.37906	*YOD1*	3.39E-05	0.630172128	Down-regulated
11924	Bt.37906	*YOD1*	0.000169627	0.630172128	Down-regulated

2027	Bt.10259	*PLK2*	0.00014694	0.618101393	Down-regulated
2287	Bt.27376	*Bag4*	4.25E-07	0.614071777	Down-regulated
2287	Bt.27376	*Bag4*	0.000106309	0.614071777	Down-regulated
10996	Bt.22526	*HSPB8*	7.83E-05	0.608102329	Down-regulated
10996	Bt.22526	*HSPB8*	0.001019679	0.608102329	Down-regulated
8731	Bt.27784	*PPP1R10*	4.21E-07	0.587989706	Down-regulated
8731	Bt.27784	*PPP1R10*	1.13E-06	0.587989706	Down-regulated
8808	Bt.53659	*ABHD13*	4.64E-05	0.550445569	Down-regulated
8808	Bt.53659	*ABHD13*	0.000432173	0.550445569	Down-regulated
320		*TMEM121*	2.64E-05	0.536403768	Down-regulated
2719	Bt.75092	*DIDO1*	0.000364325	0.530232529	Down-regulated
2719	Bt.75092	*DIDO1*	4.49E-07	0.530232529	Down-regulated

Table S4.4: Unique differentially expressed genes between HHP-treated embryos and controls. Spreadsheet 1: Up and down-regulated genes specific to 40 MPa treatment compared to control; Spreadsheet 2: Up and down-regulated genes specific to 60 MPa treatment compared to control; Spreadsheet 3: Up and down-regulated genes specific to 80 MPa treatment compared to control.

Table S4.4 Sheet 1. Genes uniquely differentially expressed at 40 Mpa*

Probe	UniGene ID	Gene name	P value	0.1 vs. 40 Mpa	0.1 vs. 60 MPa	0.1 vs. 80 MPa
7909	Bt.64563	*RFFL*	2.54E-05	-1	0	0
5332	Bt.12314	*PFKFB3*	3.97E-05	-1	0	0
5332	Bt.12314	*PFKFB3*	6.60E-05	-1	0	0
1968	Bt.53760	*LOC728216*	0.000104525	-1	0	0
1968	Bt.53760	*LOC728216*	0.000116805	-1	0	0
123	Bt.64291	*H3F3A*	0.000158185	-1	0	0
747	Bt.28249	*SEC24A*	0.000173964	-1	0	0
2986	Bt.27527	*STXBP3*	0.000277868	-1	0	0
12100	Bt.30953	*DOCK1*	0.000300917	-1	0	0
11731	Bt.22575	*TMEM167*	0.000368237	-1	0	0
3518	Bt.28502	*FAM83D*	0.000417802	-1	0	0
281	Bt.9783	*DDX3X*	0.000421378	-1	0	0
8159	Bt.89557		0.000463052	-1	0	0
123	Bt.64291	*H3F3A*	0.000471416	-1	0	0
3518	Bt.28502	*FAM83D*	0.000474399	-1	0	0
1794	Bt.64539	*QSCN6*	0.000535468	-1	0	0
538	Bt.13737	*GALK2*	0.0005438	-1	0	0
779	Bt.90218	*TES*	0.000610689	-1	0	0
7909	Bt.64563	*RFFL*	0.00069605	-1	0	0
265	Bt.20277	*TOP2A*	0.00078607	-1	0	0
281	Bt.9783	*DDX3X*	0.000835839	-1	0	0
942	Bt.6650	*GGH*	0.000841897	-1	0	0
10961	Bt.20683	*TP53INP2*	0.000874769	-1	0	0
289	Bt.21926	*LOC508142*	0.000881912	-1	0	0
8159			0.000946699	-1	0	0
7968	Bt.49157	*APOA1*	0.000973083	-1	0	0
8529	Bt.7930	*NUCB1*	0.001008492	-1	0	0
1116	Bt.25186	*BCAT1*	0.001013464	-1	0	0
6897	Bt.21510	*GPR107*	0.001032206	-1	0	0
2986	Bt.27527	*STXBP3*	0.001176051	-1	0	0
874	Bt.22035	*C5orf14*	0.001283155	-1	0	0
1794	Bt.64539	*QSCN6*	0.001353419	-1	0	0
10961	Bt.20683	*TP53INP2*	0.001621616	-1	0	0
9896	Bt.58080	*ZNF3*	1.85E-06	1	0	0
10009	Bt.48940	*PPAP2B*	1.96E-06	1	0	0
2357	Bt.5184	*Ankrd40*	2.39E-06	1	0	0

9896	Bt.58080	*ZNF3*	7.52E-06	1	0	0
6978	Bt.22374	*MGC137396*	9.39E-06	1	0	0
6978	Bt.22374	*MGC137396*	1.89E-05	1	0	0
5625	Bt.45080	*LRRC45*	2.36E-05	1	0	0
8432	Bt.5633	*TMEM129*	2.38E-05	1	0	0
8747	Bt.23521	*DUSP26*	2.74E-05	1	0	0
2575	Bt.2041	*AP2A2*	3.05E-05	1	0	0
4349	Bt.66386	*WRNIP1*	4.59E-05	1	0	0
2339	Bt.52556	*RCP9*	6.21E-05	1	0	0
4349	Bt.66386	*WRNIP1*	6.48E-05	1	0	0
8983	Bt.45485	*KLF9*	6.68E-05	1	0	0
967	Bt.73261	*LOC513955*	6.85E-05	1	0	0
8747	Bt.23521	*DUSP26*	6.93E-05	1	0	0
8250	Bt.22175	*KCNT1*	7.92E-05	1	0	0
9907	Bt.59214	*FGD5*	8.01E-05	1	0	0
4144	Bt.49595	*TMEM111*	8.44E-05	1	0	0
8983	Bt.45485	*KLF9*	9.23E-05	1	0	0
5945	Bt.26976	*COX11*	0.000101741	1	0	0
3885	Bt.51520	*C14orf94*	0.000123869	1	0	0
4656	Bt.32544	*NVL*	0.000128076	1	0	0
4986	Bt.89574	*KIAA0182*	0.00015045	1	0	0
6659	Bt.3196	*STRA6*	0.000152416	1	0	0
5633	Bt.2568	*ADFP*	0.000157168	1	0	0
393	Bt.58700	*C9orf58*	0.000166772	1	0	0
8159			0.000169244	1	0	0
10034	Bt.7363	*LOC789389*	0.000174731	1	0	0
3739	Bt.28081	*POGZ*	0.000189317	1	0	0
3885	Bt.51520	*C14orf94*	0.000191948	1	0	0
5633	Bt.2568	*ADFP*	0.000202974	1	0	0
1640	Bt.49587	*GPI*	0.000206942	1	0	0
3263	Bt.37779	*BPHL*	0.000210985	1	0	0
2193	Bt.52796	*COPS6*	0.000225488	1	0	0
2807	Bt.64366	*NDUFC2*	0.00024111	1	0	0
5727	Bt.70	*NDUFB10*	0.000260029	1	0	0
2339	Bt.52556	*RCP9*	0.000280575	1	0	0
6155	Bt.7804	*GYG1*	0.000301114	1	0	0
4802	Bt.22766	*SART3*	0.000322672	1	0	0
1689	Bt.39564	*LOC506268*	0.000332691	1	0	0
3263	Bt.37779	*BPHL*	0.000358399	1	0	0
3414	Bt.61384	*UNC84B*	0.000416827	1	0	0
2357	Bt.5184	*Ankrd40*	0.000435507	1	0	0
6830	Bt.52335	*IFT57*	0.000455475	1	0	0
11028	Bt.30639	*CNOT2*	0.000466637	1	0	0

7556	Bt.63484	*SURF2*	0.000466669	1	0	0
1858	Bt.62530	*ITPK1*	0.000482942	1	0	0
4505	Bt.15915	*IPO11*	0.000488539	1	0	0
1019	Bt.9195	*SIVA1*	0.000493087	1	0	0
8159			0.000509444	1	0	0
4532	Bt.4949	*SF3B5*	0.000553053	1	0	0
6830	Bt.52335	*IFT57*	0.000574792	1	0	0
5625	Bt.45080	*LRRC45*	0.000610941	1	0	0
10232	Bt.1858	*C20orf108*	0.000616414	1	0	0
8250	Bt.22175	*KCNT1*	0.000618956	1	0	0
1902	Bt.12743	*ABCD4*	0.00063545	1	0	0
393	Bt.58700	*C9orf58*	0.000647975	1	0	0
1640	Bt.49587	*GPI*	0.000651686	1	0	0
4400	Bt.49192	*RPL15*	0.000657603	1	0	0
5773	Bt.38862	*RALA*	0.000703113	1	0	0
2112	Bt.44470	*SATB1*	0.000705402	1	0	0
4931	Bt.27370	*HEATR3*	0.000707341	1	0	0
6501	Bt.48788	*RAB30*	0.000717957	1	0	0
5262	Bt.8206	*SFRS7*	0.000724622	1	0	0
10232	Bt.1858	*C20orf108*	0.000725454	1	0	0
2709	Bt.21070	*MRPL10*	0.000769143	1	0	0
5727	Bt.70	*NDUFB10*	0.000771141	1	0	0
4400	Bt.49192	*RPL15*	0.000774218	1	0	0
4532	Bt.4949	*SF3B5*	0.000798069	1	0	0
5987	Bt.13810	*U2AF1*	0.000815782	1	0	0
8460	Bt.5530	*DHRS3*	0.000816115	1	0	0
6339	Bt.44021	*WDR8*	0.000830074	1	0	0
1886	Bt.76825	*LOC789371*	0.000833992	1	0	0
2992	Bt.18512	*ASF1A*	0.000880779	1	0	0
2887	Bt.61617	*UHRF2*	0.000904125	1	0	0
6155	Bt.7804	*GYG1*	0.000905921	1	0	0
11275	Bt.2988	*GNG11*	0.000947206	1	0	0
11761	Bt.17953	*FOXP1*	0.000956226	1	0	0
4986	Bt.89574	*KIAA0182*	0.000987792	1	0	0
2555	Bt.5528	*SLC7A5*	0.001005994	1	0	0
2555	Bt.5528	*SLC7A5*	0.001009624	1	0	0
1858	Bt.62530	*ITPK1*	0.001071091	1	0	0
3794	Bt.44982	*CHRNA10*	0.001086009	1	0	0
3045	Bt.36226	*SMNDC1*	0.001098454	1	0	0
11668	Bt.49238	*HPRT1*	0.001115084	1	0	0
4008	Bt.26835	*RANGNRF*	0.001169618	1	0	0
11668	Bt.49238	*HPRT1*	0.00120866	1	0	0
11761	Bt.17953	*FOXP1*	0.001230279	1	0	0

4707	Bt.46116	*MXRA7*	0.001246439	1	0	0
2672	Bt.6612	*DDX51*	0.001257178	1	0	0
4834	Bt.8545	*MRPL43*	0.001262114	1	0	0
7621	Bt.9482	*LOC504245*	0.001291534	1	0	0
7122	Bt.22	*PPP1R8*	0.00131595	1	0	0
3472	Bt.49273	*POLR2F*	0.001433416	1	0	0
3739	Bt.28081	*POGZ*	0.001460923	1	0	0
5836	Bt.64649	*Znf313*	0.001516718	1	0	0
4505	Bt.15915	*IPO11*	0.001528633	1	0	0
11499	Bt.64697	*LMNB2*	0.001540602	1	0	0
8159	Bt.30026		0.001543259	1	0	0
8583	Bt.52379	*QRICH1*	0.001559509	1	0	0
10674	Bt.48127	*Tia1*	0.001569778	1	0	0
7193	Bt.3871	*LOC506318*	0.00160686	1	0	0

***: 1: up-regualted; -1: down-regulated; 0: not significant differential expressed**

Table S4.4 Sheet 2. Genes uniquely differentially expressed at 60 MPa*

Probe	UniGene ID	Gene name	P value	0.1 vs. 40 MPa	0.1 vs. 60 MPa	0.1 vs. 80 MPa
5154	Bt.52786	*LOC100008588*	0.000102515	0	-1	0
5154	Bt.52786	*LOC100008588*	0.000160742	0	-1	0
11261	Bt.9525	*RNF113A*	0.000258404	0	-1	0
11261	Bt.9525	*RNF113A*	0.000628905	0	-1	0
1895	Bt.42858	*ABCC5*	1.02E-05	0	1	0
1895	Bt.42858	*ABCC5*	2.23E-05	0	1	0
89	Bt.32543	*IPO8*	0.000191059	0	1	0
2167	Bt.20080	*WSB1*	0.000214607	0	1	0
1401	Bt.3199	*RUSC1*	0.000221096	0	1	0
9208	Bt.26958	*LOC788996*	0.000280312	0	1	0
5966	Bt.5519	*TncRNA*	0.000328087	0	1	0
8128	Bt.33726	*NPY*	0.000388758	0	1	0
7061	Bt.42818	*POLR3B*	0.000405958	0	1	0
4952	Bt.2642	*ZNF706*	0.000593589	0	1	0
5	Bt.29924	*MDM4*	0.000824537	0	1	0
2167	Bt.20080	*WSB1*	0.000831809	0	1	0
564	Bt.28460	*SNRPD3*	0.000971154	0	1	0
6217	Bt.11128	*MGC140717*	0.001127116	0	1	0
8474	Bt.89593	*PRPF4B*	0.001438541	0	1	0
7199	Bt.54456	*TFB2M*	0.001489348	0	1	0
6217	Bt.11128	*MGC140717*	0.00162156	0	1	0

***: 1: up-regualted; -1: down-regulated; 0: not significant differential expressed**

Table S4.4 Sheet 3. Genes uniquely differentially expressed at 80 MPa*

Probe	UniGene ID	Gene name	P-value	0.1 vs. 40 MPa	0.1 vs. 60 MPa	0.1 vs. 80 MPa
3757	Bt.2005	*LSM1*	1.50E-07	0	0	-1
3757	Bt.2005	*LSM1*	9.51E-06	0	0	-1
3093	Bt.53211	*LOC615039*	1.14E-05	0	0	-1
320		*TMEM121*	2.64E-05	0	0	-1
11924	Bt.37906	*YOD1*	3.39E-05	0	0	-1
8808	Bt.53659	*ABHD13*	4.64E-05	0	0	-1
1685	Bt.77469	*KCTD20*	6.84E-05	0	0	-1
3089	Bt.22538	*FZD8*	9.71E-05	0	0	-1
3093	Bt.53211	*LOC615039*	0.000103313	0	0	-1
3865	Bt.52937	*hCG_20417*	0.00010856	0	0	-1
7060	Bt.11280	*BANP*	0.00012219	0	0	-1
7060	Bt.11280	*BANP*	0.00013219	0	0	-1
1402	Bt.4279	*SHC1*	0.000135301	0	0	-1
3865	Bt.52937	*hCG_20417*	0.000139517	0	0	-1
9829	Bt.27267	*TMEM4*	0.000150732	0	0	-1
11924	Bt.37906	*YOD1*	0.000169627	0	0	-1
5840	Bt.1658	*DUSP1*	0.000276438	0	0	-1
1685	Bt.77469	*KCTD20*	0.000311099	0	0	-1
6408	Bt.52209	*RAD54B*	0.000316422	0	0	-1
4066	Bt.73096	*CTH*	0.000403511	0	0	-1
1402	Bt.4279	*SHC1*	0.00040895	0	0	-1
8808	Bt.53659	*ABHD13*	0.000432173	0	0	-1
2559	Bt.64693	*Dnaja1*	0.000442293	0	0	-1
4066	Bt.73096	*CTH*	0.000528372	0	0	-1
1109	Bt.17683	*DBF4*	0.000560078	0	0	-1
4027	Bt.11096	*KPNA4*	0.00070383	0	0	-1
4183	Bt.89506	*NSL1*	0.001051208	0	0	-1
10834	Bt.18408	*CCDC117*	0.001065327	0	0	-1
10834	Bt.18408	*CCDC117*	0.001272419	0	0	-1
2559	Bt.64693	*Dnaja1*	0.001322657	0	0	-1
5840	Bt.1658	*DUSP1*	0.001339091	0	0	-1
4027	Bt.11096	*KPNA4*	0.001347272	0	0	-1
4322	Bt.13666	*ATP5S*	0.001395012	0	0	-1
8159	Bt.65398		0.001454744	0	0	-1
337	Bt.17282	*SNIP1*	0.001581269	0	0	-1
6758	Bt.26855	*LOC506746*	0.001605013	0	0	-1
8159	Bt.89557		1.96E-05	0	0	1
42	Bt.34521	*RBM12*	0.000452853	0	0	1
2634	Bt.90254	*HDAC8*	0.000981392	0	0	1

***: 1: up-regualted; -1: down-regulated; 0: not significant differential expressed**

Table S4.5: Common genes between 40 MPa and 60 MPa treatments compared to control.

Probe	UniGene ID	Gene name	P-value	0.1 vs. 40 MPa	0.1 vs. 60 MPa	0.1 vs. 80 MPa
6779	Bt.90151	*LOC100008589*	8.45E-07	-1	-1	0
6779	Bt.90151	*LOC100008589*	2.96E-06	-1	-1	0
2402	Bt.7873	*BCAM*	6.99E-06	-1	-1	0
2402	Bt.7873	*BCAM*	1.10E-05	-1	-1	0
5355	Bt.269	*ATP2C1*	5.96E-05	-1	-1	0
5355	Bt.269	*ATP2C1*	0.000228997	-1	-1	0
5040	Bt.1786	*YME1L1*	0.00133787	-1	-1	0
3189	Bt.44977	*SLC4A1AP*	0.00145282	-1	-1	0
7528	Bt.49586	*NET1*	0.001601412	-1	-1	0
2380	Bt.23564	*C6orf120*	1.02E-06	1	1	0
3495	Bt.8279	*C3orf60*	1.30E-06	1	1	0
3495	Bt.8279	*C3orf60*	1.85E-06	1	1	0
2380	Bt.23564	*C6orf120*	2.66E-06	1	1	0
1090	Bt.7391	*APLP1*	3.84E-06	1	1	0
1536	Bt.66354	*POLR2L*	6.22E-06	1	1	0
1536	Bt.66354	*POLR2L*	1.19E-05	1	1	0
9020	Bt.4586	*GNG2*	1.81E-05	1	1	0
4237	Bt.87845	*COX5B*	2.00E-05	1	1	0
5328	Bt.55867	*HSPC152*	2.62E-05	1	1	0
9020	Bt.4586	*GNG2*	3.28E-05	1	1	0
2569	Bt.7562	*FLJ41352*	3.66E-05	1	1	0
4243	Bt.6710	*LOC440330*	4.50E-05	1	1	0
4237	Bt.87845	*COX5B*	4.62E-05	1	1	0
9505	Bt.46839	*SNRPD2*	6.96E-05	1	1	0
4927	Bt.55759	*DKFZp434K1815*	7.42E-05	1	1	0
4243	Bt.6710	*LOC440330*	7.53E-05	1	1	0
9505	Bt.46839	*SNRPD2*	8.10E-05	1	1	0
4834	Bt.8545	*MRPL43*	9.09E-05	1	1	0
5328	Bt.55867	*HSPC152*	9.14E-05	1	1	0
5262	Bt.8206	*SFRS7*	0.000104799	1	1	0
5768	Bt.49597	*LOC528919*	0.000113481	1	1	0
6385	Bt.9092	*POLR2G*	0.000122794	1	1	0
8583	Bt.52379	*QRICH1*	0.000137079	1	1	0
2654	Bt.4298	*GPR172B*	0.000194199	1	1	0
626	Bt.38424	*DST*	0.000199418	1	1	0
5945	Bt.26976	*COX11*	0.00020081	1	1	0
8159	Bt.65788		0.00021414	1	1	0
9208	Bt.26958	*LOC788996*	0.000215607	1	1	0
6385	Bt.9092	*POLR2G*	0.000220579	1	1	0

2654	Bt.4298	*GPR172B*	0.000251915	1	1	0
2887	Bt.61617	*UHRF2*	0.000439983	1	1	0
8159	Bt.732		0.000471878	1	1	0
4008	Bt.26835	*RANGNRF*	0.000511388	1	1	0
5768	Bt.49597	*LOC528919*	0.00051465	1	1	0
3585	Bt.26997	*Ptpmt1*	0.000530902	1	1	0
1090	Bt.7391	*APLP1*	0.000618699	1	1	0
5186	Bt.57711	*NDUFB8*	0.000654609	1	1	0
2471	Bt.49501	*C15orf40*	0.000680521	1	1	0
6124	Bt.43820	*RPL36*	0.000760297	1	1	0
2569	Bt.7562	*FLJ41352*	0.00076392	1	1	0
4263	Bt.23076	*C1orf93*	0.000764781	1	1	0
4263	Bt.23076	*C1orf93*	0.000783807	1	1	0
8266	Bt.16440	*USP6NL*	0.000800332	1	1	0
4345	Bt.17589	*DMTF1*	0.000879507	1	1	0
3472	Bt.49273	*POLR2F*	0.00090677	1	1	0
2877	Bt.30434	*SFRS9*	0.00091765	1	1	0
8266	Bt.16440	*USP6NL*	0.001046373	1	1	0
4884		*SEMA3G*	0.001075152	1	1	0
6997	Bt.15697	*PER2*	0.001261833	1	1	0
3543	Bt.4055	*BGLAP*	0.001321025	1	1	0
6124	Bt.43820	*RPL36*	0.001325246	1	1	0
9205	Bt.21568	*FAM118B*	0.001325776	1	1	0
11164	Bt.7658	*RPL38*	0.001398942	1	1	0
8159	Bt.6330		0.001422256	1	1	0
8654	Bt.13784	*TIMM10*	0.001437679	1	1	0
6140	Bt.27014	*FEZ2*	0.001617183	1	1	0

***: 1: up-regualted; -1: down-regulated; 0: not significant differential expressed**

Table S4.6: Common genes among 40, 60, and 80 MPa treatments compared to control.

Probe	UniGene ID	Gene name	P-value	Comment
2719	Bt.75092	*DIDO1*	4.49E-07	Down-regulated
9586	Bt.88378	*KIAA1303*	1.52E-06	Down-regulated
9586	Bt.88378	*KIAA1303*	6.94E-06	Down-regulated
3177	Bt.48951	*ASB11*	1.21E-05	Down-regulated
4305	Bt.89659	*FKBP9L*	2.02E-05	Down-regulated
3177	Bt.48951	*ASB11*	4.73E-05	Down-regulated
8644	Bt.8738	*HSPH1*	5.51E-05	Down-regulated
10996	Bt.22526	*HSPB8*	7.83E-05	Down-regulated
8644	Bt.8738	*HSPH1*	9.29E-05	Down-regulated
5271	Bt.4908	*RBBP7*	0.000265911	Down-regulated
5040	Bt.1786	*YME1L1*	0.000619741	Down-regulated

11849	Bt.9469	*DNAJC10*	0.000759213	Down-regulated
10996	Bt.22526	*HSPB8*	0.001019679	Down-regulated
3140	Bt.56130	*SIAHBP1*	0.001110559	Down-regulated
3140	Bt.56130	*SIAHBP1*	0.001220173	Down-regulated
7456	Bt.46948	*LRRFIP2*	0.001449248	Down-regulated
4927	Bt.55759	*DKFZp434K1815*	6.86E-08	Up-regulated
6288	Bt.13965	*CNTNAP1*	2.48E-07	Up-regulated
6288	Bt.13965	*CNTNAP1*	1.56E-06	Up-regulated
7223	Bt.500	*HK1*	0.000108524	Up-regulated
9205	Bt.21568	*FAM118B*	0.000132473	Up-regulated
7986	Bt.66	*NDUFA3*	0.000138015	Up-regulated
7223	Bt.500	*HK1*	0.000187069	Up-regulated
7986	Bt.66	*NDUFA3*	0.000226394	Up-regulated
4914	Bt.64707	*LOC146346*	0.000776269	Up-regulated
2865	Bt.89629	*EIF4B*	0.001144637	Up-regulated

Table S4.7: Differentially expressed genes of among different treatments of equilibration time. Spreadsheet 1: Up and down-regulated genes on 1 h treatment compared to control; Spreadsheet 2 Up and down-regulated genes on 2 h treatment compared to control.

Table S4.7 Sheet 1. Genes differentially expressed between 0 and 1h of recovery

Probe	UniGene ID	Gene name	P-value	Fold change (0 vs. 1h)	Comment
5625	Bt.45080	*LRRC45*	2.36E-05	1.269131265	Up-regulated
1109	Bt.17683	*DBF4*	0.000560078	1.243291605	Up-regulated
6288	Bt.13965	*CNTNAP1*	2.48E-07	1.23071886	Up-regulated
6288	Bt.13965	*CNTNAP1*	2.48E-07	1.23071886	Up-regulated
7986	Bt.66	*NDUFA3*	0.000138015	1.230378288	Up-regulated
7986	Bt.66	*NDUFA3*	0.000138015	1.230378288	Up-regulated
3414	Bt.61384	*UNC84B*	0.00011637	1.207528134	Up-regulated
3414	Bt.61384	*UNC84B*	0.00011637	1.207528134	Up-regulated
6779	Bt.89284	*LOC100008589*	8.45E-07	1.199860282	Up-regulated
6779	Bt.89284	*LOC100008589*	8.45E-07	1.199860282	Up-regulated
5165	Bt.2056	*APEH*	0.000356549	1.176167778	Up-regulated
5165	Bt.2056	*APEH*	0.000356549	1.176167778	Up-regulated
6830	Bt.52335	*IFT57*	0.000455475	1.166629102	Up-regulated
4237	Bt.87845	*COX5B*	2.00E-05	1.165799255	Up-regulated
4237	Bt.87845	*COX5B*	2.00E-05	1.165799255	Up-regulated
2274	Bt.44330	*LRP10*	0.000316879	1.067708969	Up-regulated
1685	Bt.77469	*KCTD20*	6.84E-05	0.893326617	Down-regulated
1685	Bt.77469	*KCTD20*	6.84E-05	0.893326617	Down-regulated
5271	Bt.4908	*RBBP7*	0.000265911	0.892729784	Down-regulated
779	Bt.90218	*TES*	0.000610689	0.875245348	Down-regulated
2559	Bt.64693	*Dnaja1*	0.000442293	0.867036407	Down-regulated

2559	Bt.64693	*Dnaja1*	0.000442293	0.867036407	Down-regulated
5633	Bt.2568	*ADFP*	0.000157168	0.856417406	Down-regulated
2887	Bt.61617	*UHRF2*	0.000439983	0.856087424	Down-regulated
9071	Bt.66625	*CNOT4*	0.000437896	0.855133904	Down-regulated
5836	Bt.64649	*Znf313*	0.001516718	0.852494945	Down-regulated
8035	Bt.38851	*RHPN2*	0.000230282	0.848287492	Down-regulated
8474	Bt.89593	*PRPF4B*	0.001438541	0.845408454	Down-regulated
1782	Bt.53344	*ARHGAP21*	0.000850676	0.843893588	Down-regulated
3093	Bt.53211	*LOC615039*	1.14E-05	0.843690779	Down-regulated
3093	Bt.53211	*LOC615039*	1.14E-05	0.843690779	Down-regulated
5	Bt.29924	*MDM4*	0.000824537	0.836946262	Down-regulated
3518	Bt.28502	*FAM83D*	0.000417802	0.815352753	Down-regulated
3518	Bt.28502	*FAM83D*	0.000417802	0.815352753	Down-regulated
8867	Bt.64897	*IGFBP3*	0.001597096	0.814325559	Down-regulated
2719	Bt.75092	*DIDO1*	4.49E-07	0.79833497	Down-regulated
2719	Bt.75092	*DIDO1*	4.49E-07	0.79833497	Down-regulated
337	Bt.17282	*SNIP1*	0.001581269	0.794623117	Down-regulated
288	Bt.28094	*CASP7*	1.68E-06	0.791653844	Down-regulated
288	Bt.28094	*CASP7*	1.68E-06	0.791653844	Down-regulated
893	Bt.59331	*UAP1*	3.55E-05	0.79045226	Down-regulated
893	Bt.59331	*UAP1*	3.55E-05	0.79045226	Down-regulated
2167	Bt.20080	*WSB1*	0.000214607	0.78232197	Down-regulated
2167	Bt.20080	*WSB1*	0.000214607	0.78232197	Down-regulated
10961	Bt.20683	*TP53INP2*	0.000874769	0.771774751	Down-regulated
10961	Bt.20683	*TP53INP2*	0.000874769	0.771774751	Down-regulated
1607	Bt.4371	*GADD45B*	0.001314898	0.770497271	Down-regulated
1897	Bt.12732	*NYREN18*	3.57E-05	0.768407693	Down-regulated
1897	Bt.12732	*NYREN18*	3.57E-05	0.768407693	Down-regulated
5143	Bt.4106	*F3*	0.00150891	0.764778651	Down-regulated
4274	Bt.11909	*DUSP6*	0.000701201	0.751863703	Down-regulated
4274	Bt.11909	*DUSP6*	0.000701201	0.751863703	Down-regulated
5332	Bt.12314	*PFKFB3*	3.97E-05	0.710209123	Down-regulated
5332	Bt.12314	*PFKFB3*	3.97E-05	0.710209123	Down-regulated
8731	Bt.27784	*PPP1R10*	4.21E-07	0.706705653	Down-regulated
8731	Bt.27784	*PPP1R10*	4.21E-07	0.706705653	Down-regulated
11924	Bt.37906	*YOD1*	3.39E-05	0.688987822	Down-regulated
11924	Bt.37906	*YOD1*	3.39E-05	0.688987822	Down-regulated
9151	Bt.77424	*Fubp1*	0.001387453	0.68726246	Down-regulated
9151	Bt.77424	*Fubp1*	0.001387453	0.68726246	Down-regulated
5840	Bt.1658	*DUSP1*	0.000276438	0.642889004	Down-regulated
5840	Bt.1658	*DUSP1*	0.000276438	0.642889004	Down-regulated
320		*TMEM121*	2.64E-05	0.492582637	Down-regulated
320		*TMEM121*	2.64E-05	0.492582637	Down-regulated

967	Bt.73261	*LOC513955*	6.85E-05	0.418946684	Down-regulated

Table S4.7 Sheet 2. Genes differentially expressed between 0 and 2h of recovery

Probe	UniGene ID	Gene name	P-value	Fold change (0h vs. 2h)	Comment
42	Bt.34521	*RBM12*	0.000452853	2.497810117	Up-regulated
4986	Bt.89574	*KIAA0182*	0.00015045	1.650827373	Up-regulated
4986	Bt.89574	*KIAA0182*	0.00015045	1.650827373	Up-regulated
385	Bt.36587	*APLN*	0.000390659	1.50615348	Up-regulated
6897	Bt.21510	*GPR107*	0.00052826	1.353259037	Up-regulated
6897	Bt.21510	*GPR107*	0.00052826	1.353259037	Up-regulated
6288	Bt.13965	*CNTNAP1*	2.48E-07	1.331670835	Up-regulated
6288	Bt.13965	*CNTNAP1*	2.48E-07	1.331670835	Up-regulated
7223	Bt.500	*HK1*	0.000108524	1.304444633	Up-regulated
7223	Bt.500	*HK1*	0.000108524	1.304444633	Up-regulated
5165	Bt.2056	*APEH*	0.000356549	1.297146386	Up-regulated
5165	Bt.2056	*APEH*	0.000356549	1.297146386	Up-regulated
265	Bt.20277	*TOP2A*	0.00078607	1.295875101	Up-regulated
4001	Bt.26739	*TOMM40L*	0.000557284	1.289914225	Up-regulated
3414	Bt.61384	*UNC84B*	0.00011637	1.279234189	Up-regulated
3414	Bt.61384	*UNC84B*	0.00011637	1.279234189	Up-regulated
6978	Bt.22374	*MGC137396*	9.39E-06	1.267036603	Up-regulated
6978	Bt.22374	*MGC137396*	9.39E-06	1.267036603	Up-regulated
4914	Bt.64707	*LOC146346*	0.000776269	1.266441112	Up-regulated
1401	Bt.3199	*RUSC1*	0.000221096	1.245308408	Up-regulated
8460	Bt.5530	*DHRS3*	0.000377253	1.225399019	Up-regulated
8460	Bt.5530	*DHRS3*	0.000377253	1.225399019	Up-regulated
3739	Bt.28081	*POGZ*	0.000189317	1.221355924	Up-regulated
2709	Bt.21070	*MRPL10*	0.000769143	1.21338753	Up-regulated
1116	Bt.25186	*BCAT1*	0.001013464	1.203687035	Up-regulated
393	Bt.58700	*C9orf58*	0.000166772	1.201367218	Up-regulated
393	Bt.58700	*C9orf58*	0.000166772	1.201367218	Up-regulated
1109	Bt.17683	*DBF4*	0.000560078	1.196284515	Up-regulated
2415	Bt.46615	*MGC138967*	0.001483978	1.181577304	Up-regulated
8159	Bt.89557		1.96E-05	1.166478459	Up-regulated
8159			1.96E-05	1.166478459	Up-regulated
4349	Bt.66386	*WRNIP1*	4.59E-05	1.146309036	Up-regulated
4349	Bt.66386	*WRNIP1*	4.59E-05	1.146309036	Up-regulated
4144	Bt.49595	*TMEM111*	8.44E-05	0.876101945	down-regulated
4144	Bt.49595	*TMEM111*	8.44E-05	0.876101945	down-regulated
4834	Bt.8545	*MRPL43*	9.09E-05	0.874209742	down-regulated
4834	Bt.8545	*MRPL43*	9.09E-05	0.874209742	down-regulated
8583	Bt.52379	*QRICH1*	0.000137079	0.867103998	down-regulated

8404	Bt.76983	*BIRC5*	0.00034293	0.865148925	down-regulated
5040	Bt.1786	*YME1L1*	0.000619741	0.861594366	down-regulated
5040	Bt.1786	*YME1L1*	0.000619741	0.861594366	down-regulated
8159	Bt.11282		1.96E-05	0.857281154	down-regulated
8159	Bt.11282		1.96E-05	0.857281154	down-regulated
1477	Bt.12399	*NEU1*	1.84E-05	0.835950196	down-regulated
3140	Bt.56130	*SIAHBP1*	0.001110559	0.834894943	down-regulated
3140	Bt.56130	*SIAHBP1*	0.001110559	0.834894943	down-regulated
3806	Bt.25997	*CABLES1*	0.000852658	0.834548327	down-regulated
11849	Bt.9469	*DNAJC10*	0.000759213	0.829061036	down-regulated
3736	Bt.49705	*PRDX6*	0.001085807	0.827086914	down-regulated
3736	Bt.49705	*PRDX6*	0.001085807	0.827086914	down-regulated
7122	Bt.22	*PPP1R8*	0.00131595	0.821455382	down-regulated
9896	Bt.58080	*ZNF3*	1.85E-06	0.81790242	down-regulated
9896	Bt.58080	*ZNF3*	1.85E-06	0.81790242	down-regulated
2887	Bt.61617	*UHRF2*	0.000439983	0.817883423	down-regulated
2887	Bt.61617	*UHRF2*	0.000439983	0.817883423	down-regulated
9907	Bt.59214	*FGD5*	8.01E-05	0.816453471	down-regulated
4400	Bt.49192	*RPL15*	0.000657603	0.814329769	down-regulated
4400	Bt.49192	*RPL15*	0.000657603	0.814329769	down-regulated
2357	Bt.5184	*Ankrd40*	2.39E-06	0.812380507	down-regulated
2357	Bt.5184	*Ankrd40*	2.39E-06	0.812380507	down-regulated
595	Bt.6522	*DNAJB6*	0.001496388	0.811001765	down-regulated
9071	Bt.66625	*CNOT4*	0.000437896	0.810971414	down-regulated
4656	Bt.32544	*NVL*	0.000128076	0.803804047	down-regulated
3472	Bt.49273	*POLR2F*	0.00090677	0.800728012	down-regulated
3472	Bt.49273	*POLR2F*	0.00090677	0.800728012	down-regulated
1090	Bt.7391	*APLP1*	3.84E-06	0.799066523	down-regulated
7060	Bt.11280	*BANP*	0.00012219	0.79017418	down-regulated
7060	Bt.11280	*BANP*	0.00012219	0.79017418	down-regulated
5836	Bt.64649	*Znf313*	0.001516718	0.786717476	down-regulated
3865	Bt.52937	*hCG_20417*	0.00010856	0.786277948	down-regulated
3865	Bt.52937	*hCG_20417*	0.00010856	0.786277948	down-regulated
5271	Bt.4908	*RBBP7*	0.000265911	0.783934201	down-regulated
1685	Bt.77469	*KCTD20*	6.84E-05	0.773831736	down-regulated
1685	Bt.77469	*KCTD20*	6.84E-05	0.773831736	down-regulated
3177	Bt.48951	*ASB11*	1.21E-05	0.773298905	down-regulated
3177	Bt.48951	*ASB11*	1.21E-05	0.773298905	down-regulated
779	Bt.90218	*TES*	0.000610689	0.763933829	down-regulated
10834	Bt.18408	*CCDC117*	0.001065327	0.759673645	down-regulated
10834	Bt.18408	*CCDC117*	0.001065327	0.759673645	down-regulated
3757	Bt.2005	*LSM1*	1.50E-07	0.746170346	down-regulated
3757	Bt.2005	*LSM1*	1.50E-07	0.746170346	down-regulated

337	Bt.17282	*SNIP1*	0.001581269	0.744157241	down-regulated
8808	Bt.53659	*ABHD13*	4.64E-05	0.742459346	down-regulated
8808	Bt.53659	*ABHD13*	4.64E-05	0.742459346	down-regulated
1897	Bt.12732	*NYREN18*	3.57E-05	0.734240435	down-regulated
1897	Bt.12732	*NYREN18*	3.57E-05	0.734240435	down-regulated
967	Bt.73261	*LOC513955*	6.85E-05	0.734214548	down-regulated
4952	Bt.2642	*ZNF706*	0.000593589	0.728207851	down-regulated
4952	Bt.2642	*ZNF706*	0.000593589	0.728207851	down-regulated
3089	Bt.22538	*FZD8*	9.71E-05	0.726791936	down-regulated
2287	Bt.27376	*Bag4*	4.25E-07	0.723285019	down-regulated
2287	Bt.27376	*Bag4*	4.25E-07	0.723285019	down-regulated
1782	Bt.53344	*ARHGAP21*	0.000850676	0.723115673	down-regulated
8731	Bt.27784	*PPP1R10*	4.21E-07	0.719613028	down-regulated
8731	Bt.27784	*PPP1R10*	4.21E-07	0.719613028	down-regulated
8644	Bt.8738	*HSPH1*	5.51E-05	0.71875201	down-regulated
8644	Bt.8738	*HSPH1*	5.51E-05	0.71875201	down-regulated
5143	Bt.4106	*F3*	0.00150891	0.718339566	down-regulated
3518	Bt.28502	*FAM83D*	0.000417802	0.713784544	down-regulated
3518	Bt.28502	*FAM83D*	0.000417802	0.713784544	down-regulated
5966	Bt.5519	*TncRNA*	0.000328087	0.710419396	down-regulated
5966	Bt.5519	*TncRNA*	0.000328087	0.710419396	down-regulated
2559	Bt.64693	*Dnaja1*	0.000442293	0.690349473	down-regulated
2559	Bt.64693	*Dnaja1*	0.000442293	0.690349473	down-regulated
10961	Bt.20683	*TP53INP2*	0.000874769	0.675668865	down-regulated
10961	Bt.20683	*TP53INP2*	0.000874769	0.675668865	down-regulated
893	Bt.59331	*UAP1*	3.55E-05	0.659111818	down-regulated
893	Bt.59331	*UAP1*	3.55E-05	0.659111818	down-regulated
4027	Bt.11096	*KPNA4*	0.00070383	0.651833157	down-regulated
4027	Bt.11096	*KPNA4*	0.00070383	0.651833157	down-regulated
7815	Bt.6432	*FAM44B*	0.001395402	0.642693905	down-regulated
6997	Bt.15697	*PER2*	0.001261833	0.614594305	down-regulated
10996	Bt.22526	*HSPB8*	7.83E-05	0.61047851	down-regulated
10996	Bt.22526	*HSPB8*	7.83E-05	0.61047851	down-regulated
11924	Bt.37906	*YOD1*	3.39E-05	0.60781931	down-regulated
11924	Bt.37906	*YOD1*	3.39E-05	0.60781931	down-regulated
2027	Bt.10259	*PLK2*	0.00014694	0.602727628	down-regulated
320		*TMEM121*	2.64E-05	0.601343868	down-regulated
5332	Bt.12314	*PFKFB3*	3.97E-05	0.598072051	down-regulated
5332	Bt.12314	*PFKFB3*	3.97E-05	0.598072051	down-regulated
288	Bt.28094	*CASP7*	1.68E-06	0.5950576	down-regulated
288	Bt.28094	*CASP7*	1.68E-06	0.5950576	down-regulated
4274	Bt.11909	*DUSP6*	0.000701201	0.593725436	down-regulated
4274	Bt.11909	*DUSP6*	0.000701201	0.593725436	down-regulated

8867	Bt.64897	*IGFBP3*	0.001597096	0.587502707	down-regulated
5633	Bt.2568	*ADFP*	0.000157168	0.544075346	down-regulated
5633	Bt.2568	*ADFP*	0.000157168	0.544075346	down-regulated
5840	Bt.1658	*DUSP1*	0.000276438	0.509476418	down-regulated
5840	Bt.1658	*DUSP1*	0.000276438	0.509476418	down-regulated
7811	Bt.39540	*IL15RA*	0.000207216	0.44008029	down-regulated
7811	Bt.39540	*IL15RA*	0.000207216	0.44008029	down-regulated

Table S4.8: Unique differentially expressed genes among different treatments of equilibration time. Spreadsheet 1: Up and down-regulated genes specific to 1 h treatment compared to control; Spreadsheet 2: Up and down-regulated genes specific to 2 h treatment compared to control.

Table S4.8 Sheet 1. Unique differentials of the 0 vs. 1h comparison

Probe	UniGene ID	Gene name	P-value	Comment
4237	Bt.87845	COX5B	4.62E-05	Up-regulated
5625	Bt.45080	LRRC45	2.36E-05	Up-regulated
4237	Bt.87845	COX5B	2.00E-05	Up-regulated
7986	Bt.66	NDUFA3	0.000138015	Up-regulated
7986	Bt.66	NDUFA3	0.000226394	Up-regulated
1109	Bt.17683	DBF4	0.000560078	Up-regulated
2274	Bt.44330	LRP10	0.000316879	Up-regulated
2719	Bt.75092	DIDO1	4.49E-07	Down-regulated
3093	Bt.53211	LOC615039	1.14E-05	Down-regulated
320		TMEM121	2.64E-05	Down-regulated
3093	Bt.53211	LOC615039	1.14E-05	Down-regulated
2167	Bt.20080	WSB1	0.000214607	Down-regulated
8035	Bt.38851	RHPN2	0.000230282	Down-regulated
2719	Bt.75092	DIDO1	4.49E-07	Down-regulated
5	Bt.29924	MDM4	0.000824537	Down-regulated
2167	Bt.20080	WSB1	0.000214607	Down-regulated
1607	Bt.4371	GADD45B	0.001314898	Down-regulated
9151	Bt.77424	Fubp1	0.001387453	Down-regulated
8474	Bt.89593	PRPF4B	0.001438541	Down-regulated
9151	Bt.77424	Fubp1	0.001387453	Down-regulated

Table S8 Sheet 2. Unique differentials of the 0 vs. 2h comparison

Probe	UniGene ID	Gene name	P-value	Comment
6978	Bt.22374	MGC137396	9.39E-06	Up-regulated
6978	Bt.22374	MGC137396	9.39E-06	Up-regulated
4349	Bt.66386	WRNIP1	4.59E-05	Up-regulated
4349	Bt.66386	WRNIP1	4.59E-05	Up-regulated

393	Bt.58700	C9orf58	0.000166772	Up-regulated
3739	Bt.28081	POGZ	0.000189317	Up-regulated
8460	Bt.5530	DHRS3	0.000377253	Up-regulated
8159	Bt.89557		1.96E-05	Up-regulated
8159			1.96E-05	Up-regulated
4001	Bt.26739	TOMM40L	0.000557284	Up-regulated
393	Bt.58700	C9orf58	0.000166772	Up-regulated
2709	Bt.21070	MRPL10	0.000769143	Up-regulated
8460	Bt.5530	DHRS3	0.000377253	Up-regulated
1109	Bt.17683	DBF4	0.000560078	Up-regulated
2415	Bt.46615	MGC138967	0.001483978	Up-regulated
9896	Bt.58080	ZNF3	1.85E-06	Down-regulated
9896	Bt.58080	ZNF3	1.85E-06	Down-regulated
1477	Bt.12399	NEU1	1.84E-05	Down-regulated
8644	Bt.8738	HSPH1	5.51E-05	Down-regulated
8644	Bt.8738	HSPH1	5.51E-05	Down-regulated
3089	Bt.22538	FZD8	9.71E-05	Down-regulated
7060	Bt.11280	BANP	0.00012219	Down-regulated
3865	Bt.52937	hCG_20417	0.00010856	Down-regulated
5966	Bt.5519	TncRNA	0.000328087	Down-regulated
8404	Bt.76983	BIRC5	0.00034293	Down-regulated
8808	Bt.53659	ABHD13	4.64E-05	Down-regulated
2887	Bt.61617	UHRF2	0.000439983	Down-regulated
5040	Bt.1786	YME1L1	0.000619741	Down-regulated
11849	Bt.9469	DNAJC10	0.000759213	Down-regulated
3472	Bt.49273	POLR2F	0.00090677	Down-regulated
10996	Bt.22526	HSPB8	7.83E-05	Down-regulated
3736	Bt.49705	PRDX6	0.001085807	Down-regulated
3140	Bt.56130	SIAHBP1	0.001110559	Down-regulated
5966	Bt.5519	TncRNA	0.000328087	Down-regulated
3140	Bt.56130	SIAHBP1	0.001110559	Down-regulated
3736	Bt.49705	PRDX6	0.001085807	Down-regulated
4834	Bt.8545	MRPL43	9.09E-05	Down-regulated
10834	Bt.18408	CCDC117	0.001065327	Down-regulated
7122	Bt.22	PPP1R8	0.00131595	Down-regulated
5040	Bt.1786	YME1L1	0.000619741	Down-regulated
595	Bt.6522	DNAJB6	0.001496388	Down-regulated

Table S4.9: Common genes among comparisons of different treatments of equilibration time.

Table S9. Overlapped differentials of the 0 vs. 1h and 0 vs. 2h comparisons*

Probe	UniGene ID	Gene name	P-value	0 vs. 1h	0 vs. 2h

6288	Bt.13965	*CNTNAP1*	2.48E-07	1	1
6288	Bt.13965	*CNTNAP1*	2.48E-07	1	1
3414	Bt.61384	*UNC84B*	0.00011637	1	1
5165	Bt.2056	*APEH*	0.000356549	1	1
3414	Bt.61384	*UNC84B*	0.00011637	1	1
5165	Bt.2056	*APEH*	0.000356549	1	1
8731	Bt.27784	*PPP1R10*	4.21E-07	-1	-1
8731	Bt.27784	*PPP1R10*	4.21E-07	-1	-1
11924	Bt.37906	*YOD1*	3.39E-05	-1	-1
893	Bt.59331	*UAP1*	3.55E-05	-1	-1
1897	Bt.12732	*NYREN18*	3.57E-05	-1	-1
5332	Bt.12314	*PFKFB3*	3.97E-05	-1	-1
5332	Bt.12314	*PFKFB3*	3.97E-05	-1	-1
893	Bt.59331	*UAP1*	3.55E-05	-1	-1
967	Bt.73261	*LOC513955*	6.85E-05	-1	-1
1897	Bt.12732	*NYREN18*	3.57E-05	-1	-1
11924	Bt.37906	*YOD1*	3.39E-05	-1	-1
1685	Bt.77469	*KCTD20*	6.84E-05	-1	-1
3518	Bt.28502	*FAM83D*	0.000417802	-1	-1
9071	Bt.66625	*CNOT4*	0.000437896	-1	-1
2559	Bt.64693	*Dnaja1*	0.000442293	-1	-1
3518	Bt.28502	*FAM83D*	0.000417802	-1	-1
779	Bt.90218	*TES*	0.000610689	-1	-1
4274	Bt.11909	*DUSP6*	0.000701201	-1	-1
1782	Bt.53344	*ARHGAP21*	0.000850676	-1	-1
10961	Bt.20683	*TP53INP2*	0.000874769	-1	-1
4274	Bt.11909	*DUSP6*	0.000701201	-1	-1
2887	Bt.61617	*UHRF2*	0.000439983	-1	-1
2559	Bt.64693	*Dnaja1*	0.000442293	-1	-1
5143	Bt.4106	*F3*	0.00150891	-1	-1
5836	Bt.64649	*Znf313*	0.001516718	-1	-1
337	Bt.17282	*SNIP1*	0.001581269	-1	-1
10961	Bt.20683	*TP53INP2*	0.000874769	-1	-1
288	Bt.28094	*CASP7*	1.68E-06	-1	-1
1685	Bt.77469	*KCTD20*	6.84E-05	-1	-1
288	Bt.28094	*CASP7*	1.68E-06	-1	-1
5633	Bt.2568	*ADFP*	0.000157168	-1	-1
5271	Bt.4908	*RBBP7*	0.000265911	-1	-1
5840	Bt.1658	*DUSP1*	0.000276438	-1	-1
5840	Bt.1658	*DUSP1*	0.000276438	-1	-1
8867	Bt.64897	*IGFBP3*	0.001597096	-1	-1
320		*TMEM121*	2.64E-05	-1	-1

***: 1: up-regualted; -1: down-regulated**

4.7. References

1 Silva, J. L., Foguel, D. & Royer, C. A. Pressure provides new insights into protein folding, dynamics and structure. *Trends Biochem Sci* **26**, 612-618 (2001).

2 Paula-Lopes, F. F. & Hansen, P. J. Apoptosis is an adaptive response in bovine preimplantation embryos that facilitates survival after heat shock. *Biochemical and biophysical research communications* **295**, 37-42 (2002).

3 Paula-Lopes, F. F. & Hansen, P. J. Heat shock-induced apoptosis in preimplantation bovine embryos is a developmentally regulated phenomenon. *Biology of reproduction* **66**, 1169-1177 (2002).

4 Horvath, I. *et al.* Heat shock response in photosynthetic organisms: membrane and lipid connections. *Progress in lipid research* **51**, 208-220, doi:10.1016/j.plipres.2012.02.002 (2012).

5 Pribenszky, C., Molnar, M., Cseh, S. & Solti, L. Improving post-thaw survival of cryopreserved mouse blastocysts by hydrostatic pressure challenge. *Animal reproduction science* **87**, 143-150, doi:10.1016/j.anireprosci.2004.09.007 (2005).

6 Du, Y. *et al.* High hydrostatic pressure: a new way to improve in vitro developmental competence of porcine matured oocytes after vitrification. *Reproduction* **135**, 13-17, doi:10.1530/REP-07-0362 (2008).

7 Pribenszky, C., Du, Y., Molnar, M., Harnos, A. & Vajta, G. Increased stress tolerance of matured pig oocytes after high hydrostatic pressure treatment. *Animal reproduction science* **106**, 200-207, doi:10.1016/j.anireprosci.2008.01.016 (2008).

8 Pribenszky, C. *et al.* Stress preconditioning of boar spermatozoa: a new approach to enhance semen quality. *Reproduction in domestic animals = Zuchthygiene* **46 Suppl 2**, 26-30, doi:10.1111/j.1439-0531.2011.01812.x (2011).

9 Huang, S. Y. *et al.* Hydrostatic pressure pre-treatment affects the protein profile of boar sperm before and after freezing-thawing. *Animal reproduction science* **112**, 136-149, doi:10.1016/j.anireprosci.2008.04.016 (2009).

10 Bock, I. *et al.* Stress Tolerance and Transcriptional Response in Mouse Embryos Treated with High Hydrostatic Pressure to Enhance Cryotolerance. *Cryoletters* **31**, 401-412 (2010).

11 Siqueira, E. *et al.* Vitrification of bovine blastocysts pretreated with sublethal hydrostatic pressure stress: evaluation of post-thaw in vitro development and gene expression. *Reprod Fert Develop* **23**, 585-590 (2011).

12 Trigal, B. *et al.* Cell counts and survival to vitrification of bovine in vitro produced blastocysts subjected to sublethal high hydrostatic pressure. *Reproduction in domestic animals = Zuchthygiene* **48**, 200-206, doi:10.1111/j.1439-0531.2012.02131.x (2013).

13 Pribenszky, C. *et al.* Stress for stress tolerance? A fundamentally new approach in mammalian embryology. *Biology of reproduction* **83**, 690-697, doi:10.1095/biolreprod.110.083386 (2010).

14 Pribenszky, C. *et al.* 89 Improved post-warming developmental cometence of open pulled straw-vitrified in vitro produced bovine blastocysts by sublethal hydrostatic pressure pretreatment. *Reproduction, Fertility and Development* **20**, 125-125, doi:http://dx.doi.org/10.1071/RDv20n1Ab89 (2008).

15 Silva, J. L., Foguel, D. & Royer, C. A. Pressure provides new insights into protein folding, dynamics and structure. *Trends Biochem Sci* **26**, 612-618 (2001).
16 Wemekamp-Kamphuis, H. H., Karatzas, A. K., Wouters, J. A. & Abee, T. Enhanced Levels of Cold Shock Proteins in Listeria monocytogenes LO28 upon Exposure to Low Temperature and High Hydrostatic Pressure. *Applied and Environmental Microbiology* **68**, 456-463, doi:10.1128/aem.68.2.456-463.2002 (2002).
17 Everts, R. E. *et al.* A 7872 cDNA microarray and its use in bovine functional genomics. *Veterinary immunology and immunopathology* **105**, 235-245, doi:10.1016/j.vetimm.2005.02.003 (2005).
18 Smith, S. L. *et al.* Global gene expression profiles reveal significant nuclear reprogramming by the blastocyst stage after cloning. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 17582-17587, doi:10.1073/pnas.0508952102 (2005).
19 Dinnyes, A., Dai, Y., Jiang, S. & Yang, X. High developmental rates of vitrified bovine oocytes following parthenogenetic activation, in vitro fertilization, and somatic cell nuclear transfer. *Biology of reproduction* **63**, 513-518 (2000).
20 Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**, 44-57, doi:10.1038/nprot.2008.211 (2009).
21 Saragusty, J. & Arav, A. Current progress in oocyte and embryo cryopreservation by slow freezing and vitrification. *Reproduction* **141**, 1-19, doi:10.1530/REP-10-0236 (2011).
22 Pribenszky, C. & Vajta, G. Cells under pressure: how sublethal hydrostatic pressure stress treatment increases gametes' and embryos' performance. *Reproduction, fertility, and development* **23**, 48-55, doi:10.1071/RD10231 (2011).
23 Loureiro, B., Brad, A. M. & Hansen, P. J. Heat shock and tumor necrosis factor-alpha induce apoptosis in bovine preimplantation embryos through a caspase-9-dependent mechanism. *Reproduction* **133**, 1129-1137, doi:10.1530/REP-06-0307 (2007).
24 Silva, C. F. *et al.* Effects of heat stress on development, quality and survival of Bos indicus and Bos taurus embryos produced in vitro. *Theriogenology* **79**, 351-357, doi:10.1016/j.theriogenology.2012.10.003 (2013).
25 Fear, J. M. & Hansen, P. J. Developmental changes in expression of genes involved in regulation of apoptosis in the bovine preimplantation embryo. *Biology of reproduction* **84**, 43-51, doi:10.1095/biolreprod.110.086249 (2011).
26 Hansen, P. J. To be or not to be--determinants of embryonic survival following heat shock. *Theriogenology* **68 Suppl 1**, S40-48, doi:10.1016/j.theriogenology.2007.03.013 (2007).
27 Du, Y. *et al.* High hydrostatic pressure treatment of porcine oocytes before handmade cloning improves developmental competence and cryosurvival. *Cloning and stem cells* **10**, 325-330, doi:10.1089/clo.2007.0089 (2008).
28 Elmore, S. Apoptosis: a review of programmed cell death. *Toxicol Pathol* **35**, 495-516, doi:10.1080/01926230701320337 (2007).
29 Ouyang, L. *et al.* Programmed cell death pathways in cancer: a review of apoptosis, autophagy and programmed necrosis. *Cell Prolif* **45**, 487-498, doi:10.1111/j.1365-2184.2012.00845.x (2012).

30 Yamagishi, N., Ishihara, K., Saito, Y. & Hatayama, T. Hsp105 but not Hsp70 family proteins suppress the aggregation of heat-denatured protein in the presence of ADP. *FEBS Letters* **555**, 390-396, doi:10.1016/s0014-5793(03)01292-4 (2003).
31 Hosaka, S. *et al.* Synthetic small interfering RNA targeting heat shock protein 105 induces apoptosis of various cancer cells both in vitro and in vivo. *Cancer Sci* **97**, 623-632, doi:10.1111/j.1349-7006.2006.00217.x (2006).
32 Li, B. *et al.* Overload of the heat-shock protein H11/HspB8 triggers melanoma cell apoptosis through activation of transforming growth factor-beta-activated kinase 1. *Oncogene* **26**, 3521-3531, doi:10.1038/sj.onc.1210145 (2007).
33 Garcia-Domingo, D., Ramirez, D., Gonzalez de Buitrago, G. & Martinez, A. C. Death inducer-obliterator 1 triggers apoptosis after nuclear translocation and caspase upregulation. *Molecular and cellular biology* **23**, 3216-3225 (2003).
34 Wu, Y. Y. *et al.* Inhibition of hepatocellular carcinoma growth and angiogenesis by dual silencing of NET-1 and VEGF. *J Mol Histol* **44**, 433-445, doi:10.1007/s10735-012-9480-5 (2013).
35 Shahbazian, D., Parsyan, A., Petroulakis, E., Hershey, J. & Sonenberg, N. eIF4B controls survival and proliferation and is regulated by proto-oncogenic signaling pathways. *Cell cycle* **9**, 4106-4109 (2010).
36 Shahbazian, D. *et al.* Control of cell survival and proliferation by mammalian eukaryotic initiation factor 4B. *Molecular and cellular biology* **30**, 1478-1485, doi:10.1128/MCB.01218-09 (2010).
37 Dennis, M. D., Jefferson, L. S. & Kimball, S. R. Role of p70S6K1-mediated phosphorylation of eIF4B and PDCD4 proteins in the regulation of protein synthesis. *The Journal of biological chemistry* **287**, 42890-42899, doi:10.1074/jbc.M112.404822 (2012).
38 Aertsen, A. *et al.* Heat shock protein-mediated resistance to high hydrostatic pressure in Escherichia coli. *Appl Environ Microbiol* **70**, 2660-2666 (2004).
39 Jiang, Z. *et al.* Transcriptional profiles of bovine in vivo pre-implantation development. *BMC genomics* **15**, 756, doi:10.1186/1471-2164-15-756 (2014).
40 Martinez-Gomariz, M., Hernaez, M. L., Gutierrez, D., Ximenez-Embun, P. & Prestamo, G. Proteomic analysis by two-dimensional differential gel electrophoresis (2D DIGE) of a high-pressure effect in Bacillus cereus. *J Agric Food Chem* **57**, 3543-3549, doi:10.1021/jf803272a (2009).

Chapter Five

Knockdown of Brm and Baf170, Components of Chromatin Remodeling Complex, Facilitates Reprogramming of Somatic Cells

Published in Stem Cells and Development, 2015 Jun 29.

(Jiang Z, Tang Y, Zhao X, Zhang M, Donovan DM &Tian XC)

5.1. Abstract

The SWI/SNF (SWItch/Sucrose NonFermentable or BAF, Brg/Brahma-associated factors) complexes are epigenetic modifiers of chromatin structure and undergo progressive changes in subunit composition during cellular differentiation. For example, in embryonic stem cells (ESCs) esBAF contains *Brg1* and *Baf155*, while their homologs, *Brm* and *Baf170,* are present in BAF of somatic cells. Here we sought to determine whether *Brm* and *Baf170* play any roles in iPSC reprogramming by using shRNA-mediated knockdown studies in the mouse model. We found that knocking down *Brm* during early, mid and late stages (Days 3, 6 and 9 after initial iPSC induction) and knocking down *Baf170* during late stage (Day 9) reprogramming improve the numbers of iPSC colonies formed. We further showed that inhibition of these somatic BAF components also promotes complete reprogramming of partially reprogrammed somatic cells (pre-iPSCs). Finally, we found that the expression of *Brm* and *Baf170* during reprogramming was regulated by Jak/Stat3 activity. Taken together, these data suggest that inhibiting somatic BAF improves complete reprogramming by facilitating the activation of the "pluripotency circuitry".

Key words: BAF complex, *Brm*, *Baf170*, Induced pluripotent stem cells, Reprogramming efficiency

5.2. Introduction

Induced pluripotent stem cells (iPSCs) are embryonic stem cell (ESC) – like cells reprogrammed using ectopic transcription factors *Oct4, Sox2, Klf4*, and *c-Myc* (OKSM) [1,2]. However, transcription factor-mediated reprogramming is a slow and inefficient process, achieved by overcoming a series of epigenetic barriers [3]. Acquisition of induced pluripotency requires an intricate interplay among specialized transcriptional circuitries, signaling pathways and chromatin remodeling. In addition to DNA and histone modifications, ATP-dependent enzymes that remodel chromatin are important controllers of chromatin structure and assembly, and are major contributors to regulations of gene expression [4,5].

The SWI/SNF (SWItch/Sucrose NonFermentable) [also known as BAF (Brg/Brahma-associated factors)] complex is consisted of at least 15 core subunits and has ATP-dependent chromatin remodeling activity. It is essential for the formation of totipotent and pluripotent cells of early embryos [6]. In addition, BAF complex is the most frequently mutated chromatin regulatory complex in human cancers and thus their manipulation constitutes a major strategy for tumor suppression [7]. The BAF complex participates in numerous developmental transitions by changing its subunit composition. For example, the BAF complex in embryonic stem cells (ESCs), esBAF, has a unique subunit composition defined by the presence of *Brg*, *Baf155*, and *Baf60a*, and the absence of their somatic cell homologues *Brm*, *Baf170*, and *Baf60c* [8]. Altering this subunit composition caused a reduction in self-renewal and pluripotency in mouse ESCs (mESCs) [8]. In addition, *Baf250a* is also essential for self-renewal and pluripotency in mESCs [9,10]. It has been shown that the mechanisms of maintaining ESC pluripotency by esBAF are mediated by conditioning the genome for LIF/STAT3 signaling and by regulating the functions of the polycomb complex [11]. Conversely, adding esBAF components to fibroblasts facilitates

their reprogramming to pluripotent cells. For example, *Brg1* and *Baf155*, combined with *Oct4*, *Sox2*, *Klf4* and *c-Myc*, synergistically increased reprogramming efficiency by enhancing the binding of *Oct4* to target promoters [12]. These data also suggest that specific components of the BAF complex serve to facilitate the activation of the "pluripotency circuitry".

Given the influence of epigenetic factors over reprogramming fate and the documented role of SWI/SNF complexes in pluripotency, we sought to test the roles of somatic *Brm* and *Baf170* in mouse iPSC generation through shRNA-mediated knockdown studies. Using mouse embryonic fibroblasts (MEFs) harboring the green fluorescence protein (GFP) driven by the Oct4-promoter (OG-MEFs), we found that inhibiting components of the somatic BAF improves complete reprogramming by facilitating the activation of the "pluripotency circuitry".

5.3. Materials and Methods

5.3.1. Chemicals and Protein Expression Constructs

Jak inhibitor I (Jaki) and doxycycline were purchased from EMD Millipore (Billerica, MA, USA). Erk inhibitor PD0329501 and GSK3β inhibitor CHIR99021 (CHIR) were obtained from SelleckChem (Houston, TX, USA). The vectors for *pMXs-Oct4, Klf4, Sox2,* and *c-Myc* [1], pLKO.1-puro, pLKO.1-scramble shRNA control [13], Retro- and Lenti-viral packaging constructs *pUMVC, pCMV-VSV-G*, and *psPAX2* [14] were all purchased from Addgene (Cambridge, MA, USA). DNA oligos designed against the mouse *Brm* and *Baf170* cDNA (shBrm_1, shBrm_2 and shBaf170_1, shBaf170_2) and scramble sequence (shCtl) (Table S5.1) were subcloned into pLKO.1-puro vector. All DNA subcloning was performed using the standard restriction-enzyme digestion or Infusion PCR Cloning Kit (Clontech, Mountain View, CA, USA) and expression constructs of shBrm and shBaf170 were verified by DNA sequencing. The human embryonic kidney cell line 293T for viral packaging was obtained from Invitrogen (Grand Island, NY, USA).

5.3.2. Cell Culture, Viral Preparation, and Reprogramming Assay

OG-MEFs, as well as MEFs from CD1 mice were generated from E13.5 embryos as described [15]. OG-MEFs up to passage 4 were used for reprogramming. Briefly, *pMXs, pMCs*- or *pLKO.1*-constructs, together with packaging vectors *pUMVC* (for retrovirus), *psPAX2* (for lentivirus), and *pCMV-VSV-G* plasmids, were co-transfected into 293T cells according to Addgene protocols. Retrovirus OKSM and lentiviral short hairpin RNA were collected 48 and 72 hours after transfection. The iPSC induction from OG-MEFs using viral OKSM and reprogramming medium was conducted as described [15]. Briefly, OG-MEFs were plated on 6-

well-plates and transduced (Day 0) with retroviral OKSM with 10 mg/ml polybrene (American Bioanalytical, Natick, MA, USA). After 24 hours of viral transduction (Day 1), cells were trypsinized and passaged onto two 2-well-plates pre-seeded with mitomycin C treated CD1 MEF feeders. Lentiviral shBrm and shBaf170 constructs were then transduced to reprogrammed cells at different stages after OKSM transduction according to experiment design (Fig. S5.1A). Dox was added after lentiviral shBrm and shBaf170 infection, and GFP+ colonies were counted 3 weeks after OKSM transduction under a Nikon fluorescence microscope. GFP+ iPSC colonies were picked 3 weeks after retroviral transduction and expanded in 2i/LIF medium until passage 5 when further characterization was conducted. The average size of colonies (areas; mean s $\pm$ SD) was measured by using the ImageJ software *(http://imagej.nih.gov/ij/index.html)*.

Somatic cells partially reprogrammed by transfection of OKM (pre-iPSCs) were generated from our previous study [16]. These cells have ESC colony morphology but are GFP negative. For complete reprogramming of pre-iPSCs to iPSCs, shBrm and shBaf170 were transduced (Day 0) and GFP+ colonies were counted 10 days after transfection (Fig. S5.1B) and mRNA of these reprogrammed cells were also subjected for quantitative real-time reverse transcription – polymerase chain reaction (qRT-PCR) analysis. shCtl constructs were transduced to reprogrammed cells and pre-iPSCs as controls.

For Jak inhibition, OG-MEFs were reprogrammed by OKSM and cultured with either DMSO control or a low dosage (1 mM) of Jaki, stating on Day 3 of reprogramming (Fig. S5.1A). mRNA of these cells were then collected for gene expression analysis at Day 18.

5.3.3. FACS analysis

Cells reprogrammed from pre-iPSCs were trypsinized and Flow cytometry analysis was performed using BD FACS ARIA III. All data analysis was carried out using the Flowjo software.

5.3.4. Embryoid Body (EB) Formation

Established iPSC cell lines at passage 6 were grown on CD1 MEF feeders. After two days, the cells were trypsinized and replated to the original plate for 2 hours to allow MEFs to attach. The iPSCs remained in the medium were collected and subsequently plated onto Petri-dishes containing DMEM + 10% FBS (Invitrogen) without LIF. Upon 7 days of differentiation, EBs formed and were transferred to 0.1% gelatin-coated cell culture dishes (Invitrogen). The cells were allowed to re-attach and to continue differentiation for another 7 days before proceeding for RNA extraction, qRT-PCR, or immunostaining as described below.

5.3.5. Alkaline Phosphatase (AP) Staining and Immunostaining

AP staining was performed using a Vector Red Alkaline Phosphate Substrate Kit I (Vector Laboratories, Burlingame, CA, USA) according to manufacturer's instruction. For immunostaining, cells were grown on 12 mm glass coverslips (Fisher Scientific, Waltham, MA, USA) in 6-well-plates containing CD1 MEFs as feeders. Cells were fixed in 4% paraformaldehyde with 1% sucrose in PBS for 15 min at room temperature. The cell membranes were permeabilized with 0.5% TX-100 in PBS-T, then incubated for 2 hours at 37ºC in 5% donkey or goat serum with mouse anti-SSEA1 IgM (1:100), rabbit anti-Sox2 IgG (1:100), or rabbit anti-Nanog IgG (1:100) (All from Millipore), or rabbit anti-Oct4 antibody (1:100, Santa Cruz, Dallas, Texas, USA), washed in PBS-T, and then incubated with Alexa Fluor 594 conjugated donkey anti-rabbit or goat anti-mouse secondary antibodies (1:500, Invitrogen).

Cells incubated in serum without primary antibodies were kept as negative controls and R1-ESCs were stained as positive controls. After the washes, cells were counterstained with DAPI and mounted under coverslips. Fluorescence images were taken using a Nikon fluorescence microscope.

5.3.6. Western Blot Analyses

OG-MEFs were transduced with lentiviral-vector, shCtl, shBrm or shBaf170 and cultured in medium containing 10% FBS for four days. OG-MEFs, Pre-iPSCs, iPSCs and R1-ESCs were cultured as described above. Total cellular proteins were extracted using RIPA buffer (Thermo Scientific, Pittsburg, PA, USA) with 1 x proteinase and phosphatase inhibitors (Thermo Scientific). Proteins were quantified with Pierce™ BCA Protein Assay Kit (Thermo Scientific), and subjected to 10% SDS-PAGE gel electrophoresis using BioRad mini-gel system and subsequently transferred to PVDF membranes.

The blotted membranes were then blocked with 5% non-fat dry milk in TBS-T and incubated with primary antibodies at 4ºC overnight. The antibodies used were as follows: anti-*Brm* (1:1,000, Cell Signaling, Danvers, MA, USA), anti-*Baf170* (1:1,000, Cell Signaling), and anti-*GAPDH* (1:2,000, Abcam Inc., Cambridge, MA, USA). Membranes were then washed and blotted with HRP conjugated goat anti-mouse or goat anti-rabbit secondary antibodies (1:5,000, Santa Cruz). Blotting signals were detected by chemiluminescence using Pierce ECL Western-Blot Substrate (Thermo Scientific) and quantified by ChemiDoc XRS+ imaging system with Image Lab™ Software (BIO-RAD).

5.3.7. Quantitative Real Time-Reverse Transcription Polymerase Chain Reaction

Total RNA was extracted using the RNeasy Mini kit (Qiagen, Valencia, CA, USA) and reverse-transcribed using the SuperScript III Reverse Transcription Kit (Invitrogen). qRT-PCR was performed with specific primers (Table S5.2) using SYBR Green PCR Master Mix (Applied Biosystems, Grand Island, NY, USA) and the ABI 7500 Fast instrument. Combined and endogenous levels of *Oct4*, *Sox2* and *Klf4* were distinguished by primers designed at different regions. Data were analyzed using the 7500 software version 2.0.2 provided with the instrument. Quantification was normalized to the endogenous glyceraldehyde 3-phosphate dehydrogenase *(GAPDH)* as the internal control and relative mRNA expressions were calculated using R1-ESC as the reference.

5.3.8. Statistical Analysis

Data were analyzed using One Way ANOVA with Tukey's multiple comparisons or the Student's t-test. All experiments were performed at least twice ($N \geq 2$). Values in figures were presented as mean ± standard deviation (SD). A p-value < 0.05 (*) or < 0.01 (**) was considered statistically significant.

5.4. Results

5.4.1. Knockdown of Brm or Baf170 at Different Stages Promotes Reprogramming

To investigate the roles of *Brm* and *Baf170* in reprogramming and differentiation, we first measured their mRNA levels in OG-MEFs, OG-MEF-derived iPSC clones, OKM-reprogrammed pre-iPSCs and R1-ESCs. The levels of mRNAs for *Brm* and *Baf170* in OG-MEFs and in pre-iPSCs were 5-8 times higher than those in iPSCs and R1-ESCs (Figure 5.1A). We further found that subunits of esBAF (*Baf60a, Baf200, Baf250a, Brg1 and Baf155*) were up-regulated in iPSCs and R1-ESCs compared to OG-MEFs and pre-iPSCs (Figure S5.2), suggesting that BAF complex undergoes subunit changes while cells transit from differentiated to pluripotent states. Western blot analyses also revealed a dramatic higher level of proteins for *Brm* and *Baf170* in OG-MEFs and pre-iPSCs than those in iPSCs and R1-ESCs (Figure 5.1D). The inverse correlation between *Brm*, *Baf170* and the state of differentiation, especially in partially reprogrammed pre-iPSCs, suggests that *Brm* and *Baf170* may interfere with somatic reprogramming. We then asked whether knocking down *Brm* or *Baf170* would improve reprogramming. Using lentiviral shRNA constructs designed specifically against either *Brm* or *Baf170*, we were able to knock down *Brm* and *Baf170* mRNA levels in OG-MEFs by more than 95% (Figure 5.1B, C, Figure S5.3A, B), which were also confirmed at protein levels by Western blot analyses (Figure 5.1E, F). We then reprogrammed OG-MEFs with OKSM (Day 0) in the presence of either shBrm or shBaf170 or shCtl on Day 3 and cultured the cells in 2i/LIF medium. Two weeks after OKSM induction, dome-shaped, ESC-like GFP+ colonies started to appear in all treatment groups. Knocking down of *Brm* resulted in a significant increase in GFP+ colonies compared to controls (Figure 5.1G). On the contrary, knocking down *Baf170* had minimal impact on GFP+ colony formation (Figure 5.1G). To further dissect the effect of somatic BAF

components on reprograming, we knocked down *Brm* and *Baf170* at later stages, i.e., on Days 6 or 9. Interestingly, similar results were obtained when treatments were given on Day 6 (Figure 5.1H) while knocking down *Baf170* on Day 9 produced a significant gain on reprogramming efficiency (Figure 5.1I). Collectively, these observations demonstrate that *Baf170* inhibits reprogramming at a later stage while removal of *Brm* at any stage improves reprogramming (Figure 5.1G, H and I).

Subsequently we characterized the GFP+ colonies (Figure 2A) for their reprogramming status. Interestingly, we found that the colonies induced by shBrm were significantly smaller than those treated with shBaf170 or shCtl (Figure 5.2B). Furthermore, we observed that GFP+ colonies induced by shBrm and shBaf170 shared the following properties with the control ESC line, R1: 1) formation of tight, compact colonies with strong alkaline phosphatase (AP) activity (Figure5.2C); 2) propagation in 2i/LIF medium and positive immunostaining for pluripotent genes and surface markers including *Oct4, Sox2* and SSEA1 (Figure 5.2D); and 3) formation of EBs in vitro (Figure 5.2E) and differentiation to cells expressing markers of the three germ layers (Data not shown). Of note, no differences were found in the numbers of AP, *Sox2*, *Oct4*, and SSEA1 positive colonies or EBs between shBrm or shBaf170 and shCtl treated cells expect for a slightly smaller EB size in shBrm and shBaf170 treated cells. The reprogramming was also relatively complete because the retroviral transgenes were virtually silenced in the induced iPS colonies (Figure 5.2F). Taken together, the above data demonstrate that inhibition of *Brm* or *Baf170* is critical for efficient reprogramming in a stage-specific fashion.

5.4.2. Knockdown of Brm or Baf170 Promotes Complete Reprogramming of pre-iPSCs

Because we had determined that knockdown of *Brm* and *Baf170* is critical for efficient late-stage reprogramming of OG-MEFs, and also *Brm* and *Baf170* were highly expressed in pre-

iPSCs (Figure 5.1A and D), we asked whether their inhibition is also required for the complete reprogramming of these cells. We knocked down *Brm* and *Baf170* in two stable pre-iPS cell lines from our previous study [16], and found that suppressing *Brm* and *Baf170* increased the numbers of GFP+ colonies by 5.5- and 11-fold, respectively, compared to controls (Figure 5.3A). Similarly, FACS analysis showed that reduced *Brm* and *Baf170* resulted in 49.1% and 70.6% GFP+ cells, respectively, compared to controls (5.57%; Figure 5.3B). Characterization of GFP+ colonies from the treated pre-iPSCs demonstrated that these colonies grew readily in 2i/LIF medium, maintained bright GFP expression and ESC-like colony morphology under repeated passaging, stained positive for alkaline phosphatase activity and expressed ESC-specific gene markers such as *Oct4, Sox2, Nanog* and SSEA1 (Data not shown). These data corroborated with those in OG-MEF reprogramming that inhibition of either *Brm* or *Baf170* efficiently converted pre-iPSCs to iPSCs, while *Baf170* knockdown at this late stage of reprogramming is more effective than *Brm* knockdown.

5.4.3. Knockdown of Brm or Baf170 Promotes Reprogramming by Activating the "Pluripotency Circuitry"

As depletion of *Brm* or *Baf170* enhanced reprogramming efficiency and promoted complete late stage reprogramming, we sought to investigate the mechanisms of these effects by examining the expression of ESC/iPSC-specific genes. We found that early removal of *Baf170* (Days 3 and 6) did not affect the expression of core pluripotency genes (Figure 5.4A and B), while knockdown of *Baf170* on Day 9 resulted in significant up-regulation of *Sox2, Nanog, Esrrb* and *Tbx3* (Figure 5.4C). On the contrary, inhibiting *Brm* at any stage of reprogramming significantly increased the expression of *Nanog, Esrrb* and *Tbx3* (Figure 5.4A, B and C). *Brm* inhibition also elicited stage-specific up-regulation of pluripotent genes. For example, *Oct4* and

Rex1 were induced on Day 3, *Sox2* on Day 6, and *Sox2* and *Klf4* on Day 9. Additionally, we also showed that *Sox2, Nanog, Rex1, Esrrb* and *Tbx3* were significantly up-regulated when *Brm* and *Baf170* were inhibited in pre-iPSCs (Figure 5.4D). Taken together, these results demonstrated that inhibition of *Brm* or *Baf170* promotes reprogramming efficiency by activating the "pluripotency circuitry" in a stage-specific manner.

5.4.4. Inhibition of Brm or Baf170 Promotes Jak/Stat3 Signaling during Complete Reprogramming

It has been reported that esBAF facilitates pluripotency by conditioning the genome for *LIF/Stat3* signaling [11]. We have also shown that *Jak/Stat3* signaling plays an essential role in epigenetic regulation of late stage somatic cell reprogramming [16]. Here we sought to determine if somatic BAF components are involved in the intricate interplay of *Stat3* signaling during reprogramming. We analyzed *Stat3* mRNA transcripts 10 days after shBrm or shBaf170 transduction of pre-iPSCs. Interestingly, we found that expression of *Stat3* was significantly increased by inhibiting *Brm* or *Baf170* in pre-iPSCs (Figure 5.5A). We further revealed that knockdown of *Brm* or *Baf170* significantly stimulated the expression of *Socs3*—the direct target of *Stat3* (Figure 5.5A). These results suggest that the presence of *Brm* and *Baf170* inhibited the *Stat3* pathway during reprogramming.

Retroviral silencing is a prerequisite for pluripotency establishment in retroviral transgene-mediated reprogramming and is achieved by *de novo* DNA methylation through DNA methyltransferases (Dnmts) 3a, 3b and Dnmt3L [16-20]. We have also shown previously that *Jak/Stat3* activity facilitates *Dnmt3L* expression. This, in turn, stimulates *de novo* DNA methylation which silences retroviral transgenes and possibly lineage commitment genes [16]. Examining the expression of *Dnmts* here revealed that *Brm* and *Baf170* knockdown drastically

increased the expression of *de novo* methyltransferase *Dnmt3a* and *3L* (Figure 5.5B). These results suggested that reducing somatic *Brm* or *Baf170* stimulates *de novo* DNA methylation, which is correlated with the elevated *Stat3* pathway. Therefore somatic BAF reduction may coordinate with the *Stat3*-regulated epigenetic network for pluripotency establishment.

To further dissect the intricate interplay of *Brm/Baf170* and *Jak/Stat3* signaling, we examined whether inhibiting *Jak/Stat3* would affect *Brm* and *Baf170*. We reprogrammed OG-MEFs and cultured them with either DMSO control or a low dosage (1 mM) of the reversible Jak inhibitor I (Jaki), starting on Day 3 of reprogramming. Jaki treatment virtually blocked the formation of all GFP+ colonies, however, this did not change the number of GFP negative colonies formed (data not shown). Furthermore, mRNAs for components of esBAF, *Baf60a, Baf250a, Brg1, Baf155, Baf47, Baf200, Baf60b* and *Baf57*, were consistently expressed and unchanged while the somatic BAF components, *Brm* and *Baf170*, were significantly up-regulated in Jaki-treated cells compared to DMSO-treated control cells (Figure 5.5C). To minimize the heterogeneity of cells in the analysis, three typical colonies from each treatment were collected three weeks after viral transduction and analyzed for their gene expression at passage 2 (P2). A dramatic up-regulation of mRNAs for *Brm* and *Baf170* were consistently observed in Jaki-treated cells, compared to the DMSO controls (Figure 5.5D). Thus, a positive feedback loop exists between *Jak/Stat3* and the inhibition of *Brm* and *Baf170* during ground state pluripotency establishment.

5.5. Discussion

Induced pluripotent stem cells can be generated from somatic cells by ectopic expression of transcription factors such as *Oct4, Sox2, Klf4* and *Myc* (OKSM) [1], but the efficiency remains low. A variety of chromatin modifiers, such as *Brg1/Baf155*, *Utx* and *MBD3/NuRD* have been implicated in facilitating epigenetic changes leading to authentic iPSC reprogramming [12,21-24]. In this study we identified a negative role by *Brm* and *Baf170* on OKSM-mediated reprogramming of OG-MEFs. We found that shRNA-mediated knockdown of *Brm* or *Baf170* led to an increase in reprogramming efficiency. Specifically, *Baf170* seems to be more inhibitory to reprogramming at the late stage, when the pluripotency network is becoming more stably established, while *Brm* inhibits reprogramming at all phases. Also interesting to note was that colonies induced by *Brm* knockdown were significantly smaller than those from *Baf170* knockdown. Sizes of ES colonies and EBs have been shown to affect their differentiation trajectories [25] through different gradients of signaling molecules [26]. The size difference observed may reflect a mechanism that cells employ to regulate their fates upon changes of BAF compositions.

Components of esBAF have been shown to be important in both maintenance of mESCs and iPSC induction. Deficiency in *Brg1, Baf47, Baf155,* or *Baf250* impaired the ability of mESCs to proliferate and to maintain pluripotency [8-11,27,28]. *Brg1* and *Baf155*, combined with *Oct4*, *Sox2*, *Klf4* and *c-Myc*, can synergistically increase reprogramming efficiency [12]. Because iPSCs have similar properties with ESCs, it is logical to infer that homologues of *Brg1* and *Baf155*, *Brm* and *Baf170*, obstruct reprogramming, which was proven in the current study. Our results are also in agreement with the recent finding that levels of *Brm* increase during differentiation of ESCs [10]. Moreover, the dynamic composition of BAF seems to correlate with

the differentiation status of the cells, i.e., *Baf170* is more important in differentiated cells, whereas *Baf155* is more important in undifferentiated ES cells [8,10]. Competition between *Baf170* and *Baf155* subunits within the BAF complex was also observed during progression of neurogenesis and affected euchromatin structure and thereby modulated the binding efficiency of their targets [29]. Our results expanded this notion that BAF complex undergoes subunit changes while cells transit from differentiated to pluripotent states, i.e., acquisition of esBAF and disposition of the somatic BAF components (*Brm* and *Baf170*). Moreover, it suggested somatic BAF inhibits the formation of esBAF, therefore serves as a barrier during reprogramming.

We discovered that removal of *Brm* and *Baf170* aids in full reprogramming. At selective loci pre-iPSCs and somatic cells possess similar chromatin structures which are condensed and arrested epigenetically during reprogramming [30,31]. Pre-iPSCs can be completely reprogrammed to pluripotency by overexpressing *Sox2* or *Nanog* [32]. The pre-iPSC stage is an epigenetically stable landmark along the journey of reprogramming and these cells can progress toward the authentic iPSCs through the removel of epigenetic determinants for the intermediate state [31]. Here we provide a new insight that the presence of the somatic *Brm* and *Baf170* are potential epigenetic barriers during reprogramming. First, high levels of both *Brm* and *Baf170* are present in pre-iPSCs. Secondly knockdown of *Brm* or *Baf170* up-regulated the expression of pluripotency-related core transcription factors such as *Sox2, Nanog, Rex1, Tbx3* and *Esrrb*, suggesting that the somatic BAF complex may be inhibitory to the expression of pluripotency related genes. Interestingly, during mouse somatic reprogramming, *Brg1* and *Baf155* were recruited by *Oct4* in order to relax chromatin structure and facilitate the binding of other transcription factors that enhance reprogramming [12]. Genome-wide co-localization of *Brg1* with

Oct4, Sox2 and *Nanog* [27,33] further suggests that the somatic components of BAF are not compatible with reprogramming.

An important contribution of this study was to provide evidence that the suppressed *Brm* and *Baf170* may coordinate with the *Stat3*-regulated epigenetic network during pluripotency establishment. Activation of Jak/Stat3 signal pathway is essential for maintaining mouse ESC pluripotency [34] and for reactivating the endogenous pluripotency network in somatic cells [16,35]. Although the role of SWI/SNF complex in *Stat3* signaling was discovered previously, the finding was only limited for *Brg1* [11,36,37]. More recently, binding sites of *Stat3* and esBAF had been shown to frequently overlap in the ESC genome [8,27]. Moreover, Ho et al. [11] provided evidence that esBAF is required to establish chromatin accessibility at *Stat3* binding targets and to prepare these sites to respond to LIF signaling, therefore allowing in vitro preservation of ground-state pluripotency. These studies established a strong connection between esBAF complex and *Stat3* activities. For the first time, we demonstrated here a link between *Brm, Baf170* and *Stat3* activities. We not only showed that knocking down *Brm* and *Baf170* stimulated *Stat3* activity, but also inhibiting *Jak/Stat3*, in turn, increased the expression of *Brm* and *Baf170* during reprogramming, thus revealing an interplay between *Stat3* and somatic BAF during reprogramming. In addition, we demonstrated that reducing somatic *Brm* and *Baf170* drastically increased the expression of *de novo* methyltransferase *Dnmt3a* and *3L*. These changes were also seen when the *Stat3* pathway was activated [16]. Therefore *Brm* and *Baf170* inhibition integrates in the *Stat3*-regulated epigenetic network for pluripotency establishment. Our results provide new insights that *Stat3* may promote esBAF formation by inhibiting somatic BAF. As *Stat3* activation is key to ground-state pluripotency establishment during somatic cell reprogramming [16], our data here indicate that inhibiting *Brm* and *Baf170* coordinates with the determining effects

of *Stat3* in pluripotency establishment, and a mutually suppressive interaction between *Stat3* and somatic BAF is one of the mechanisms involved.

Figure 5.1. Knockdown of Brm or Baf170 promotes reprogramming in a stage-specific fashion. (A) The levels (means + SD) of Brm and Baf170 mRNAs are low in induced pluripotent stem cells (iPSCs) and murine embryonic stem cells (R1-ESCs), but high in OG-MEFs, pre-iPSCs induced by ectopic Oct4/Klf4/c-Myc (OKM) infection (n = 3). The OG-MEFs are mouse embryonic fibroblasts harboring the green florescence protein (GFP) driven by the Oct4-promoter. Transfection of OG-MEFs with lentiviruses expressing different small hairpin RNAs against Brm (shBrm; B) and Baf170 (shBaf170; C) reduced Baf170 and Brm by 95% and 96%, respectively. (D) Western blots showing the protein levels of Baf170 and Brm in OG-MEFs, Pre-iPSCs, iPSCs and R1-ESCs. Both Baf170 and Brm are present in OG-MEFs and Pre-iPSCs, while nearly absent in iPSCs and R1-ESCs. GAPDH was used as the loading control. Transfection of OG-MEFs with shBrm (E) and shBaf170 (F) efficiently reduced the protein levels of Brm and Baf170, respectively. Knocking down Brm on Days 3 (G), 6 (H) and 9 (I) and knocking down Baf170 only on Day 9 (I) significant increased the numbers of GFP+ colonies (means + SD) induced by OKSM (** P values < 0.01; n = 3).

Fig. 5.1

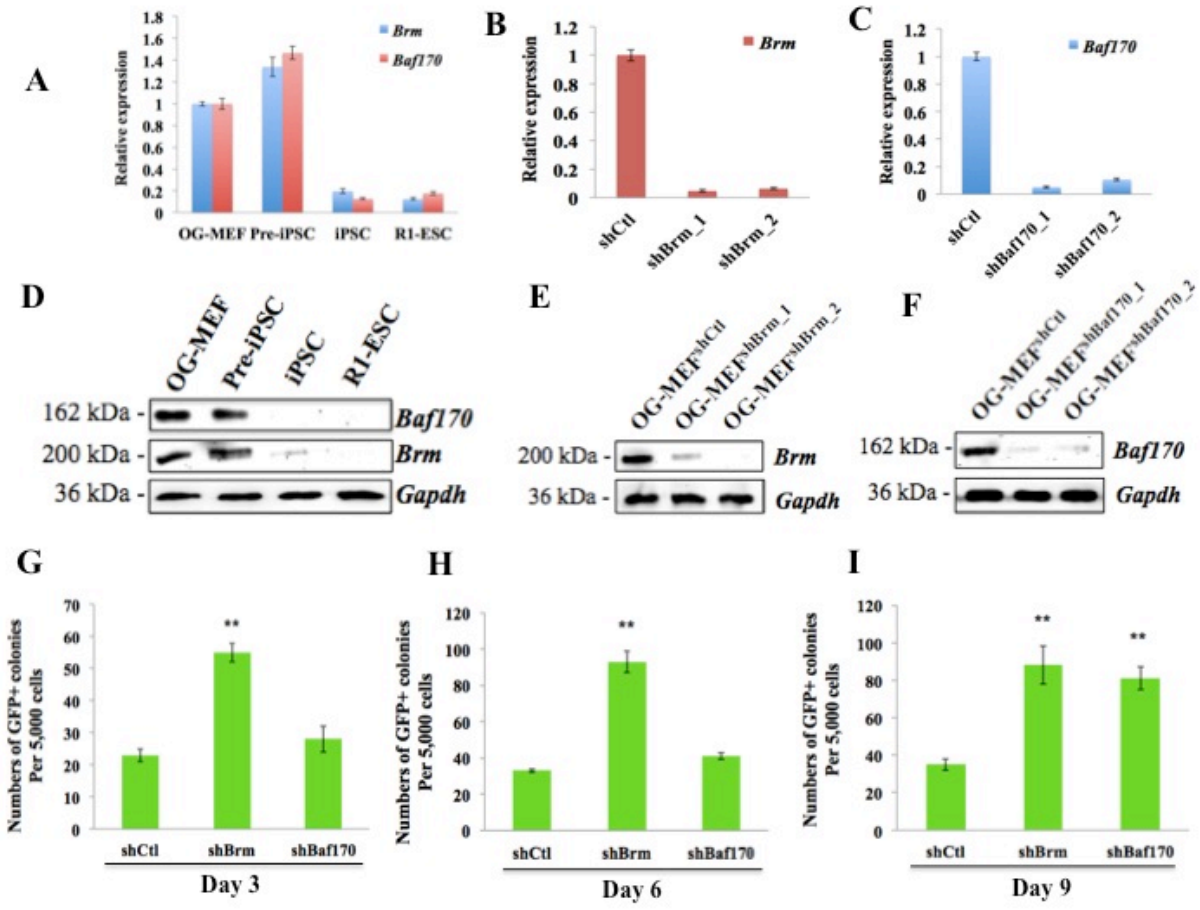

A
Relative expression
Brm
Baf170
OG-MEF Pre-iPSC iPSC R1-ESC
B
Relative expression
Brm
shCtl
shBrm_1
shBrm_2
C
Relative expression
Baf170
shCtl
shBaf170_1
shBaf170_2
D
OG-MEF
Pre-iPSC
iPSC
R1-ESC
162 kDa -
Baf170
200 kDa -
Brm
36 kDa -
Gapdh
E
OG-MEFshCtl
OG-MEFshBrm_1
OG-MEFshBrm_2
200 kDa -
Brm
36 kDa -
Gapdh
F
OG-MEFshCtl
OG-MEFshBaf170_1
OG-MEFshBaf170_2
162 kDa -
Baf170
36 kDa -
Gapdh
G
Numbers of GFP+ colonies Per 5,000 cells
**
shCtl
shBrm
shBaf170
Day 3
H
Numbers of GFP+ colonies Per 5,000 cells
**
shCtl
shBrm
shBaf170
Day 6
I
Numbers of GFP+ colonies Per 5,000 cells
**
**
shCtl
shBrm
shBaf170
Day 9

Figure 5.2. Cells reprogrammed by OKSM and *Brm*- or *Baf170*-knockdown are pluripotent. (A) Representative iPSC colonies (passage 5) induced by OKSM plus shBrm or shBaf170 and cultured in 2i/LIF medium. The GFP+ colonies indicate activation of the internal *Oct4* gene (Scale bar = 250 μm, ** P values < 0.01; n = 3). The average size of colonies (areas; means ± SD) induced by shBrm is significantly smaller than that by shBaf170 and shCtl (B). The induced colonies stained strongly for AP (C; passaged 5) and pluripotency markers including SSEA1, *Oct4 and Sox2* (D; bar = 25 μm; passage 6). The induced colonies also formed EBs after 7 days of differentiation (E; bar = 250 μm). Nuclei of the cells were counterstained with DAPI. All GFP+ colonies induced by shBrm and shBaf170 shared similar properties with the control R1-ESCs. (F) Similar levels of combined and endogenously (*Oct4*, Sox2** and *Klf4**) expressed transcription factors in OKSM-infected OG-MEFs also transfected with shBaf170 or shBrm were observed in the reprogrammed cells, suggesting the exogenous induction factors were silenced. Values were normalized with glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) and relative to R1-ESCs (n = 3).

Fig. 5.2

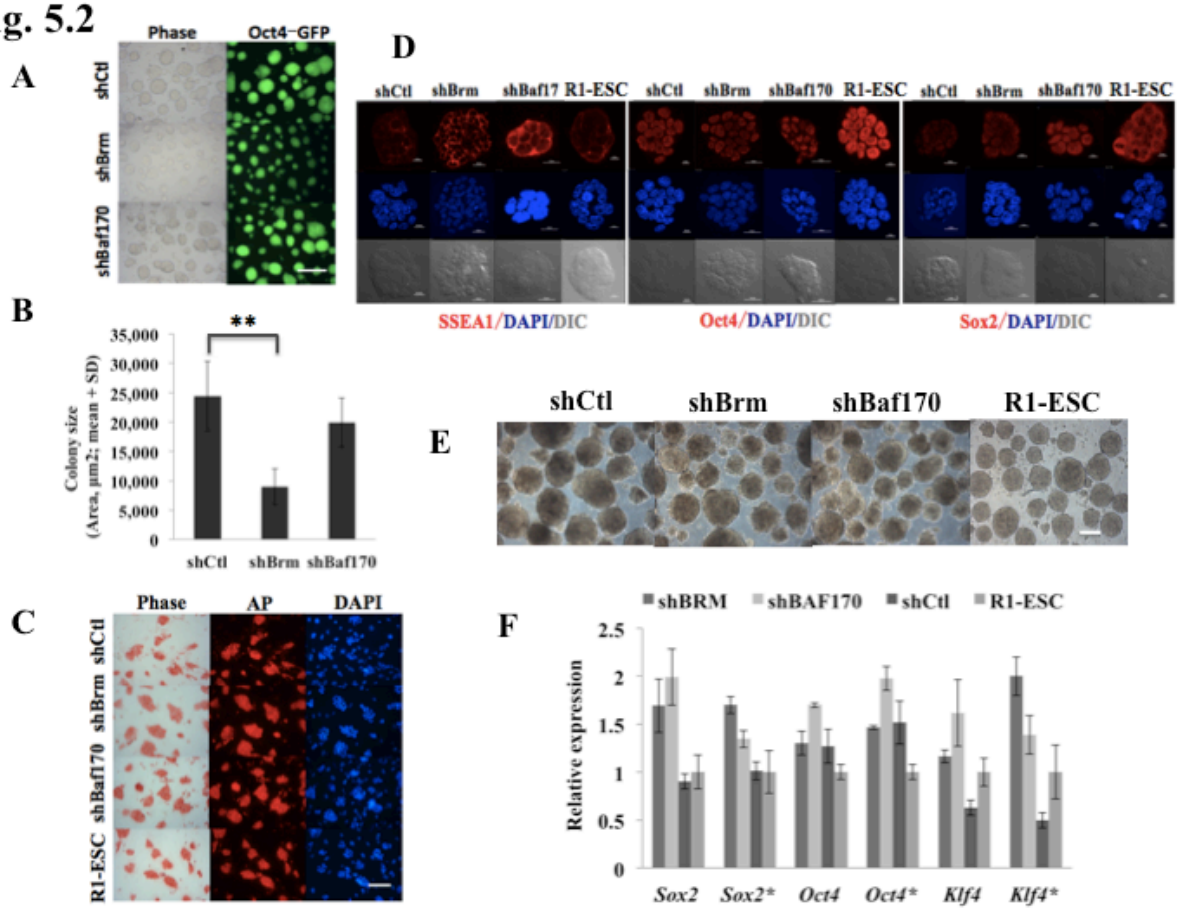

A
Phase
Oct4-GFP
shCtl
shBrm
shBaf170
B
**
Colony size (Area, μm2; mean + SD)
35,000
30,000
25,000
20,000
15,000
10,000
5,000
0
shCtl
shBrm
shBaf170
C
Phase
AP
DAPI
shCtl
shBrm
shBaf170
R1-ESC
D
shCtl
shBrm
shBaf17
R1-ESC
shCtl
shBrm
shBaf170
R1-ESC
shCtl
shBrm
shBaf170
R1-ESC
SSEA1/DAPI/DIC
Oct4/DAPI/DIC
Sox2/DAPI/DIC
E
shCtl
shBrm
shBaf170
R1-ESC
F
shBRM
shBAF170
shCtl
R1-ESC
Relative expression
2.5
2
1.5
1
0.5
0
Sox2
Sox2*
Oct4
Oct4*
Klf4
Klf4*

Figure 5.3. Knocking down *Brm* or *Baf170* promotes complete reprogramming of pre-iPSCs. The numbers of GFP+ colonies (means ± SD; A) significantly increased 10 days after pre-iPSCs were infected with shBrm or shBaf170. Mean values of three independent experiments from two pre-iPS cell lines are shown (**P values < 0.01; n = 3). The reprogrammed colonies contained different percentages of GFP+ cells as analyzed by fluorescence-activated cells sorting (B; n = 2).

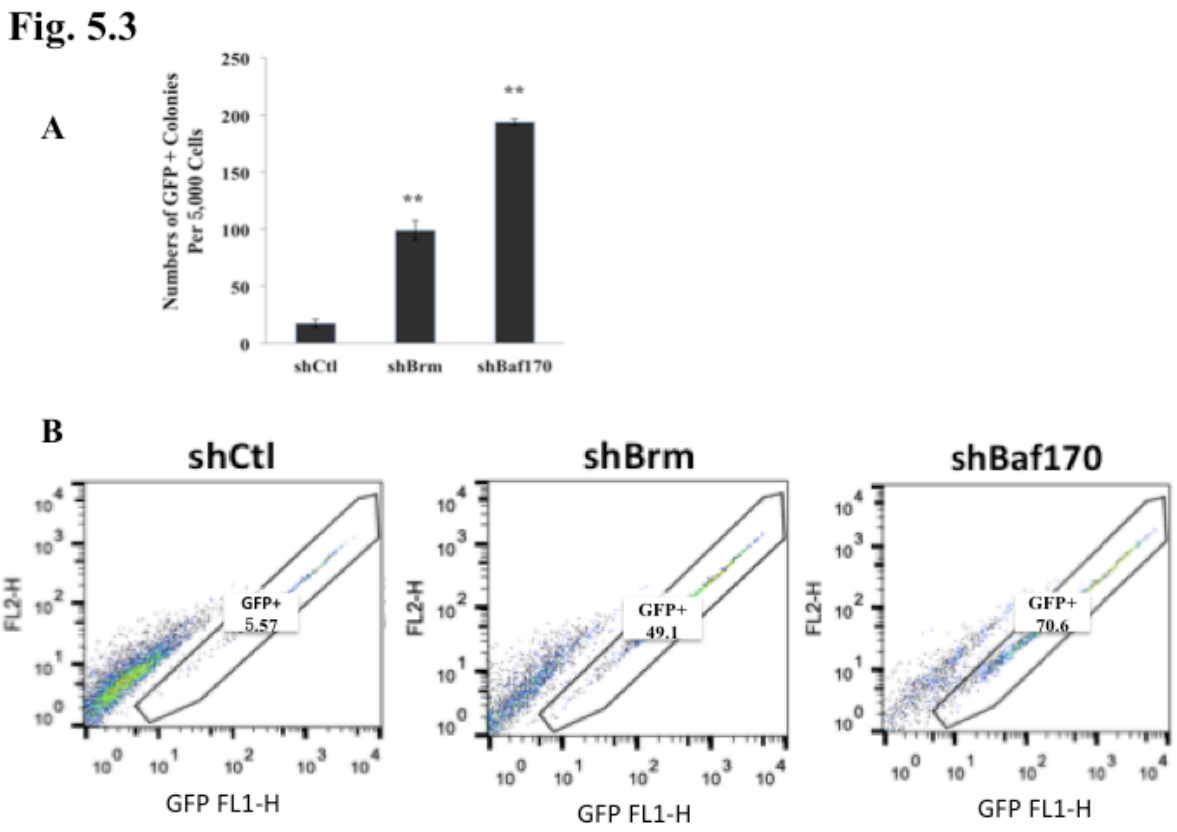


Figure 5.4. Knocking down *Brm* or *Baf170* promotes reprogramming by activating the "pluripotency circuitry". Relative levels of mRNA for selected ESC-specific genes in OG-MEFs transfected with shBaf170, shBrm or shctl on Days 3 (A), 6 (B) and 9 (C) of OKSM reprogramming. *Oct4**, *Sox2** and *Klf4** represent mRNA expressed endogenously. Values were normalized with glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) and relative to shCtl (*: *p* values <0.05; n = 3). (D) Relative levels of mRNA for selected ESC-specific genes in pre-iPSCs 10 days after transfection with shBaf170, shBrm or shCtl (n = 3). Results were normalized with *GAPDH* and relative to shCtl.

Fig. 5.4

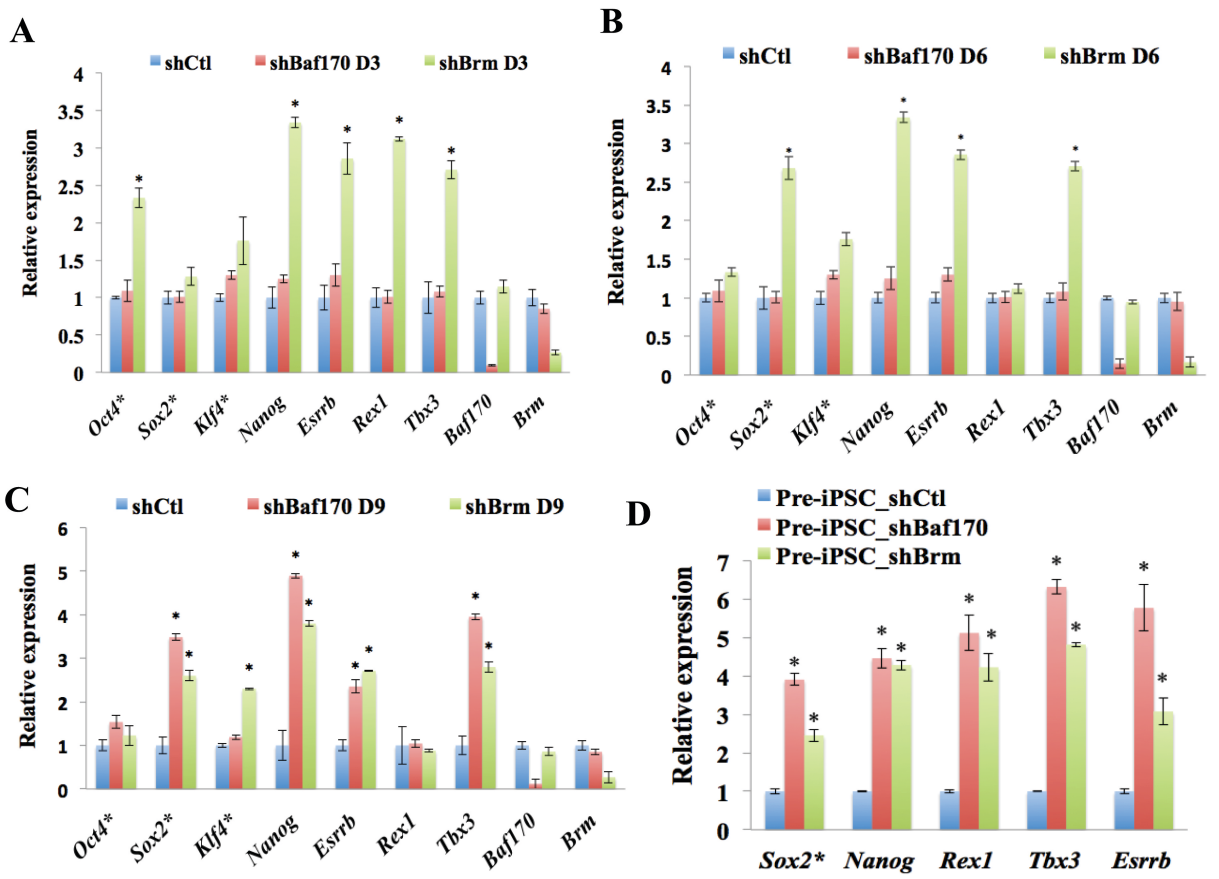
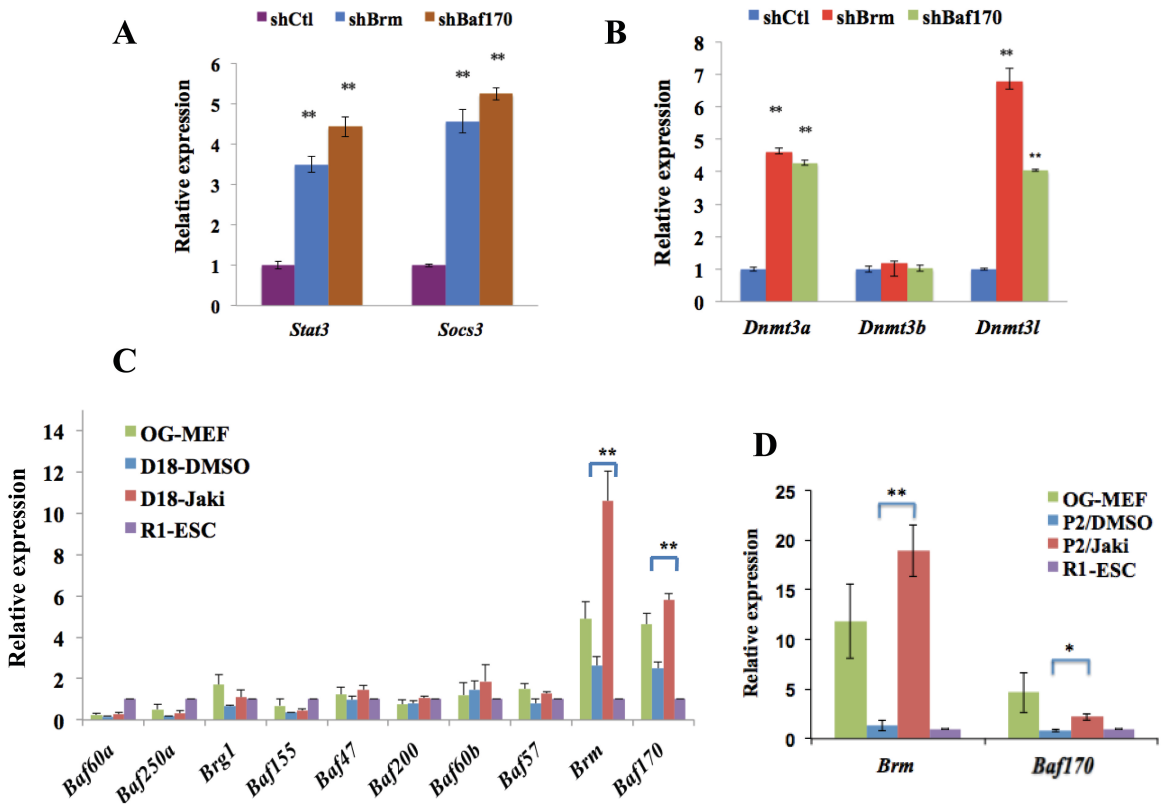


Figure 5.5. Inhibition of *Brm* or *Baf170* promotes Jak/Stat3 signaling during complete reprogramming. Relative levels of mRNA for (A) *Stat3* and its direct target, *Socs3,* as well as *do novo Dnmts* (B) in pre-iPSCs 10 days after transfected with shBaf170 or shBrm. Results were normalized with *GAPDH* and related to shCtl (**: *p* values <0.01; n = 3). (C) Relative levels of mRNA for different subunits of BAF complex in OG-MEFs, R1-ESCs, and OKSM-transduced OG-MEFs 18 days after initial OKSM viral transduction. The reprogrammed cells were treated with either DMSO or 1 μM Jaki. Results were normalized with *GAPDH* and relative to those in R1-ESCs (*: p values <0.05; n = 3). (D) Relative levels of mRNA for *Brm* and *Baf170* in OG-MEFs, R1-ESCs, and passage two (P2) of reprogrammed cells. The reprogrammed cells were treated with either DMSO or 1 μM Jaki. Results were normalized with *GAPDH* and relative to those in R1-ESC (*: p values <0.05; n = 3).

Fig. 5.5


A
shCtl
shBrm
shBaf170
Relative expression
Stat3
Socs3
B
shCtl
shBrm
shBaf170
Relative expression
Dnmt3a
Dnmt3b
Dnmt3l
C
OG-MEF
D18-DMSO
D18-Jaki
R1-ESC
Relative expression
Baf60a
Baf250a
Brg1
Baf155
Baf47
Baf200
Baf60b
Baf57
Brm
Baf170
D
OG-MEF
P2/DMSO
P2/Jaki
R1-ESC
Relative expression
Brm
Baf170

5.6. Supplementary Information

Table S5.1: Primer sequences for sub-cloning of lentiviral constructs.

Lentiviruses	Sequences (5' – 3')
shBrm_1F	CCGGGCTCGCATCGCTCATAGGATACTCGAGTATCCTATGAGCGATGCGAGCTTTTTG
shBrm_1R	AATTCAAAAAGCTCGCATCGCTCATAGGATACTCGAGTATCCTATGAGCGATGCGAGC
shBrm_2F	CCGGGGAAGAGAACGAACAGTCAGACTCGAGTCTGACTGTTCGTTCTCTTCCTTTTTG
shBrm_2R	AATTCAAAAAGGAAGAGAACGAACAGTCAGACTCGAGTCTGACTGTTCGTTCTCTTCC
shBaf170_1F	CCGGGAACCGCCAACCAACAAGTCTCTCGAGAGACTTGTTGGTTGGCGGTTCTTTTTG
shBaf170_1R	AATTCAAAAAGAACCGCCAACCAACAAGTCTCTCGAGAGACTTGTTGGTTGGCGGTTC
shBaf170_2F	CCGGGAACTTCATGATTGACACTTACTCGAGTAAGTGTCAATCATGAAGTTCTTTTTG
shBaf170_2R	AATTCAAAAAGAACTTCATGATTGACACTTACTCGAGTAAGTGTCAATCATGAAGTTC
shCtl_F	CCGGGCAGTCTACGTCGGCTTACAACTCGAGTTGTAAGCCGACGTAGACTGCTTTTTG
shCtl-R	AATTCAAAAAGCAGTCTACGTCGGCTTACAACTCGAGTTGTAAGCCGACGTAGACTGC

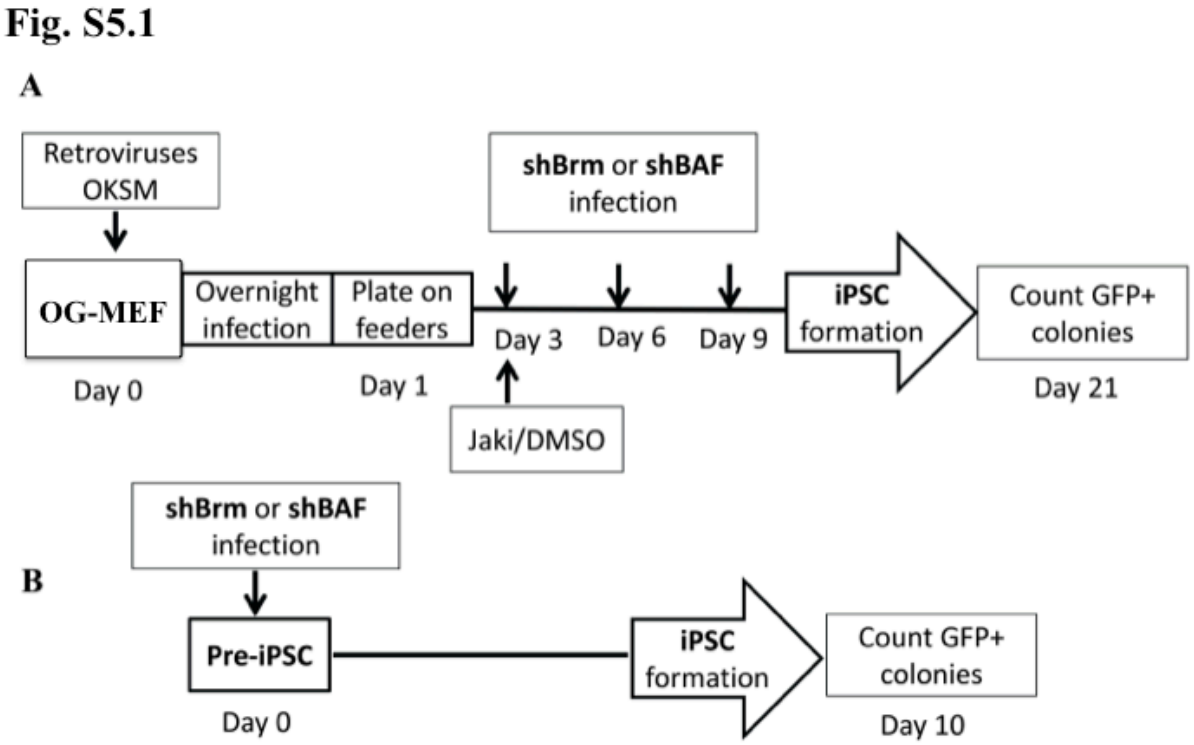
Table S5.2: Primer sequences for real time qRT-PCR.

Genes	Sequences (5' – 3')	Accession Numbers
*Oct4**	AACCTTCAGGAGATATGCAAA ACCTCACACGGTTCTCAATGC	NM_013633.2
Oct4	CAAGGCAAGGGAGGTAGACA GCTCCTGATCAACAGCATCA	
*Sox2**	AACTTTTGTCCGAGACCGAGA CGCGGCCGGTATTTATAATC	NM_011443.3
Sox2	AAGGGTTCTTGCTGGGTTTT AGACCACGAAAACGGTCTTG	
*Klf4**	CCAGATGCAGTCACAAGTCCC ACGACCTTCTTCCCCTCTTTG	NM_010637.3
Klf4	CCAGCAAGTCAGCTTGTGAA GGGCATGTTCAAGTTGGATT	
Esrrb	TTTCTGGAACCCATGGAGAG CAGCCAGCACCTCCTTCTAC	NM_001159500.1
Nanog	AGCCTCCAGCAGATGCAAGA GCACTTCATCCTTTGGTTTTGA	NM_028016.2
Rex1	GACACGTGGCAAAAGAAGATAGTC AGTGAGGCGATCCTGCTTTC	NM_009556.3

GAPDH	TGTGTCCGTCGTGGATCTGA GATGCCTGCTTCACCACCTT	NM_008084.2
Socs3	ATTTCGCTTCGGGACTAGC AACTTGCTGTGGGTGACCAT	NM_007707.3
Stat3	GAGGAGCTGCAGCAGAAAGT TCGTGGTAAACTGGACACCA	NM_213660.2
Brg1	CACCTACCATGCCAACACTG CGCTTGTCCTTCTTCTGGTC	NM_001174078.1
Baf155	CTGCATCTTCGATCACCTCA ACTCCCTTCCAGCACTAGCA	NM_009211.2
Baf57	GGGACCAAGTAAAGGCTTCC CAAGGTACGCAGGGGAATTA	NM_020618.4
Baf170	CACCCACCTCTCACTTCCAT GTCTGCTGGTTTCTCCTTGC	NM_001114097.1
Baf60a	CCAGCGAGAGTTCATGTTGA CCACCACATCAGTCATCGTC	NM_031842.2
Baf250a	AACAAGGCAGATGGAACACC TCCACCCACATCTTCCTCTC	NM_001080819.1
Baf47	CCTGGTGGACCAGTTTGAGT CACTGTGGGAAGTGGGTTCT	NM_011418.2
Baf200	ACGGCAGAGGTTCTCTTTCA GATTTGCCACTTGTCCTGGT	NM_175251.3
Baf60b	TGATTGGAAACCCTGAGGAG GAGGAAGGAAGGAGGCAATC	NM_001130187.1
Brm	CAACTTGGAAGGATCCCAGA CAGATTTCGCCTCTGACTCC	NM_011416.2

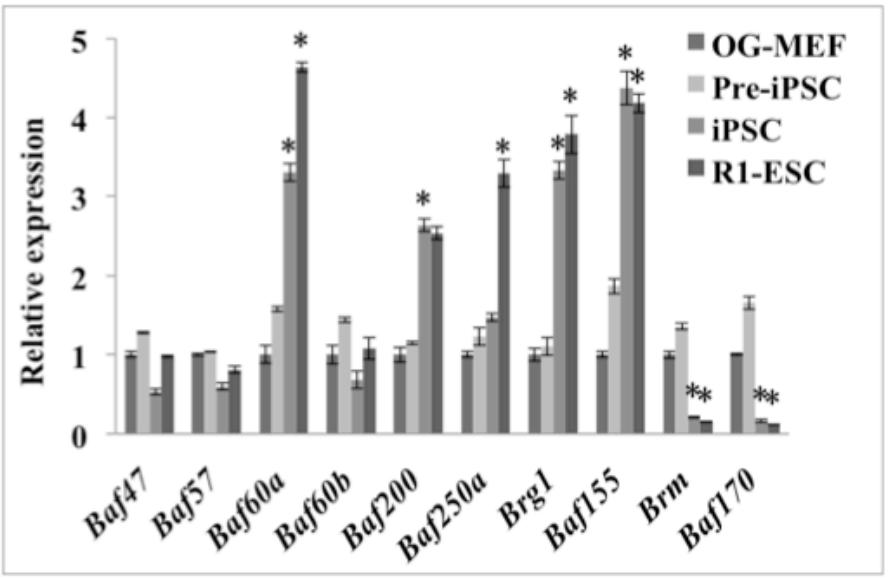
* denotes primers designed within the coding sequences of the genes and amplified both viral transgene and endogenous gene expression, primers of the same genes without the "*"were designed from 3' UTR regions of the genes and amplified only endogenous gene expression.

Figures S5.1. Schemes depicting the generation of iPSCs by OKSM and knocking down *Brm* or *Baf170* at different stages of reprogramming in OG-MEFs (A) and pre-iPSCs (B).



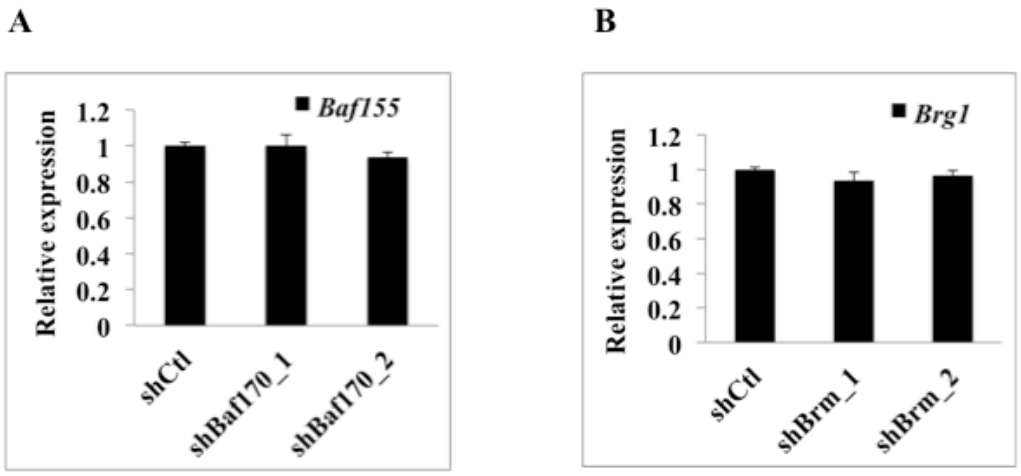
Figures S5.2. Levels (means ± SD) of subunits of *esBAF* mRNAs are high in induced pluripotent stem cells (iPSCs) and murine embryonic stem cells (R1-ESCs), but low in OG-MEFs, pre-iPSCs induced by ectopic *Oct4*/*Klf4*/*c-Myc* (OKM) infection. In contrast, levels of somatic BAF components, *Brm* and *Baf170* (*P values < 0.05; n = 3) showed the reverse patterns.

Fig. S5.2



Figures S5.3. Transfection of OG-MEFs with lentiviruses expressing different small hairpin RNAs against *Baf170* (shBaf170; A) and *Brm* (shBrm; B) does not affect expression of homologs of *Baf170* and *Brm*, or *Baf155* and *Brg1*, respectively.

Fig. S5.3



5.7. References

1 Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676, doi:10.1016/j.cell.2006.07.024 (2006).

2 Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313-317, doi:10.1038/nature05934 (2007).

3 Apostolou, E. & Hochedlinger, K. Chromatin dynamics during cellular reprogramming. *Nature* **502**, 462-471, doi:10.1038/nature12749 (2013).

4 Jaenisch, R. & Young, R. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* **132**, 567-582, doi:10.1016/j.cell.2008.01.015 (2008).

5 Ho, L. & Crabtree, G. R. Chromatin remodelling during development. *Nature* **463**, 474-484, doi:10.1038/nature08911 (2010).

6 Wang, W. *et al.* Diversity and specialization of mammalian SWI/SNF complexes. *Genes & development* **10**, 2117-2130 (1996).

7 Kadoch, C. *et al.* Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nature genetics* **45**, 592-601, doi:10.1038/ng.2628 (2013).

8 Ho, L. *et al.* An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 5181-5186, doi:10.1073/pnas.0812889106 (2009).

9 Gao, X. *et al.* ES cell pluripotency and germ-layer formation require the SWI/SNF chromatin remodeling component BAF250a. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 6656-6661, doi:10.1073/pnas.0801802105 (2008).

10 Yan, Z. *et al.* BAF250B-associated SWI/SNF chromatin-remodeling complex is required to maintain undifferentiated mouse embryonic stem cells. *Stem Cells* **26**, 1155-1165, doi:10.1634/stemcells.2007-0846 (2008).

11 Ho, L. *et al.* esBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating polycomb function. *Nature cell biology* **13**, 903-913, doi:10.1038/ncb2285 (2011).

12 Singhal, N. *et al.* Chromatin-Remodeling Components of the BAF Complex Facilitate Reprogramming. *Cell* **141**, 943-955, doi:10.1016/j.cell.2010.04.037 (2010).

13 Buganim, Y. *et al.* Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* **150**, 1209-1222, doi:10.1016/j.cell.2012.08.023 (2012).

14 Stewart, S. A. *et al.* Lentivirus-delivered stable gene silencing by RNAi in primary cells. *RNA* **9**, 493-501 (2003).

15 Tang, Y., Lin, C. J. & Tian, X. C. Functionality and transduction condition evaluation of recombinant Klf4 for improved reprogramming of iPS cells. *Cell Reprogram* **13**, 99-112 (2011).

16 Tang, Y. *et al.* Jak/Stat3 signaling promotes somatic cell reprogramming by epigenetic regulation. *Stem Cells* **30**, 2645-2656, doi:10.1002/stem.1225 (2012).

17 Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-257 (1999).
18 Wernig, M. *et al.* In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448**, 318-324, doi:10.1038/nature05944 (2007).
19 Ooi, S. K. *et al.* Dynamic instability of genomic methylation patterns in pluripotent stem cells. *Epigenetics Chromatin* **3**, 17, doi:10.1186/1756-8935-3-17 (2010).
20 Theunissen, T. W. *et al.* Nanog overcomes reprogramming barriers and induces pluripotency in minimal conditions. *Current biology : CB* **21**, 65-71, doi:10.1016/j.cub.2010.11.074 (2011).
21 Mansour, A. A. *et al.* The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature* **488**, 409-413, doi:10.1038/nature11272 (2012).
22 Rais, Y. *et al.* Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* **502**, 65-70, doi:10.1038/nature12587 (2013).
23 Gafni, O. *et al.* Derivation of novel human ground state naive pluripotent stem cells. *Nature* **504**, 282-286, doi:10.1038/nature12745 (2013).
24 dos Santos, R. L. *et al.* MBD3/NuRD facilitates induction of pluripotency in a context-dependent manner. *Cell stem cell* **15**, 102-110, doi:10.1016/j.stem.2014.04.019 (2014).
25 Bauwens, C. L. *et al.* Control of human embryonic stem cell colony and aggregate size heterogeneity influences differentiation trajectories. *Stem Cells* **26**, 2300-2310, doi:10.1634/stemcells.2008-0183 (2008).
26 Warmflash, A., Sorre, B., Etoc, F., Siggia, E. D. & Brivanlou, A. H. A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nature methods* **11**, 847-854, doi:10.1038/nmeth.3016 (2014).
27 Ho, L. *et al.* An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 5187-5191, doi:10.1073/pnas.0812888106 (2009).
28 Schaniel, C. *et al.* Smarcc1/Baf155 couples self-renewal gene repression with changes in chromatin structure in mouse embryonic stem cells. *Stem Cells* **27**, 2979-2991, doi:10.1002/stem.223 (2009).
29 Tuoc, T. C. *et al.* Chromatin regulation by BAF170 controls cerebral cortical size and thickness. *Developmental cell* **25**, 256-269, doi:10.1016/j.devcel.2013.04.005 (2013).
30 Mattout, A., Biran, A. & Meshorer, E. Global epigenetic changes during somatic cell reprogramming to iPS cells. *Journal of molecular cell biology* **3**, 341-350, doi:10.1093/jmcb/mjr028 (2011).
31 Chen, J. *et al.* H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nature genetics* **45**, 34-42, doi:10.1038/ng.2491 (2013).
32 Ichida, J. K. *et al.* A small-molecule inhibitor of tgf-Beta signaling replaces sox2 in reprogramming by inducing nanog. *Cell stem cell* **5**, 491-503, doi:10.1016/j.stem.2009.09.012 (2009).
33 Kidder, B. L., Palmer, S. & Knott, J. G. SWI/SNF-Brg1 regulates self-renewal and occupies core pluripotency-related genes in embryonic stem cells. *Stem Cells* **27**, 317-328, doi:10.1634/stemcells.2008-0710 (2009).

34 Niwa, H., Burdon, T., Chambers, I. & Smith, A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes & development* **12**, 2048-2060 (1998).
35 Yang, J. *et al.* Stat3 activation is limiting for reprogramming to ground state pluripotency. *Cell stem cell* **7**, 319-328, doi:10.1016/j.stem.2010.06.022 (2010).
36 Giraud, S., Hurlstone, A., Avril, S. & Coqueret, O. Implication of BRG1 and cdk9 in the STAT3-mediated activation of the p21waf1 gene. *Oncogene* **23**, 7391-7398, doi:10.1038/sj.onc.1207972 (2004).
37 Ni, Z. & Bremner, R. Brahma-related gene 1-dependent STAT3 recruitment at IL-6-inducible genes. *Journal of immunology* **178**, 345-351 (2007).

Chapter Six

Conclusions

This dissertation evaluated three critically important areas to further the success of animal biotechnology: 1) the biology and the underlying molecular mechanisms of embryogenesis; 2) the molecular mechanisms of HHP treatment on embryo competence; and 3) the epigenetic mechanisms of complete reprogramming. Specifically, the following conclusions are made:

1. We obtained the first complete, timeless "gold standards" for all bovine embryos including those from assisted biotechnologies such as cloning and in vitro fertilization, and conducted the first co-expression module analysis among human, mouse and bovine pre-implantation development. Moreover, we identified potential master regulators of embryo development that are not well-studied or even annotated, thus providing directions of future studies in embryogenesis.

2. Twenty-six confirmed imprinted genes in the bovine were quantified in in vivo produced oocytes and embryos and further compared them to their counterparts in humans, mice and pigs. We found differences in the imprinting status, levels and dynamics of expression among these four species.

3. We showed that 40 and 60 MPa of HHP promoted the in vitro developmental competence of bovine embryos post-cryopreservation. Furthermore, we identified specific gene expression changes induced by a positive effect of HHP on bovine IVF blastocysts.

4. Knocking down somatic BAF components, *Brm* and *Baf170,* inhibited reprogramming in a stage-specific fashion. We further showed that this effect was exerted by facilitating the activation of the "pluripotency circuitry" and by constituting to the Stat3-regulated epigenetic network during pluripotency establishment.

Our understanding of transcriptional machinery of early embryonic development and epigenetic mechanisms of reprogramming will provide insights to bypassing hurdle of bovine

iPSCs generation and to designing the optimal conditions of embryo biotechnologies. Future investigations include comprehensive characterization of transcriptome, methylome, proteome and metabolome of pre-implantation development. Our current results combined with future studies will also lead to the isolation of biomarkers for the ability of oocytes to fertilize and to develop to competent embryos.