IDENTIFICATION OF PLURIPOTENCY GENES IN THE FISH MEDAKA

WANG DANKE

(B.ENG)

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF BIOLOGICAL SCIENCE

NATIONAL UNIVERSITY OF SINGAPORE

2011

ACKNOWLEDGEMENT

My greatest gratitude goes to my supervisor, Associate Professor Hong Yunhan, for his support and supervision during the whole project. Especially thank him for his discussion and understanding of every decision I made for my future.

My thanks also go to Dr Li Zhendong, for his patience and expert guidance throughout my research work in NUS. Thanks also to Dr Guan Guijun, Miss Hong Ni, Dr Li Mingyong, Mr Yan Yan, Dr Yi Meisheng, Dr Yuan Yongming and Dr Zhao Haobin for their assistance, suggestions, and troubleshooting.

In addition, I would like to thank Mdm Deng Jiaorong for fish breeding and Miss Foong Choy Mei for laboratory management. To all my friends in the laboratory, Mr Jason Tan, Mr Kwok Chee Keong, Mr Lin Fan, Miss Manali A. Dwarakanath, Miss Narayani Bhat, and Mr Wang Tiansu, thank them for all the laughter they brought to my life in the lab. Also my thanks go to Miss Ho Danliang, for her guidance and help in my part-time teaching assistant position. Finally, I would like to thank Miss Justina Shihui Tong, who taught me a lot in the experiment in my lab rotation.

TABLE OF CONTENTS

ACKNOWLEDGEMENT I

ABSTRACT IV

LIST OF TABLES VI

LIST OF FIGURES VII

CHAPTER 1: INTRODUCTION 1

1.1 STEM CELL 1

1.1.1 Stem cell *1*

1.1.2 Pluripotent stem cell *2*

1.2 TRANSCRIPTION FACTORS 3

1.2.1 Oct4 .... *4*

1.2.2 Nanog *5*

1.2.3 Other transcription factors *7*

1.3 MEDAKA 13

1.4 MIDBLASTULA TRANSITION 14

AIM 15

CHAPTER 2: METHOD 16

2.1 ANIMAL STOCK AND MAINTENANCE 16

2.2 EXPRESSION PATTERN ANALYSIS 16

2.2.1 Collection of adult tissue *16*

2.2.2 Madaka embryo collection *17*

2.2.3 Cell culture *18*

2.2.4 Isolation of total RNA *19*

2.2.5 Sequence analysis and gene identification. *20*

2.2.6 Reverse-Transcription Polymerase Chain Reaction (RT-PCR) *21*

2.2.7 qPCR *24*

2.3 MOLECULAR CLONING 25

2.3.1 Recovery of PCR products from agarose gel *25*

2.3.2 TA cloning *25*

2.3.3 Preparation of competent cells (RbCl method) *26*

2.3.4 Transformation of competent cells *27*

2.3.5 Minipreps of plasmids (Alkaline lysis method) *27*

2.3.6 Plasmid screening by restriction enzyme digestion *28*

2.3.7 Sequencing *29*

2.3.8 Midiprep of plasmids *30*

2.4 *IN SITU* HYBRIDIZATION 30

2.4.1 Synthesis of RNA probes *30*

2.4.2 Whole mount in situ hybridization (WISH) *31*

2.4.3 Section in situ hybridization (SISH) ... *33*
2.4.4 Fluorescent in situ hybridization(FISH) ... *34*
2.4.5 Microscopy and photography ... *35*
2.5 ZYGOTIC EXPRESSION EXAMINATION ... 35
2.5.1 Gene cloning in O. celebensis ... *35*
2.5.2 Sequence analysis and species specific primer design ... *36*
2.5.3 In vitro fertilization ... *37*
2.5.4 Zygotic expression examination ... *38*

CHAPTER 3: RESULTS ... 39

3.1 GENE IDENTIFICATION ... 39
3.2 EXPRESSION PATTERN ANALYSIS ... 43
3.2.1 RT-PCR analysis of expression in tissues ... *43*
3.2.2 RT-PCR analysis of expression in embryos ... *45*
3.2.3 in vivo expression in embryos and the adult gonad ... *46*
3.2.4 Expression in ES cell culture ... *50*
3.2.5 Quantitative-PCR analysis of expression in ES cell culture ... *52*
3.3 ZYGOTIC EXPRESSION EXAMINATION ... 54
3.3.1 Gene Cloning in O. celebensis ... *54*
3.3.2 Sequence analysis ... *55*
3.3.3 RT-PCR analysis using species-specific primer ... *63*

CHAPTER 4: DISCUSSION ... 65

4.1 GENE IDENTIFICATION ... 65
4.2 IDENTIFICATION OF PLURIPOTENCY MARKERS ... 68
4.2.1 nanog and oct4 ... *68*
4.2.2 Other genes ... *70*
4.2.3 Somatic expression of pluripotency genes ... *71*
4.2.4 Conserved expression of pluripotency genes in vertebrates ... *73*
4.3 ZYGOTIC EXPRESSION PATTERN ... 74
4.3.1 Zygotic expression at midblastula ... *74*
4.3.2 O. celebensis in building the zygotic expression examining model ... *76*
4.3.3 Hierarchical expression among pluripotency genes ... *77*

CHAPTER 5: CONCLUSION AND FUTURE WORK ... 80

5.1 CONCLUSION ... 80
5.2 FUTURE WORK ... 81

CHAPTER 6: REFERENCES ... 82

ABSTRACT

Stem cell cultures can be derived directly from early developing embryos and indirectly from differentiated cells by forced expression of pluripotency transcription factors. Pluripotency genes are routinely used to characterize mammalian stem cell cultures at the molecular level. However, such genes have remained unknown in lower vertebrates. This study made use of the laboratory fish medaka as a model, because of its unique embryonic stem (ES) cells and sequenced genome as useful experimental tools and genetic resource. Seven medaka pluripotency genes were identified by homology search and expression in vivo and in vitro. By RT-PCR analysis, they fall into three groups of expression pattern. Group I includes *nanog* and *oct4* showing gonad-specific expression; Group II contains *sall4* and *zfp281* displaying gonad-preferential expression; Group III has *klf4*, *ronin* and *tcf3* exhibiting expression also in several somatic tissues apart from the gonads. The transcripts of the seven genes are maternally supplied and persist at a high level during early embryogenesis. Early embryos and adult gonads were used to examine expression in stem cells and differentiated derivatives by in situ hybridization. Strikingly, *nanog* and *oct4* are highly expressed in pluripotent blastomeres of 16-cell embryos. In the adult testis, *nanog* expression was

specific to spermatogonia, the germ stem cells, whereas *tcf3* expression occurred in spermatogonia and differentiated cells. Most importantly, all the seven genes are pluripotency markers in vitro, because they showed high expression in undifferentiated ES cells but dramatic down-regulation upon differentiation. Therefore, these genes have conserved their pluripotency-specific expression in vitro from mammals to lower vertebrates. Furthermore, by using sequence differences between two medaka species: *O. latipes* and *O. celebensis*, I built a model to examine the timing of zygotic expression of six pluripotency genes by determining their expression from the paternal alleles. All those genes showed the onset of zygotic expression around the midblastula transition, suggesting their critical roles in early embryogenesis. Specifically, *nanog* and *oct4* show earlier expression than the other remainder. Data obtained suggest the feasibility to study the hierarchical expression patterns of genes involved in pluripotency, cell fate decision and other processes.

LIST OF TABLES

Table 1	Genes and Primers used in RT-PCR for expression pattern analysis
Table 2	Genes and Primers used in qPCR
Table 3	Genes and Primers used for gene cloning
Table 4	Genes and Primers used in RT-PCR for examining zygotic expression
Table 5	Summary of RNA expression in medaka adult tissues
Table 6	Summary of zygotic expression in medaka embryos
Table 7	The hierarchy of zygotic activation

LIST OF FIGURES

Figure 1 Sequence comparison of E2A (E12/E47) proteins on alignment

Figure 2 Phylogenetic comparison of Tcf3/Tcf7l1 proteins

Figure 3 RT-PCR analysis of RNA expression in adult tissues

Figure 4 RT-PCR analysis of RNA expression in embryos

Figure 5 RNA expression by in situ hybridization

Figure 6 RNA expression in ES cell culture (RT-PCR)

Figure 7 RNA expression in ES cell culture (qPCR)

Figure 8 RT-PCR results of seven genes in *O. latipes* and *O. celebensis*

Figure 9 Alignment of *nanog* cDNAs between *O. celebensis* and *O. latipes*

Figure 10 Alignment of *oct4* cDNAs between *O. celebensis* and *O. latipes*

Figure 11 Alignment of *sall4* cDNAs between *O. celebensis* and *O. latipes*

Figure 12 Alignment of *klf4* cDNAs between *O. celebensis* and *O. latipes*

Figure 13 Alignment of *ronin4* cDNAs between *O. celebensis* and *O. latipes*

Figure 14 Alignment of *tcf3* cDNAs between *O. celebensis* and *O. latipes*

Figure 15 Alignment of *zfp281* cDNAs between *O. celebensis* and *O. latipes*

Figure 16 RT-PCR analysis of zygotic RNA expression in embryos (*O. celebensis* male X *O. latipes* female)

Figure 17 RT-PCR analysis of zygotic RNA expression in embryos (*O. latipes* male X *O. celebensis* female)

CHAPTER 1: INTRODUCTION

1.1 Stem cell

1.1.1 Stem cell

Stem cells, found in all multicellular organisms, can divide through mitosis and differentiate into diverse specialized cell types and can self renew to produce more stem cells. Stem cells are defined based on three characteristics: they can undergo self-renewing cell divisions in which at least one of the daughter cells is a stem cell; they can give rise to multiple types of cell; they must be capable of repopulating a tissue in vivo (Verfaillie, 2009).

Potency specifies the differentiation potential of the stem cell. Generally, there are five kinds of potency in stem cell. Totipotency is the ability of a single cell to divide and produce all the differentiated cells in an organism. Pluripotency refers to the potential of the stem cells to differentiate into any of the three germ layers: endoderm, mesoderm or ectoderm. However, pluripotent stem cells lack the potential to contribute to extraembryonic tissue. Multipotent stem cells can differentiate into a number of cells, but only those of a closely related family of cells. In addition, there are also oligopotent and unipotent cells.

All mammals are derived from a single stem cell, which is known as the totipotent stem cell, such as zygote. During embryogenesis, stem cells undergo specification to give rise to extraembryonic endoderm or cells of the

inner cell mass (ICM), which are pluripotent. Cells in the ICM then differentiate to three germ layers cells and finally to tissue specific stem cells. With each additional step of specification, cells become more and more limited in their differentiation ability. These stem cells present in tissues of the three germ layers cannot generate cells of other tissues within the same germ layer or of the other germ layers, but the cells of that tissue. These cells are therefore multipotent (Verfaillie, 2009).

1.1.2 Pluripotent stem cell

In mammals, there are two broad types of stem cells: embryonic stem (ES) cells that are isolated from the ICM of blastocysts, and adult stem cells that are found in various tissues. ES cells have the ability to grow indefinitely while maintaining pluripotency and the ability to differentiate into cells of all three germ layers (Evans and Kaufman, 1981).

Pluripotent stem cell cultures can be derived from developing embryos and adult tissues. In particular, ES cells from early developing embryos are an excellent model for analyzing vertebrate development *in vitro*, and provide a versatile resource for induced differentiation into many desired cell types in large quantity for regenerative medicine. In mouse, ES cells have widely been used in gene targeting for generating knockout animals to study the functions of genes and to model human diseases (Wobus and Boheler, 2005). Since the production of first mouse ES cell lines in 1981 (Evans and Kaufman, 1981), numerous attempts have been made in ES cell derivation in different

vertebrate organisms (Munoz et al., 2009; Yi et al., 2010b). The success in the production of ES cells from the fish medaka (Hong et al., 1996) demonstrated that the ability for ES cell derivation is present from fish to mammals. In 1998, human ES cells were successfully established (Thomson et al., 1998). Previously, somatic cells can be reprogrammed by transferring their nuclear into oocytes (Wilmut et al., 1997) or by fusing the somatic cells with ES cells (Cowan et al., 2005; Tada et al., 2001). At present, induced pluripotent stem (iPS) cells can be generated from differentiated cells by introduced expression of reprogramming transcription factors (Takahashi et al., 2007; Takahashi and Yamanaka, 2006; Yu et al., 2007).

Although interest in ES cell derivation has steadily been increasing and putative ES-like cells have been reported in several species (Hong N, 2011; Yi et al., 2010b), including short-term ES cell cultures capable of germline transmission in zebrafish(Ma et al., 2001a), stable lines of real ES cells have been limited to few organisms. One of the challenges is that there are no suitable molecular markers to monitor and regulate the pluripotency of putative ES cell cultures.

1.2 Transcription factors

Recent studies have established that mouse and human ES cells share a core transcriptional network consisting of Oct4, Nanog and Sox2 (Boyer et al., 2005b; Loh et al., 2006b). Oct4 and Nanog, are essential regulators of early development and ES cell identity (Chambers et al., 2003; Nichols et al., 1998).

1.2.1 Oct4

In 2006, Shinya Yamanaka demonstrated induction of pluripotent stem cells from mouse embryonic and adult fibroblasts by introducing four factors, Oct4, Sox2, c-Myc and Klf4, under ES cell culture conditions (Takahashi and Yamanaka, 2006). The second year, his group reported the induction of pluripotent stem cells from adult human fibroblasts using the same factors (Takahashi et al., 2007). The same year, James A. Thomson demonstrated that transcription factor Oct4, Sox2, Nanog and Lin28 are sufficient to reprogram human somatic cells to pluripotent stem cells that exhibit the essential characteristics of embryonic stem (ES) cells (Yu et al., 2007). In all those three studies, Oct4 is the key factor in induced pluripotent stem (iPS) cells.

The mouse *oct4* gene (also known as *pou5f1*) encodes a POU domain-containing and an octamer-binding protein and represents the prototype of pluripotency genes in mammals, because it is maternally supplied and expressed specifically throughout the totipotent cycle, including the inner cell mass (ICM), epiblast and primordial germ cells (PGCs) of early developing embryo and spermatogonia and oocytes (Pesce et al., 1998b). These embryonic and adult cells are all capable of producing stem cell cultures, in which quantitative *Oct4* expression defines differentiation, dedifferentiation or self-renewal(Niwa et al., 2000b). Oct4-null embryos develop abnormal ICM cells which are not pluripotent and have a greater propensity to express

trophoblast markers and subsequently die at the peri-implantation stage of development (Nichols et al., 1998). When Oct4 expression is repressed in ES cells, cells lose their self-renewing state and spontaneously differentiate to the trophectodermal lineage (Nichols et al., 1998; Niwa et al., 2000a).

Oct4 can act either to repress or to activate target gene transcription (Pesce and Schöler, 2001). It regulates expression of multiple genes (Saijoh et al., 1996) via interactions with transcription factors present in pluripotent cells. For example, Oct4 can heterodimerize with Sox2, to affect the expression of several genes in mouse ES cells (Botquin et al., 1998).

It was also reported that ES cells maintain Oct4 at an appropriate level in order to remain pluripotency. Either increase or decrease in expression of *oct4* may induce ES cells to differentiate (Niwa et al., 2000a). ES cells with *oct4* downregulation differentiate to the trophectodermal lineage, while ES cells with an overexpression of *oct4* tend to differentiate to multiple cell types (Niwa et al., 2000a).

1.2.2 Nanog

Nanog is one of the transcription factors used in human somatic cells reprogramming (Yu et al., 2007).

The gene *nanog* encodes a divergent protein that contains a homeobox. Its expression commences at the morula stage and later on occurs in the ICM, epiblast, PGCs and ES cell cultures, albeit its expression is absent in spermatogonia of the adult testis (Cavaleri and Scholer, 2003; Chambers et al.,

2003). Nanog is one of the key regulators essential for the formation and maintenance of the ICM during mouse pre-implantation development and for self-renewal of pluripotent ES cells (Loh et al., 2006a)

Embryos devoid of Nanog were unable to form epiblasts. (Mitsui et al., 2003) Such embryos seem to be able to initially give rise to the pluripotent cells, but these cells then immediately differentiate into the extraembryonic endoderm lineage (Loh et al., 2006a).

Similar to *oct4*, *nanog* is downregulated upon ES cells differentiation and ES cells deficient in Nanog differentiate into cells of the extraembryonic endoderm lineage (Mitsui et al., 2003). Hence, it is commonly stated that both Nanog and Oct4 are critical in maintaining pluripotency in ES cells.

However, contrary to previous findings, in other studies, Nanog-deficient ES cells were able to self-renew with Oct4 and Sox2 maintained (Chambers et al., 2007). Therefore, such result gives rise to a conclusion that Nanog is essential for establishing pluripotency but is dispensable for the maintenance of self-renewal and pluripotency in ES cells (Heng and Ng, 2010). In the further study of the key transcription factor, namely Oct4, Nanog and Sox2, to maintain pluripotency in ES cells, researcher concluded that although each of these proteins has been described as a “master regulator” of pluripotency, only Oct4 appears absolutely essential, while both Sox2 and Nanog appear dispensable, at least in certain molecular contexts (Masui et al., 2007).

1.2.3 Other transcription factors

Additional pluripotency genes have recently been described in mammals. These include *klf4* encoding a Krüppel-like factor(Takahashi and Yamanaka, 2006), *ronin* encoding the THAP domain containing 11 protein (Thap11) (Dejosez et al., 2008b), *sall4* encoding the Sal-like protein 4 (Sall4) (Lim et al., 2008; Tsubooka et al., 2009a), *tcf3 (*also called *tcf7l1*) encoding the T-cell transcription factor 3 (Cole et al., 2008a; Tam et al., 2008a) and *zfp281* encoding the zinc-finger protein 281(Wang et al., 2008a). These genes have been identified in mammalian ES cells. Their homologs/orthologs remain to be identified and characterized in lower vertebrates.

1.2.3.1 Klf4

Klf4, which is a zinc finger transcription factors, has a defining role in maintaining self-renewal in ES cells (Jiang et al., 2008). Overexpression of *klf4* can promote the self-renewal of ES cells (Hall et al., 2009). It was also shown that *klf4* mRNA and protein expression were down-regulated during human ES cells differentiation (Chan et al., 2009).

As a key factor in reprogramming, Klf4 functions as both a transcriptional activator and a repressor to regulate proliferation and differentiation of different cell types (Evans et al., 2007). Klf4 is the activator of the target gene *nanog*. Overexpression of *klf4* up-regulates *nanog* promoter activity and the endogenous *nanog* protein expression in human ES cells. Knockdown of *klf4* or mutation of Klf4 binding motifs significantly down-regulated *nanog* promoter

activity (Chan et al., 2009).

Genome-wide chromatin immunoprecipitation with microarray analysis demonstrates that the DNA binding profile of Klf4 overlaps with that of Oct4 and Sox2 on promoters of genes specifically underlying establishment of iPS cells, suggesting transcriptional synergy among these factors (Sridharan et al., 2009). Such finding was confirmed by other findings. A dominant negative mutant of Klf4 can compete with wild-type Klf4 to form defective Oct4/Sox2/Klf4 complexes and strongly inhibit reprogramming (Wei et al., 2009). This finding reinforces the idea that direct interactions between Klf4, Oct4, and Sox2 are critical for somatic cell reprogramming (Wei et al., 2009). Another example was also provided: Klf4 was identified as a mediating factor that cooperates with Oct4 and Sox2 on the distal enhancer in activating the *lefty1* promoter in ES cells (Nakatake et al., 2006).

A recent study showed that introduction of single reprogramming factor, Klf4, can induce pluripotent stem cells from epistem cell (EpiSCs), a cell line that is from post-implantation epithelialized epiblast and unable to colonize the embryo even though they express the core pluripotency genes, *oct4, sox2* and *nanog*. These EpiSC-derived induced pluripotent stem (Epi-iPS) cells activated expression of ES cell-specific transcripts including endogenous Klf4, and down-regulated markers of lineage specification (Guo et al., 2009).

1.2.3.2 Ronin

The protein Ronin possesses a Thap domain, which is associated with

sequence-specific DNA binding and epigenetic silencing of gene expression (Dejosez et al., 2008a). It is expressed primarily during the earliest stages of mouse embryonic development, and its deficiency in mouse produces peri-implantational lethality and defects in the ICM (Dejosez et al., 2008a). Meanwhile, *ronin* knockout ES cells were found to be nonviable. On the other hand, the overexpression of *ronin* can inhibit ES cells differentiation and allows ES cells to self-renew under conditions that normally suppress self-renewal (Heng and Ng, 2010). Interestingly, it was found that ectopic expression of *ronin* was able to partly compensate for *oct4* knock-down in ES cells (Dejosez et al., 2008a). The authors also demonstrated that Ronin is a transcriptional repressor of multiple genes that are either directly or indirectly involved in differentiation (Dejosez et al., 2008a). Furthermore, it was shown that Ronin exerts its anti-differentiation effects through epigenetic silencing of gene expression (Dejosez et al., 2008a). In addition, *ronin*, like *nanog*, is also targeted by Caspase-3, a component of the cell death system that compels ES cells to exit their self-renewal phase and induces differentiation (Fujita et al., 2008). Those findings reinforce that ronin is essential for the maintenance of pluripotent stem cells.

1.2.3.3 Sall4

A spalt family member, *sall4*, is required for the pluripotency of ES cells (Zhang et al., 2006). Previously, microarray studies revealed that *sall4* and *oct4* share similar expression patterns during early embryo development (Hamatani et al.,

2004). Similarly to Oct4, a reduction in Sall4 levels in mouse ES cells results in respecification, under the appropriate culture conditions, of ES cells to the trophoblast lineage (Zhang et al., 2006). Also similar to *oct4*, in mouse ES cells, both overexpression and underexpression of Sall4 cause differentiation suggesting that ES cells maintain Sall4 at an appropriate level in order to remain pluripotency (Yang et al., 2010).

Sall4 is important for early embryonic cell-fate decisions (Zhang et al., 2006). Sall4-null embryos died shortly after implantation (Tsubooka et al., 2009b).

Sall4 plays positive roles in the generation of pluripotent stem cells from blastocysts and fibroblasts. Although ES-like cell lines could be established from Sall4-null blastocysts, the efficiency was much lower. The knockdown of *sall4* significantly decreased the efficiency of iPS cell generation from mouse fibroblasts. Furthermore, retroviral transduction of *sall4* significantly increased the efficiency of iPS cell generation in mouse and some human fibroblast lines (Tsubooka et al., 2009b).

By using chromatin immunoprecipitation coupled to microarray hybridization, researchers have identified a total of 3,223 genes that were bound by the Sall4 protein on duplicate assays with high confidence (Yang et al., 2008). Many of these genes have major functions in developmental and regulatory pathways (Yang et al., 2008). For example, Sall4 is a transcriptional activator of *oct4*, and has a critical role in the maintenance of ES cell pluripotency by modulating *oct4* expression. Microinjection of *sall4* small interfering (si) RNA into mouse

zygotes resulted in reduction of *sall4* and *oct4* mRNAs in pre-implantation embryos (Zhang et al., 2006).

Although most studies generally showed the important role Sall4 plays in ES cells, Shunsuke Yuri stated that Sall4 does not contribute to the central machinery of the pluripotency. Instead it stabilizes ES cells by repressing aberrant trophectoderm gene expression (Yuri et al., 2009). Such statement provided a new point of view of the unique function of Sall4.

1.2.3.4 Tcf3

The Wnt signaling pathway is necessary both for maintaining undifferentiated stem cells and for directing their differentiation. In mouse ES cells, Wnt signaling preferentially maintains “stemness” under certain permissive conditions (Tam et al., 2008b)

In searching how external signals connect to this regulatory circuitry to influence ES cell fate, Cole found that a terminal component of the Wnt pathway in ES cells, T-cell transcription factor-3 (Tcf3), co-occupies promoters throughout the genome in association with the pluripotency regulators Oct4 and Nanog (Cole et al., 2008b). In mouse pre-implantation development embryos, Tcf3 expression is co-regulated with Oct4 and Nanog and becomes localized to the ICM of the blastocyst (Tam et al., 2008b).

Up-regulation of *nanog* upon *tcf3* depletion showed that Tcf3 acts to repress *nanog* under normal ES cell growth conditions (Pereira et al., 2006). Furthermore, it was also found that *tcf3* depletion caused increased

expression of Oct4, Nanog, and other pluripotency factors and produced ES cells (Cole et al., 2008b). Comparing effects of *tcf3* ablation with *oct4* or *nanog* knockdown revealed that Tcf3 counteracted effects of both Nanog and Oct4 (Yi et al., 2008). However, Tcf3 is still important in maintaining pluripotency in ES cells. Through repressing pluripotency factors, Tcf3 prevents overactivation of transcriptional circuits, promoting pluripotent cell self-renewal (Yi et al., 2008). Meanwhile, by using a whole-genome approach, researchers found that Tcf3 transcriptionally repressed many genes important for maintaining pluripotency and self-renewal, as well as those involved in lineage commitment and stem cell differentiation (Tam et al., 2008b). Thus, it was concluded that Wnt pathway, through Tcf3, brings developmental signals directly to the core regulatory circuitry of ES cells to influence the balance between pluripotency and differentiation (Cole et al., 2008b).

1.2.3.5 Zfp281

The zinc finger transcription factor Zfp281 was first implicated as a regulator of pluripotency in ES cells since it is expressed in undifferentiated ES cells and less in differentiated ES cells (Brandenberger et al., 2004). Then, Zfp281 was identified as a key component of the pluripotency regulatory network in a series of studies.

Zfp281 was shown to activate *nanog* expression directly by binding to a site in the promoter (Wang et al., 2008b). Its binding sites for *oct4*, *sox2* were also identified (Chen et al., 2008). Data in the analysis of protein interaction

networks in ES cells showed that Zfp281 physically interacts with Oct4, Sox2, and Nanog in regulating pluripotency (Wang et al., 2006).

Through Chromatin immunoprecipitation, 2417 genes were identified to be direct targets by Zfp281, including several transcription factors that are known regulators of pluripotency (Wang et al., 2008b). Upon knockdown of *zfp281*, some of the Zfp281 target genes were activated, whereas others were repressed, suggesting that this transcription factor plays bifunctional roles in regulating gene expression within the network (Wang et al., 2008b).

1.3 Medaka

The medaka (*Oryzias latipes*) is well-suited for analyzing vertebrate development (Wittbrodt et al., 2002). This laboratory fish is used as a lower vertebrate model for stem cell biology (Hong N, 2011; Yi et al., 2010b). In this organism, a feeder-free culture system has been previously established that allowed for derivation of diploid ES cells(Hong et al., 1996; Hong et al., 1998) from midblastula embryos as the equivalent of the mammalian blastocysts, male germ stem cells from the adult testis (Hong et al., 2004a) and even haploid ES cells from gynogenetic embryos (Yi et al., 2009; Yi et al., 2010a). Most recently, it was demonstrated that in medaka early embryos even up to the 32-cell stage are capable of cell culture initiation (Li et al., 2011). In addition, the medaka genome has been fully sequenced and partially annotated (http://www.ensembl.org/index.html). The availability of sequence data and well-characterized stem cell lines makes medaka a unique model

organism to identify pluripotency genes *in vivo* and *in vitro*.

1.4 Midblastula transition

Early development of the embryo is directed by maternal gene products and has limited zygotic gene activity (O'Boyle et al., 2007). The cell divisions are characterized by synchrony, short phases and no cell motility (O'Boyle et al., 2007). At the midblastula transition (MBT), a series of event happens: zygotic gene transcription is activated; the cell cycle lengthens; cell divisions lose their synchrony; cell motility begins (Newport and Kirschner, 1982a; Newport and Kirschner, 1982b). In the Xenopus embryos, the developmental changes termed MBT begin after the 12th cell division (Newport and Kirschner, 1982a). In zebrafish, it begins at the 512-cell stage (cell cycle10), two cycles earlier than in Xenopus (Bree et al., 2005). Realization of critical nucleocytoplasmic ratio is thought to trigger the beginning of MBT (Aizawa et al., 2003). However, the mechanism of MBT has not been clarified.

Many of the genes that are first expressed at this stage will play critical roles in later events such as gastrulation and segmentation (Bree et al., 2005). So in order to obtain more insight of the function of certain gene, it is important to study the activation of zygotic transcription.

There have been several studies to identify zygotically expressed genes in mouse. Those methods include large scale sequencing of expressed sequence tags from staged pre-implantation cDNA libraries (Ko et al., 2000), mRNA differential display (Ma et al., 2001b) and suppression subtractive

hybridization to prepare subtract zygotic cDNA libraries (Yao et al., 2003). However, in mouse embryo, the activation of zygotic gene expression occurs by the 2 cell stage (Aizawa et al., 2003). In addition, transcription of some genes even begins at 1 cell stage, such as *hsp70.1* (Aizawa et al., 2003). Because of the fast zygotic activation, it is not possible to examine more details about the activation of genes at MBT.

Aim

Here I planned to identify several medaka pluripotency genes and examine their candidacy as pluripotency markers by analyzing RNA expression patterns in adult tissues, developing embryos and ES cell culture. Furthermore, I intended to build a model to examine the timing of zygotic expression of those identified genes, and explore some features about the paternal expression pattern.

CHAPTER 2: METHOD

2.1 Animal stock and maintenance

Work with fish followed the guidelines on the Care and Use of Animals for Scientific Purposes of the National Advisory Committee for Laboratory Animal Research in Singapore (permit number 27/09). Medaka was maintained under an artificial photoperiod of 14-h/10-h light/darkness at 26°C (Li et al., 2009). Adult fish are fed two to three times with artemia nauplii and dry food. Medaka strains HB32C and i1 were used for gene expression analysis by RT-PCR and in situ hybridization. HB32C is a wild-type pigmentation strain from which diploid ES cell lines MES1 were derived (Hong et al., 1996), whereas i1 is an albino strain from which haploid ES cell lines HX1 were generated (Yi et al., 2009; Yi et al., 2010a).

2.2 Expression pattern analysis

2.2.1 Collection of adult tissue

Adult male and female medaka fish were anaesthetized in ice-cold water for 2 min. Dissection was then performed under a stereomicroscope (Leica MZ125). Tissues and organs were excised and collected into eppendorf tubes. Several adult tissues were selected in this experiment: brain, skin, heart, kidney, liver, gut, testis and ovary.

2.2.2 Madaka embryo collection

To obtain egg production in the laboratory, young medaka fish were kept after hatching under resting condition at high density in large containers for up to 8 weeks. When females started spawning with about 5-10 eggs per day, males and females were separated for several days before they were brought together for the production of large numbers of eggs.

When experiment started, the separated males and females were mixed together. Spawning took place within the first 30 min. The eggs stuck together through hairy filament and attached to the belly of female fish for several hours. Fertilized eggs were obtained by carefully scraping the egg clusters from females by hand. The clusters of eggs were transferred into a petri-dish with embryo rearing medium (ERM:17 mM NaCl, 0.4 mM KCl, 0.3 mM $CaCl_2 \cdot H_2O$, 0.6mM $MgSO_4 \cdot 7H_2O$, 1 ppm methylene blue). Single eggs were obtained either by rolling egg clusters on the petri-dish or by using needle to remove hairy filaments. Once the eggs were separated, the dead and injured embryos were removed.

Single embryos were kept in 28°C until the desired stage of embryonic development. Embryos were staged according to Iwamatsu (Iwamatsu, 2004). Several developmental stages were chosen: 16-cell, morula, early blastula, late blastula, pre-mid gastrula, 34 somite and prehatch.

2.2.3 Cell culture

Maintenance and induced differentiation by embryoid body formation in suspension culture were done essentially as described for the diploid ES cell line MES1 (Hong et al., 1996) and the haploid ES cell line HX1 (Yi et al., 2009; Yi et al., 2010a). ES cells were cultured in medium ESM4 in gelatin-coated tissue culture plastic ware (BD biosciences, NJ) for undifferentiated growth. The exponentially growing cells were washed with 2ml PBS. After PBS was removed, 1 ml 1X trypsin-EDTA was added to trypsinize cells at room temperature for 5 min. The cells were spun down, and trypsin-EDTA was aspirated. Cells were resuspended in 1 ml ESM4 to form a single-cell suspension. Single cells were transferred into three to six wells of a gelatin-coated six-well plate or a 10-cm tissue culture dish.

For induced differentiation, ES cells were trypsinized and separated into single cells and small aggregates, seeded into non-adherent bacteriological Petri dishes for suspension culture in growth factor-depleted ESM4 containing all-trans retinoic acid (Sigma, final 5 μM) for 7 days before harvest for RNA isolation. In suspension culture, cells formed aggregates of varying sizes (mostly 100~20 μm in diameter) at day 1 and subsequently developed into spherical embryoid bodies. The dishes of suspension cultures were gently swirled twice a day to prevent any attachment.

2.2.4 Isolation of total RNA

After adult tissue collection, samples were collected in eppendorf tubes respectively. For each developmental stage, approximately 20-30 medaka embryos were collected in eppendorf tubes. Excess water was removed with pipette. The tissues and embryos were homogenized in 1 ml of TRIZOL RNA isolation reagent (Invitrogen) with plastic pestles. 200 ml of chloroform was added in each tube. Each tube was inverted several times until the liquid was homogenized. The aqueous phase and organic phase were separated by centrifugation at 12000X g for 20 min at 4°C. The upper aqueous phase was transferred to a fresh RNase free tube. RNA was precipitated by 1 volume of isopropanol at room temperature for 30 min and centrifuged at 12000X g for 20 min at 4°C. After centrifugation, supernatant was removed. The RNA pellet was washed with diethylpyrocarbonate (DEPC, Sigma) treated 70% ethanol (ethanol was dissolved by 0.1% DEPC treated water) and centrifuged at 12000X g for 20 min at 4°C. Supernatant was removed as much as possible. The RNA pellet was air-dried and dissolved in 20 μl of 0.1% DEPC treated water. The quality of RNA was ascertained by gel electrophoresis, and the concentration was determined by Nanophotometer (WPA BioWave II+). The RNA samples were stored at -80°C if not immediately used for RT-PCR.

2.2.5 Sequence analysis and gene identification.

BLAST searches were run against public databases by using BLASTN for nucleotide sequences, BLASTP for protein sequences and tBLASTN from protein queries to nucleotide sequences. Multiple sequence alignment was conducted by using the Vector NTI suite 11 (Invitrogen). Phylogenetic trees were constructed by the DNAMAN package (Lynnon BioSoft). Chromosomal locations were investigated by examining corresponding genomic sequences.

Several medaka genes homologous/orthologous to the mammalian pluripotency genes have been previously described (Yi et al., 2009), including *oct4*, *nanog*, *klf4*, *ronin, sall4*, *zfp281a* and *tcf3*. However, the previous medaka *tcf3* according to the genome annotation (ENSORLG00000004923 and ENSORLG00000015259) was found to encode transcription factor E2A instead (Figure 1). A BLAST search by using the human *tcf3* as a query against the medaka genome sequence (http://www.ensembl.org/index.html) led to the identification of a gene annotated as *tcf3* (ENSORLG00000011813), to which more than 30 expressed sequence tags (http://blast.ncbi.nlm.nih.gov/) displayed ≥96% identity in nucleotide sequence.

2.2.6 Reverse-Transcription Polymerase Chain Reaction (RT-PCR)

2.2.6.1 DNase treatment

In order to reduce the genomic DNA contamination, DNase I (Invitrogen) was used to treat the total RNA before cDNA synthesis.

The following were added to an RNase-free, eppendorf tube on ice: 1 μg of RNA sample, 1 μl of 10X DNase I Reaction Buffer, 1 μl of DNase I (Amp Grade, 1 U/μl) and DEPC-treated water to 10 μl. The tube was incubated at room temperature for 15 min. Then the DNase I was inactivated by 1 μl of 25 mM EDTA solution to the reaction mixture, followed by 10 min heating at 65°C. Now the total RNA was ready for cDNA synthesis.

2.2.6.2 Reverse Transcription (RT) for cDNA synthesis

Synthesis of cDNA templates was primed with oligo $(dT)_{25}$ by using M-MLV reverse transcriptase (Invitrogen). A 20 μl reaction volume was used for 2 μg of total RNA. 1 μl of oligo $(dT)_{25}$ (500 μg/ml), 1 μl of 10 mM dNTP Mix (10 mM each dATP, dGTP, dCTP and dTTP) and 2 μg of total RNA were added into a nuclease-free eppendorf tube. Sterile, distilled water was added to adjust the final volume to 12 μl. The mixture was heated to 65°C for 5 min and quickly chilled on ice thereafter. After brief centrifugation, 4 μl of 5X First-Strand Buffer, 2 μl of 0.1 M dithiothreitol (DTT) and 1 μl of RNaseOUT™ Recombinant Ribonuclease Inhibitor (40 units/μl) were added into the reaction. The mixture was incubated for 2 min at 37 °C. 1 μl (200 units) of M-MLV reverse transcriptase was added at last, followed by 50 min incubation at 37°C. In

order to inactivate the reaction, the final product was heated at 70°C for 15 min. The final cDNA product was stored in -20°C for further use.

2.2.6.3 Polymerase Chain Reaction (PCR)

Standard PCR was performed in a 25 μl reaction using PTC-100 Thermal Cyclers (Bio-Rad). The following reaction components were added to a PCR tube for a final reaction volume of 25 μl: 2.5 μl of 10X Ex Taq Buffer(Mg2+plus), 2 μl of dNTP Mix (2.5 mM each), 1 μl of amplification primer 1 (10 μM), 1 μl of amplification primer 2 (10 μM), 0.2 μl of Taq DNA polymerase (5 U/μl), 1 μl of cDNA (from first-strand reaction) and 17.3 μl of autoclaved, distilled water.

The parameters for the amplification reaction were as below:

94 ^{o}C for 5 min (first denaturation);

35 cycles of amplification process:

94 ^{o}C for 10 sec (denaturation),

58 ^{o}C for 20 sec (annealing),

72 ^{o}C for 60 sec (extension);

72 ^{o}C for 7 min (final extension).

PCR was run for 28 and 35 cycles for β-actin and other genes, respectively. Primers and gene accession numbers are listed in Table 1.

Table 1 Genes and Primers used in RT-PCR for expression pattern analysis

Gene		Primer sequence		size (bp)
Name	Accession	Forward primer	Reverse primer	
nanog	FJ436046	CTCCACATGTCCCCCCTTATC	AGGATAGAATAGTCACATCAC	**591**
oct4	NM_001104869	GCTTTCTTTGGCGTAAACTCGTC	TCATCCTGTCAGGTGACCTACC	**777**
klf4	ENSORLT00000007097	CATCCTCTCACCCAGATGC	TCATAAGTGCCTCTTCATGTGG	**447**
sall4	ENSORLG00000016130	ATGTCGAGGCGCAAACAAG	AGCCACTTTAGCGTCAGGTATG	**501**
tcf3	HQ705658	ATGCCTCAACTGAACGGAGG	CTGCAGAGCTGGGAACATCC	**433**
zfp281	ENSORLG00000005292	ATGAGTATTATCCAAGACAAGATAGGC	TGTGTCCTTTTGTGTCGCTCC	**854**
ronin	ENSORLG00000008903	AACTGAGAAGCGACGAGTACTC	CATTTTCTTTCTGAAACCAAC	**302**

2.2.6.4 Agarose gel electrophoresis

The PCR products were separated on 1% agarose gels. 1% agarose was prepared by melting 5 g of agarose (1st BASE) in 500 ml 1X TAE electrophoresis running buffer (50X stock: 2M Tris-acetic acid, 10 mM EDTA, pH 8.0). 1% agarose was mixed with ethidium bromide. Agarose gel was prepared in tray mould using proper comb. The PCR products (25 μl) were thoroughly mixed with 5 μl of 6X loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol, and 60% glycerol), and a 10 μl aliquot was loaded into each well in the agarose gel. Electrophoresis was run at 100 V, 400 mA for 30 min via PowerPac Basic Power Supply(Bio-Rad). The electrophoresis results were documented with a bioimaging system (Synoptics).

2.2.7 qPCR

Quantitative real time PCR analysis was carried out in triplicate on the IQ5 system (BioRad). Each reaction consisted of 25 μl of 1X qPCR SuperMix (SYBR GreenER), 1 μl of forward primer (10 μM), 1 μl of reverse primer (10 μM), 10 μl of template and 13 μl of DEPC-treated water. Primers are listed in Table 2. A standard 50-μl reaction size is provided; component volumes can be scaled as desired.

The program of the real-time instrument was as below:

50°C for 2 min hold (UDG incubation);

95°C for 10 min hold (UDG inactivation and DNA polymerase activation);

40 cycles of:

95°C, 15 sec ,

60°C, 60 sec ,

Melting curve analysis.

The reaction PCR plate was centrifuged briefly to make sure that all components were at the bottom of the plate. The PCR plate was sealed before it was placed in the preheated real-time instrument programmed as described above. When the reaction stopped, the data were collected and analyzed.

Table 2 Genes and Primers used in qPCR

Gene	Primer sequence		size
Name	Forward primer	Reverse primer	(bp)
nanog	CACAATGCTGTCCAATGAGGTAG	GCCAGGATAGAATAGTCACATCAC	232
oct4	TGTCTCCCATCATCCTGTCA	AGTGCGTCTCCACTCAACCT	174

2.3 Molecular cloning

I chose gene *nanog*, *oct4* and *tcf3* to do *in situ* hybridization. The PCR products of medaka *nanog*, *oct4* and *tcf3* were used.

2.3.1 Recovery of PCR products from agarose gel

The PCR products were excised from the gel under UV light and extracted by the QIAquick Gel Extraction Kit (QIAGEN). Briefly, for every 100 mg excised gel, 300 μl of buffer QG was added. Gel was heated at 50°C until the gel slice has completely dissolved. After that, 100 μl of isopropanol was added into the sample. The sample was applied to the QIAquick column and centrifuged for 1 min. The flowthrough was discarded, and the column was washed with buffer PE. Centrifugation was carried out for 1 min at 17900 x g (13000 rpm). To remove residual wash buffer, the QIAquick column was centrifuged once more for 2 min. PCR products were eluted from the column with 30 μl of buffer EB. The concentration of the final product was determined by Nanophotometer (WPA BioWave II+).

2.3.2 TA cloning

Purified PCR products were cloned into pGEM-T Easy vectors (Promega). The ligation reaction consisted of 2X ligation reaction buffer, pGEM-T Easy vector (50 ng), 1-2 μl T4 ligase (3U/ul) and insert. The insert : vector molar ratio was 3 : 1. The total volume of the ligation reaction was adjusted to 10ul or 20ul according to the volume of insert and vector.

2.3.3 Preparation of competent cells (RbCl method)

Stock of E. coli was streaked onto LB plates and incubated overnight at 37° C. A single colony was inoculated into 3 ml of LB broth and incubated overnight at 37° C. The next day, 1 ml of such culture was inoculated into 100 ml of LB broth in a 250 ml flask and incubated at 37°C for 2-3 h until OD_{600} of the bacterial culture got to 0.35-0.4. The bacterial culture was transferred into a 50ml falcon tube and chilled on ice for 15min. The bacteria were spun down by centrifugation at 2700 rpm, 4°C for 20 min. From this step onwards, remaining steps were carried out on ice. The supernatant was discarded, and 20 ml ice-cold sterile TfbI solution (30 mM KAc, 100 mM RbCl, 10 mM $CaCl_2.2H_2O$, 50 mM $MnCl_2.4H_2O$, 15% v/v glycerol) was added in falcon tube. The bacterial pellet was resuspended gentlely in the TfbI solution and chilled on ice for 30 min. Centrifugation was carried out again in order to pellet the bacteria. After supernatant was removed, 2 ml ice-cold sterile TfbII solution (10 mM MOPS, 75 mM $CaCl_2$, 10 mM RbCl, 15% v/v glycerol) was used to resuspend the bacteria. The suspension was dispensed into 100 μl aliquots in 1.5 ml eppendorf tubes and frozen in liquid nitrogen before storing at -80° C. Efficiency of competent cells was tested before they were applied in the experiment. 100 pg, 10 pg and 0 pg plasmids were transformed into 100 μl of those cells respectively. The efficiency was calculated according to: number of colonies per μg of DNA in 100 μl of competent cells.

2.3.4 Transformation of competent cells

Competent cells were thawed on ice for 15 min. Meanwhile, ligation products were also chilled on ice. 5 μl of ligation product was added to the cells, followed by 30 min incubation on ice. The eppendorf tube containing competent cells was transferred quickly from ice to 42° C water bath for 90 s and then transferred back immediately on ice. 1 ml of LB medium was added into the tube, and the bacteria were incubated at 37°C for 40 min for recovery. Bacteria were spun down by centrifugation at 5000 rpm for 1 min. After the supernatant was removed, the bacteria were resuspended in 100 μl of LB medium. The LB ampicillin (100 μg/ml) plate was pre-coated with 50 μl of X-gal (2% 5-bromo-4-chloro-3-20 indolyl-β-D-galactoside) and IPTG (isopropylthio-β-D-galactoside). The bacterial suspension was then spread on the plate. The plate was incubated for 16 h at 37°C.

2.3.5 Minipreps of plasmids (Alkaline lysis method)

Single white colonies were inoculated into 3 ml LB+ ampicillin (100 μg/ml) medium in 15 ml Falcon snap cap tubes. The bacteria were incubated overnight at 37°C. 1.5 ml of bacterial culture was transferred into a 1.5 ml eppendorf tube and centrifuged at 6000 rpm for 5 min. The supernatant was discarded, and the pellet was resuspended in 100 μl of ice-cold buffer P1 (50 mM glucose, 25 mM Tris-Cl pH 8.0, 10 mM EDTA pH 8.0, 100 μg/ml RNase A) by vortex. 200 μl of freshly-prepared alkaline lysis buffer P2 (0.2 N NaOH, 1% w/v SDS) was added, followed by gentle inversions of the tube. 150 μl of

ice-cold buffer P3 (5 M potassium acetate, 5 M acetic acid) was quickly added and gently mixed with the bacterial lysate to stop lysis reaction. After 50 μl of chloroform was added, the eppendorf tube was put on ice for more than 5 min. Then, the bacterial lysate was centrifuged at 15000 rpm for 5 min at 4°C. The supernatant was transferred into a new eppendorf tube. An equal volume of 100% isopropanol was added to precipitate the plasmid DNA. The mixture was incubated on ice for 10 min and centrifuged at 15000 rpm for 5 min at 4°C. The supernatant was removed, and the white pellet was washed by 500 μl of 70% ethanol. After one more centrifugation, ethanol was discarded, leaving the pellet. The plasmid pellet was air-dried and dissolved in 50 μl of TE+ RNase buffer (10 mM Tris-Cl, 1 mM EDTA, pH 7.4, 20 μg/ml RNAse A).

2.3.6 Plasmid screening by restriction enzyme digestion

To identify whether the right insert was cloned in the plasmid, test digestion was performed with appropriate restriction enzymes. Normally, EcoRI (Promega) was used. The digestion reaction consisted of 10X buffer, plasmid and enzyme (10 U/μl). The amount of DNA and enzyme was determined according to the principle that 1 unit enzyme digests 1 μl of DNA at 37 °C for 1 h. The total volume of the digestion reaction was adjusted to 10 or 20 μl. The reaction was incubated at 37 °C for more than 1 h. Finally the digested products were evaluated through gel electrophoresis to check the released insert and backbone.

2.3.7 Sequencing

Desired plasmid clones were sequenced by the BigDye Terminator V3.0 Cycle Sequencing Kit (Applied Biosystems).

Total volume of 5 μl was used in the sequencing reactions. For each reaction, the following reagents were added to a separate 0.2 ml PCR tube: 2 μl of Terminator Ready Reaction Mix (Big Dye ABI), 1 μl of M13 forward (or reverse) primer (10 μM) and 2 μl of template (pGEM plasmid). PCR tubes were placed in the PTC-100 Thermal Cyclers (Bio-Rad) and subjected to the program as follows:

95°C for 5 min;

25 cycles of:

95°C, 10 sec,

50°C, 10 sec,

60°C, 2.5 min.

Extension products were purified by ethanol precipitation. In each tube of product, 2.5 volume of 100% ethanol and 0.1 volume of sodium acetate (3 M, pH 4.6) were added and mixed gently. The tubes were left at room temperature for >15 min (and < 24 hrs) to precipitate products. Thereafter, the precipitate was spun down for a minimum of 20 min at maximum speed. The supernatant was discarded. The final product was washed by 70% ethanol and centrifuged as before for 5 min at maximum speed. Sample was air-dried for 15 min and redissolved in 15 μl of Hidi formamide. The sample was sequenced by using

the Applied Biosystems 3130xl (Applied Biosystems, MA).

2.3.8 Midiprep of plasmids

Bacterial cultures with successful insertions were inoculated into 50 ml of LB+ampicillin medium and incubated overnight at 37°C. Midiprep was performed with Nucleobond AX kit (Macherey-Nagel). Plasmid concentration was determined by Nanophotometer (WPA BioWave II+).

2.4 *In situ* hybridization

All solutions and buffers used from this point were prepared with DEPC-treated water, and reactions were all performed under RNase-free conditions.

2.4.1 Synthesis of RNA probes

As described, the amplified products of medaka *nanog*, *oct4* and *tcf3* were cloned into pGEM-T. The plasmid was linearized with ApaI for the synthesis of RNA probes.

The linearized plasmid DNA was purified by phenol/chloroform extraction. 0.1 volume of Sodium Acetate was added to the digestion product. The template DNA was extracted once with a equal volume of phenol and chloroform/isoamyl alcohol(IAA)(24:1), and then twice more with chloroform/IAA. 2 volume of ethanol was used to precipitate the DNA, and the template was left at -20°C for approximately 1 h. The DNA pellet was spun down, washed with 70% ethanol and redissolved in TE (10 mM Tris-Cl, 1 mM EDTA, pH 7.4). The concentration of linearized product was determined by Nanophotometer (WPA BioWave II+).

The purified linearized plasmid was used as a template for synthesis of RNA probes labelled with digoxigenin (DIG) or fluorescein isothiocyanate (FITC). The *nanog* and *tcf3* probes were labeled with DIG while the *oct4* probes were labeled with FITC. The synthesis reaction was performed at 37°C for 2 h in a total volume of 20 μl. The whole reaction consisted of 1 μg of template DNA, 2 μl of 100 mM dithiothreitol(DTT), 2 μl of 10 mM NTP mix with Dig-UTP/Fluorescein-UTP (Roche), 0.5 μl of RNase inhibitor, 4 μl of 5X transcription buffer, 1 μl of T7 (Sp6) RNA polymerase, and DEPC-treated water. Following the reaction, 1 μl of Turbo RNase-free DNase (Ambion) was added to digest the DNA template at 37°C for 15 min. After digestion, probe was precipitated by 30 μl RNasefree water and 30 μl Lithium Chloride Precipitation Solution (7.5 M lithium chloride, 50 mM EDTA, pH 8.0), washed with 70% ethanol, air-dried and dissolved in 25 μl DEPC-treated water. At last the probe was quantified, diluted to a final concentration of 1ng/μl in hybridization buffer (50% formamide (Sigma), 5xSSC, 50 μg/ml heparin, 0.1%Tween20, 5 mg/ml torula RNA, pH 6.0-6.5) and examined by normal agarose gel running.

2.4.2 Whole mount *in situ* hybridization (WISH)

Medaka ovary was fixed in 4 % PFA (paraformamide) / 0.85 X PTW (prepare 16% PFA as stock) for 48 h at 4°C, washed three times with PBS and stored in 50% formamide/2 X SSC (pH 6.5) at -20°C. After the outer layer membrane was removed, the ovary was transferred into 100% methanol and stored in -20°C for at least one overnight. When hybridization was started, the ovary

was transferred into 24 or 48 well cell culture dish, rehydrated through a series of methanol dilutions: 75% MeOH (methanol)/PTW (0.1% tween in 0.85 X PBS), 50% MeOH/PTW, 25% MeOH/PTW and rinsed 3 times, each for 5min, in 1XPTW. Following rehydration, the ovary was digested by Proteinase K (PTK: 10 μg/ml in PTW, prepare 20 mg/ml for stock; Roche) at room temperature for following probe penetration. After PTK digestion, the ovary was rinsed twice in freshly prepared 2mg/ml glycine/PTW, re-fixed in 4% PFA/PTW for 20 min and washed 5 times in PTW.

For pre-hybridization, the ovary was rinsed in hybridization buffer and incubated at 68°C for 2 h. The probe was diluted in 200 μl hybridization buffer to a final concentration of 1-5 ng/μl and denatured at 80°C for 10 min, followed by 2 min incubation in ice water bath. After the pre-hybridization buffer was removed, the denatured hybridization probe was quickly added to the sample. Hybridization was performed at 68°C in a water bath for 16 h with shaking. On the second day, the ovary was washed in 50% formamide/2xSSCT (diluted from 20 X SSC, pH 7.0, 0.1 % Tween) at 68°C for two times, each for 30 min, 2xSSCT at 68°C for 30 min and 0.2xSSCT at 68°C for two times, each for 30 min. After hybridization, RNase free condition is not required. Subsequently, the ovary was washed twice with PTW at room temperature and blocked with 5% sheep serum/PTW at room temperature for 1 hour on a shaker. After blocking solution was removed, the ovary was incubated with anti-Dig-AP Fab fragments (Roche) at a 1: 2500 dilution in 5% sheep serum/PTW at room

temperature. The ovary was washed in PTW 6 times, each for 15 min and washed again in TBST (100 mM TrisCl, pH 9.5, 100 mM NaCl, 0.1%Tween-20) twice, each for 2 min. Then the ovary was equilibrated for 5 min in staining buffer NTMT (100 mM TrisCl, pH 9.5, 100 mM NaCl, 50 mM $MgCl_2$, 0.1%Tween20, prepared freshly) and incubated with NTMT containing BCIP (final 175μg/ml, stock: 50 mg/ml in 100% DMF) and NBT (final 337.5μg/ml, stock: 75 mg/ml in 70% DMF/H_2O) in darkness without shaking at 4°C from several hours to days until signals appeared. The staining process was stopped by washing with PTW. For immediately photography, the ovary was rinsed in 87% glycerol overnight.

2.4.3 Section *in situ* hybridization (SISH)

The tissue testis was obtained as described. The tissues were fixed in 4% PFA overnight at 4°C. After fixation, the tissues were dehydrated by rinsing in 20% sucrose/PBS overnight, 30% sucrose/PBS for 2 h, 1:1 of 30% sucrose:O.C.T (embedding medium) for 30 min and embedded in O.C.T with liquid nitrogen in a special mould under stereomicroscope. Embedded tissues were kept below -20°C and fixed to a supporting base for cryostat (CM1850, Leica). Section was done at 4 μm for testis. The sections were collected onto pre-cleaned slides (Fisher scientific) and stored at -80°C. For hybridization, the sections were incubated on a 42°C heating block for 30 min and directly rinsed in hybridization buffer. The following steps were similar to WISH.

2.4.4 Fluorescent *in situ* hybridization(FISH)

For embryo FISH, 16-cell stage embryos were obtained as described. The fixation of embryos and hybridization were similar to WISH. After post-hybridization wash, the hybridized embryos were blocked and incubated with anti-fluorescein POD for 4 h at room temperature followed by six washes for 20 min each with PBS. These embryos were then incubated in TSA plus fluorescein solution for 60 min, rinsed orderly in 30%, 50%, 75% and 100% methanol/PBS each for 5 min and then incubated in 1% H_2O_2 in 100% methanol for 30 min. After this incubation, these embryos were rinsed orderly in 75%, 50% and 30% methanol/PBS and washed with PBS 3 times, each for 10 min. For second probe detection, the embryos were blocked again and incubated with anti-DIG POD for 4 h at room temperature followed by 6 washes, each for 20 min, with PBS. These embryos were then incubated in TSA plus Cy3 solution for 60 min followed by 3 washes with PBST. Finally, the embryos were deyolked, stained with DAPI and mounted with anti-fade reagent for photography.

For testicular section FISH, the tissue processing and hybridization were similar to SISH. After post-hybridization wash, the probe detection was performed with the same method as that in embryo FISH. Instead of two different antibodies for two probe detection, only anti-DIG POD was used to detect single *tcf3* probe. The sections were finally stained with DAPI and mounted with anti-fade reagent for photography.

2.4.5 Microscopy and photography

Observation and photography on Leica MZFIII stereo microscope, Zeiss Axiovertinvert 2 invert microscope and Axiovert 200 upright microscope with a Zeiss AxioCam MRc5 digital camera (Zeiss Corp., Germany) were as described previously (Yi et al., 2009).

2.5 Zygotic expression examination

In order to examine the zygotic expression of pluripotency genes, I made use of the sequence difference between two species, *O. latipes* and *O. celebensis*. F1 embryos obtained by crossing *O. celebensis* with *O. latipes* were used. Since the F1 embryos have DNA from both parents, the paternal transcripts can be detected by RT-PCR analysis using paternal-species-specific primer, while maternal transcripts cannot be detected. Therefore zygotic expression can be examined.

2.5.1 Gene cloning in *O. celebensis*

The ovary was obtained as described in the collection of adult tissue. Total RNA was isolated by using the Trizol Reagent (Invitrogen). Synthesis of cDNA templates was primed with oligo $(dT)_{25}$ by using M-MLV reverse transcriptase (Invitrogen).

In order to amplify the fragment of the seven genes in *O. celebensis*, the primers in Table 3 were used in RT-PCR. PCR was run in a 25 µl volume containing 10 ng of cDNA reaction for 37 cycles (20 s at 94ºC, 20 s at 58ºC and 120 s at 72ºC). The PCR products were separated on 1% agarose gels

and documented with a bioimaging system (Synoptics).

The PCR products were purified by QIAquick Gel Extraction Kit (QIAGEN) and cloned into pGEM-T Easy vectors (Promega). The insert was sequenced by the Applied Biosystems 3130xl (Applied Biosystems).

Table 3 Genes and Primers used for gene cloning

Gene	**Primer sequence**	
	Forward primer	Reverse primer
nanog	ATGGCGGAGTGGAAAACTCAGGTC	TCATTGGACAGCATTGTGCAAAATG
oct4	ATGTCTGACAGGCCGCACAGCC	TCATCCTGTCAGGTGACCTACC
klf4	ATGGCGTTAGGCGGAACAC	TCATAAGTGCCTCTTCATGTGGA
sall4	ATTGCACTAAAAGCCAGCAGG	TCTCCAGGCAGTCCACCATC
tcf3	ATGCCTCAACTGAACGGAGG	AGGGCTGCCTGATGGAAC
zfp281	ATGAGTATTATCCAAGACAAGATAGGC	TGACTTGTCATGCTTGAGATCCAG
ronin	ATGCCCGGGTTCACGTG	TCCCGCTGCTCGTTCAG

2.5.2 Sequence analysis and species specific primer design

Sequence alignment between *O. latipes* and *O. celebensis* was carried out in Vector NTI. The regions on the gene fragments with obvious sequence differences were chosen to design species specific primers. The specific-primers designed are shown in Table 4.

Table 4 Genes and Primers used in RT-PCR for examining zygotic expression

Gene		**Primer sequence**	
Name	**species**	Forward primer	Reverse primer
nanog	**Oc**	GTTTCAGAACCGTAGGATGAAGC	AGGTAGAACGCCAAGTTCTGC
oct4	**Oc**	ATATTTCGGCGCAAACGCAG	TCATCCTGTCAGGTGACCTACC
klf4	**Oc**	GAAGGAGGAGCTGTTCCACC	GTACTCTCCTGCATCCATTGCC
sall4	**Oc**	GAGATGTTGAGCCTCCCGC	CCAGTCCCATCTCTTTGCCA
tcf3	**Oc**	GCCTCTCCTGGATGTCCCAG	CTCTGGCGAGAGGTGCGAT
ronin	**Oc**	CTTCCAGCCCACCACAGGT	TTGAGTTCCTCCTCGATGCTGAT
nanog	**Ol**	GTTTCAGAACCGTAGGATGAAGC	AGATAGAAGGCCAGGTTGTGT
oct4	**Ol**	TTTCTTTGGCGTAAACTCGT	TCATCCTGTCAGGTGACCTACC

Oc, *O. celebensis*; Ol, *O. latipes*

2.5.3 *In vitro* fertilization

O. latipes female fish that produced many good eggs were isolated. On day of experiment, *O. latipes* female and *O. celebensis* male fish were anaesthetized in the ice-cold water, and surface-dried by a paper towel. The fish were placed in a 10 cm plastic petri-dish and gonads were individually dissected. Ovaries were put in medaka oocyte medium(MOM; 9.8 g/L medium M199 and antibiotics (Invitrogen), pH 8.0.) in a 60 mm Petri dish, and mature oocytes were released and transferred into 5 ml of MOM in a 35mm dish at room temperature. Work with testis and sperm preparation was done in separate places and with different sets of forceps and tweezers. A single testis was used for up to 40 eggs. The testis of adult *O. celebensis* was isolated, rinsed in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO4, 2 mM KH_2PO_4, pH7.4), transferred into a drop of 100-μl cold PBS in a dish on ice, and manually dissociated with a pair of forceps into single cells and small aggregates. The preparation was transferred into a 1.5 ml eppendorf tube and allowed to settle

down for 5 min on ice. The top supernatant rich in single sperm (20-50 μl) was immediately added to 10-40 eggs in 0.5 ml of MOM for activation. After incubation for 2 min in darkness, the activated eggs were rinsed and incubated at 28ºC in embryo rearing medium. Successful *in vitro* fertilization was examined by two criteria. Firstly, fertilized eggs will become transparent while unfertilized eggs will remain opaque. Second, successfully fertilized embryos will begin development, therefore embryogenesis can be observed according to the time line of embryo development. On the other hand, some unfertilized eggs will stop their development at one-cell stage. The hybrid embryos were kept in 28ºC. Embryos were staged according to Iwamatsu(Iwamatsu, 2004). Several developmental stages are chosen: 16-cell (stage 6), early morula (stage 8), late morula (stage 9), early blastula (stage 10) and late blastula (stage 11). For each stage, 8-10 embryos were collected. The same work was also performed for *O. latipes* male and *O. celebensis* female fish.

2.5.4 Zygotic expression examination

As described, total RNA of hybrid embryos was isolated by the Trizol Reagent (Invitrogen). Synthesis of cDNA templates was primed with oligo $(dT)_{25}$ by using M-MLV reverse transcriptase (Invitrogen). PCR was run in a 25 μl volume containing 10 ng of cDNA reaction for 30 and 35 cycles (20 s at 94ºC, 20 s at 58ºC and 60 s at 72ºC) for β-actin and other genes, respectively. Primers are listed in Table 4. The PCR products were separated on 1% agarose gels and documented by a bioimaging system (Synoptics).

CHAPTER 3: RESULTS

3.1 Gene identification

I adopted a homology approach to identify medaka pluripotency genes by taking advantage of genome sequence data. Through BLAST searches against public database, putative homologs/orthologs of several mammalian pluripotency genes including *nanog, oct4, klf4, sall4, zfp281, ronin* and *tcf3* have previously been identified *(Yi et al., 2009).* In literature and several annotated genomes, *tcf3* (also *tcf7l1*) is confused with *e2a* (also called *e12* or *e47*) encoding transcription factor E2A (E2A immunoglobulin enhancer binding factor). The improvement in the genome sequence and the availability of more cDNA sequences led to revisit gene identification in medaka. Extensive reciprocal BLAST searches and sequence alignments revealed that the two previous *tcf3* genes actually encode E2A proteins instead (Figure 1).

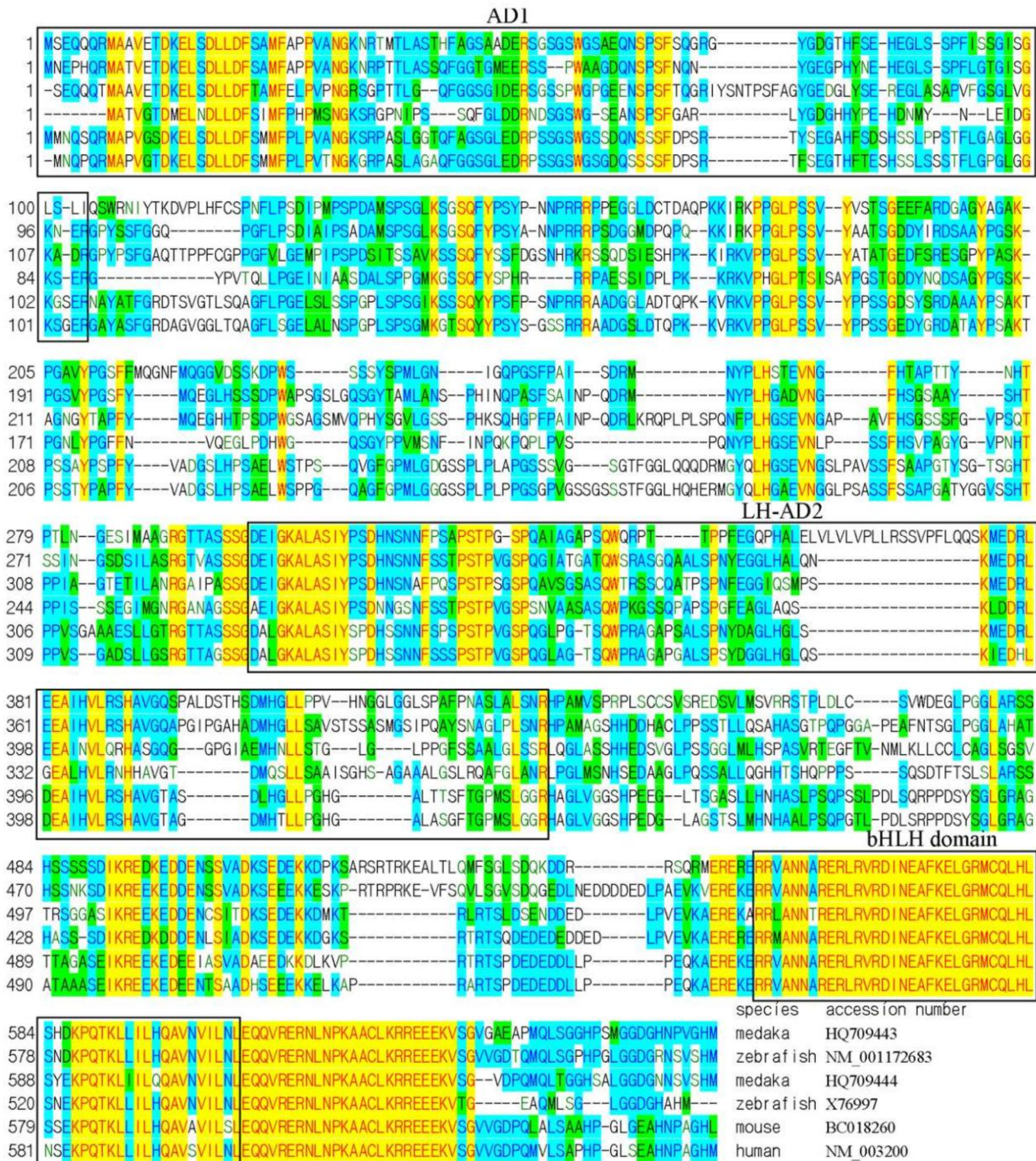

Figure 1. Sequence comparison of E2A (E12/E47) proteins on alignment. Species names and gene accession numbers are given at the end of alignment. The two medaka e2a genes, e2a1 (HQ709443) and e2a2 (HQ709444), have been submitted to GenBank. The activation domain 1 (AD1), the loop-helix activation domain 2 (LH-AD2) and the basic helix–loop–helix (bHLH) are highlighted in frame, respectively.

Meanwhile the new *tcf3* was identified. However, only single *tcf3* gene is found in the medaka genome, while two *tcf3* genes, namely *tcf3a* and *tcf3b*, exist in zebrafish(Dorsky et al., 2003; Kim et al., 2000). To see whether the medaka *tcf3* is homologous or orthologous to the mammalian gene, I compared its protein sequence and its chromosome location. On sequence alignment, the medaka Tcf3 protein shows a high overall conservation and has a beta-catenin binding domain and a class-I HMG box (Figure 2A). By pairwise comparison, the medaka Tcf3 is 65.7% and 67.5% identical to the mouse and human Tcf3 protein, respectively (Figure 2A). On a phylogenetic tree, the medaka Tcf3 clusters with known zebrafish Tcf3 protein, while the mouse and human Tcf3 proteins form a second cluster (Figure 2B). More importantly, medaka *tcf3* exhibits a syntenic relationship to its human counterpart (Figure 2C). These data suggest that the medaka *tcf3* identified in this study is a homolog of the mammalian *tcf3* gene.

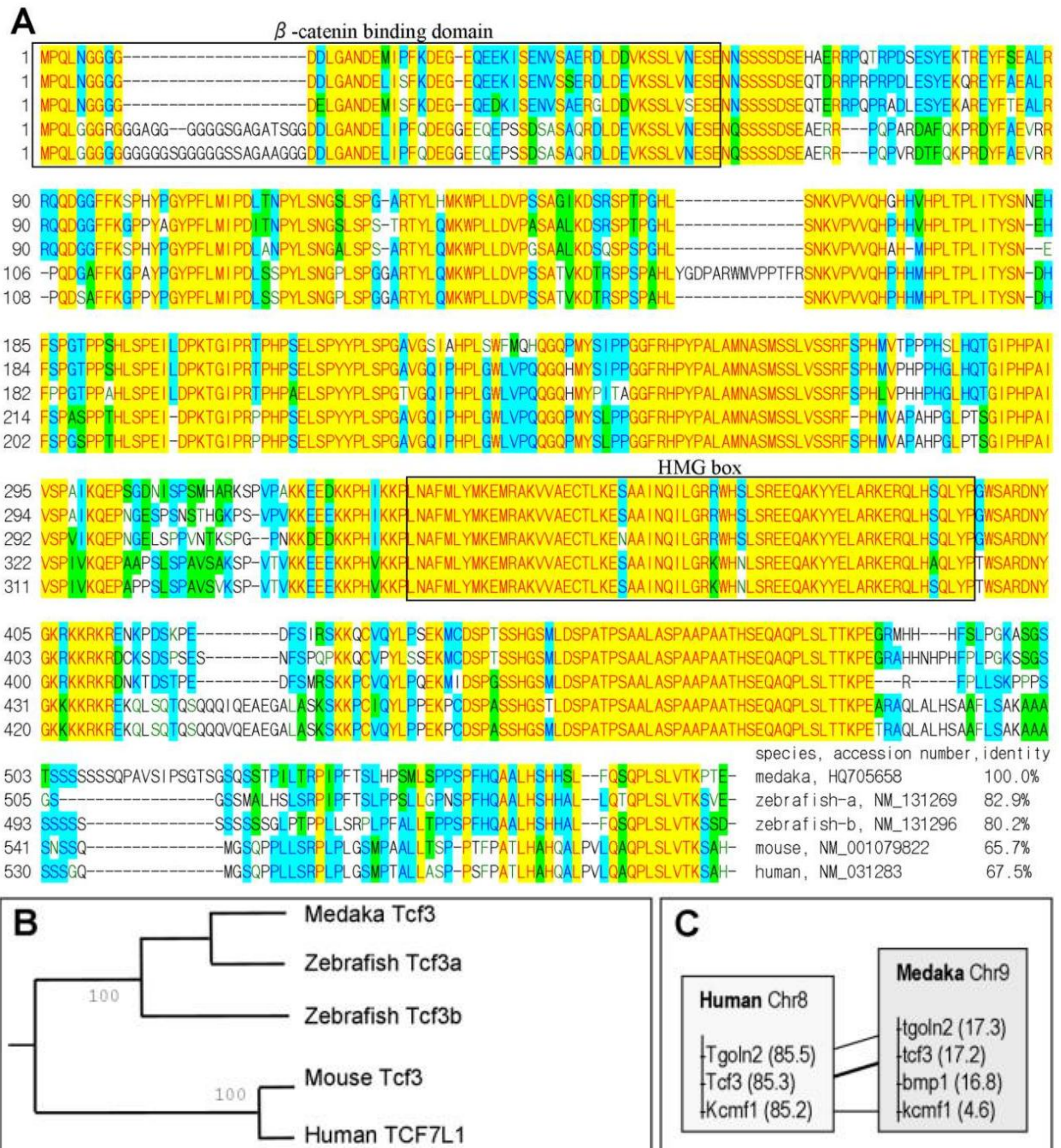

Figure 2. Phylogenetic comparison of Tcf3/Tcf7l1 proteins. (A) Sequence alignment. At the end of the alignment are species, gene accession numbers and amino acid identity values. The β-catenin binding domain and high mobility group (HMG) box are highlighted in frame. (B) Phylogenetic tree of Tcf3/TCF7l1 proteins. (C) Chromosome location of the tcf3/tcf7l1 gene in human and medaka. Chromosomal positions are in parenthesis.

3.2 Expression pattern analysis

3.2.1 RT-PCR analysis of expression in tissues

There is a set of seven medaka genes in this study. I carried out sequence analysis using CDS and genomic sequence of these pluripotency genes, to identify exon and intron within the genes. Based on the sequence analysis results, intron-spanning gene specific primers were designed. By using primers listed in Table 1, RT-PCR analysis was performed to analyze their expression patterns in eight representative tissues: brain, skin, heart, kidney, liver, gut, testis and ovary. According to expression patterns in adult tissues, the seven genes fall into three groups (Figure 3): Group I contains *nanog* and *oct4*, which show gonad-specific expression; Group II comprises *sall4* and *zfp281*, which display high expression in the gonads and detectable expression in 1 or 2 somatic tissues; Group III consists of *klf4, ronin* and *tcf3*, which exhibit wide expression at variable levels in 6 to 7 tissues. The expression patterns of the seven genes are compared in more detail in Table 5.

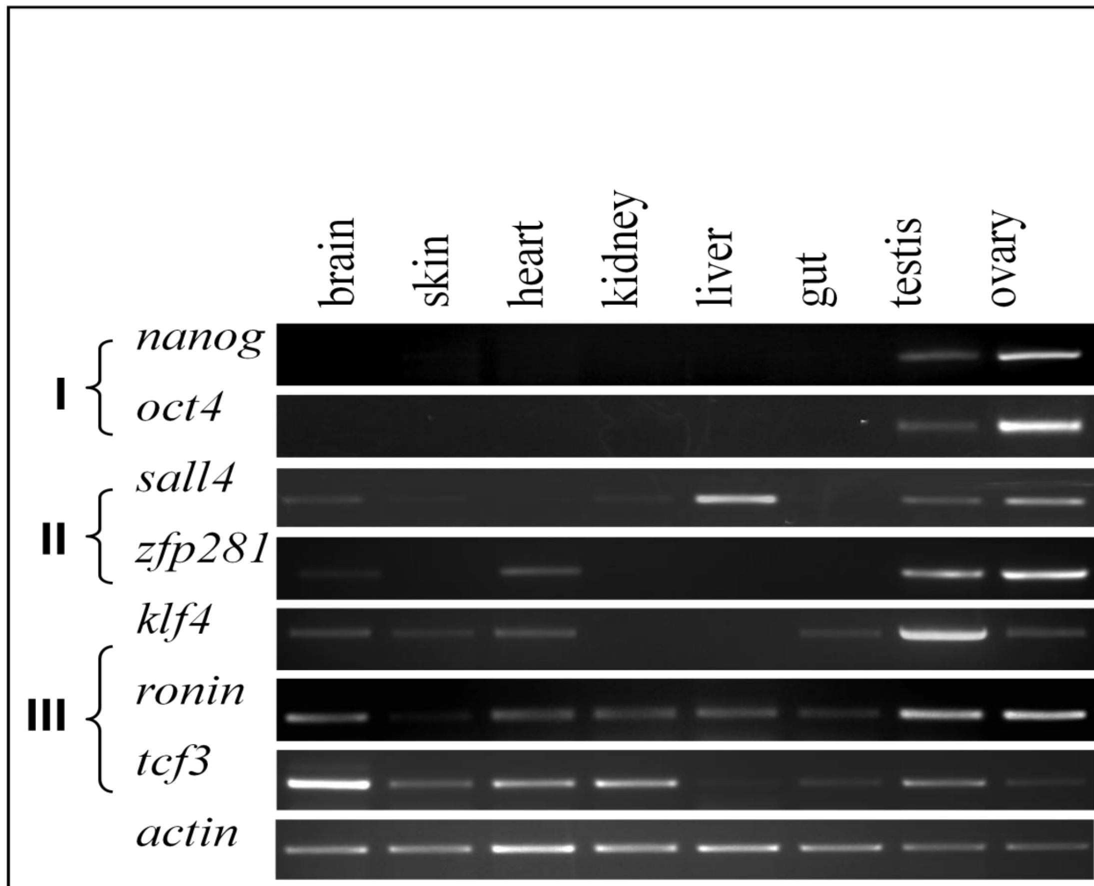

Figure 3. RT-PCR analysis of RNA expression in adult tissues

Table 5 Summary of RNA expression in medaka adult tissues

Gene	brain	skin	heart	kidney	liver	gut	testis	ovary
Oct4	-	-	-	-	-	-	++	+++
Nanog	-	-	-	-	-	-	++	+++
klf4	+	+	+			+	++++	+
sall4	+	+	-	+	+++		+	++
tcf3	+++	++	+++	+++	+	++	+++	++
zfp281a	+	-	++	-	-	-	++	+++
ronin	++	+	++	++	++	+	+++	+++

Relative levels of expression are indicated by -, +, ++ and +++ for barely detectable, faintly detectable, easily detectable and high expression

3.2.2 RT-PCR analysis of expression in embryos

RT-PCR analysis was performed to analyze their expression patterns in seven critical stages of embryos: 16-cell, morula, early blastula, late blastula, pre-mid gastrula, 34 somite and prehatch. The transcripts of the seven genes were easily detected in developing embryos ranging from cleavages to the late blastula stage (Figure 4), suggesting that their RNAs are maternally supplied since zygotic transcription in medaka does not occur until the midblastula transition (Aizawa et al., 2003). When embryonic development proceeds through gastrulation to the prehatching stage, differences in RNA level became apparent among the genes: *nanog* and *oct4* sharply declined and ultimately became undetectable; *sall4, zfp281* and *klf4* exhibited a dramatic decrease, whereas *ronin* and *tcf3* displayed little change. Since cell lineage restriction and differentiation occurs during and after gastrulation, expression patterns shown by the RT-PCR analysis suggest that *nanog* and *oct4* appear to be expressed mainly in pluripotent cells of developing embryos, while the other genes are expressed and required in other events of embryonic development.

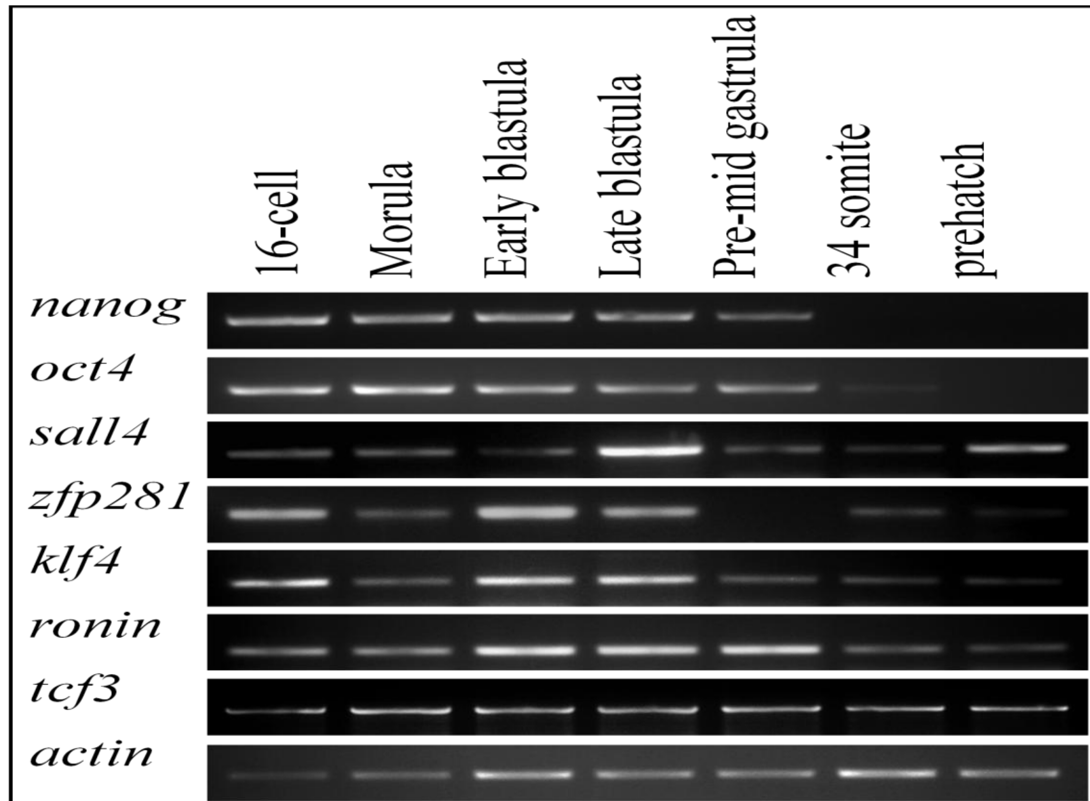

Figure 4. RT-PCR analysis of RNA expression in embryos.

3.2.3 *in vivo* expression in embryos and the adult gonad

In the mouse, the totipotent cycle begins with the zygote, proceeds through a series of transient structures (morula, inner cell mass and epiblast) and ends with the germline, which establishes PGCs and ultimately produces eggs and sperm in the adult gonads, the ovary in the female and the testis in the male(Pesce et al., 1998a). All the transient embryonic structures and PGCs have the potential to produce pluripotent stem cell cultures. Even adult spermatogonia, the male germ stem cells, are also capable of generating stem cell cultures in fish (Hong et al., 2004a) and mouse (Guan et al., 2006). Therefore, I made use of early embryos and adult gonads as *in vivo* systems to

study gene expression in stem cells and their differentiated derivatives.

Through the two color fluorescent *in situ* hybridization procedure (Liu et al., 2009; Xu et al., 2009), the transcripts of both *nanog* and *oct4* were found to be present in blastomeres of the 16-cell staged embryos (Figure 5A-C), in accordance with a high level of their RNAs by RT-PCR analysis (Figure 4). It was also observed that a majority (75%, n = 12) of 16-cell embryos exhibited a significantly more intense signal for both *nanog* and *oct4* in the four inside blastomeres compared to the 12 peripheral blastomeres (Figure 5A-C). Because of the early expression of both *nanog* and *oct4,* the expression of the two genes is associated with embryonic pluripotency.

In medaka, the female produces eggs every day, and the mature ovary comprises a small number of oogonia and oocytes at ten ontogenic stages (Iwamatsu et al., 1988). The adult medaka testis is composed of spermatogonia and differentiating spermatogenic cells in the seminiferous cysts. Spermatogonia are single cells or small clusters and seminiferous cysts comprise germ cells at various stages of spermatogenesis. Spermatogenesis proceeds synchronously in each cyst through primary spermatocytes (meiosis I), secondary spermatocytes (meiosis II), spermatids (meiosis completion) and sperm (terminally morphological differentiation). The testis is a highly ordered organ. Spermatogonia are located at the outmost periphery. The cysts consisting of germ cells at the most advanced stages of development reside closer to the efferent duct, which is in the central region. The unique

architecture and bipolarity of spermatogenesis make possible the unambiguous distinction of stem cells and of their progressive differentiation.

I chose *nanog* as a representative of group-I genes for chemical *in situ* hybridization in the ovary and testis. In the ovary, *nanog* was found to be expressed at a high level in the female germ stem cells oogonia (Figure 5D). A weaker signal was also detected in small oocytes (Figure 5E). In the testis, the *nanog* signal is detectable only in spermatogonia (Figure 5F). In both ovary and testis, *nanog* expression is absent in somatic cells (Figure 5D-F).

I also examined the expression of *tcf3* in the testis. The *tcf3* signal appears moderate in spermatogonia, peaks in spermatocytes and declines in spermatids and sperm (Figure 5G). The wide *tcf3* expression in various stages of spermatogenesis is coincident with its stronger expression in the testis than in the ovary, as revealed by RT-PCR analysis (Figure 3). Apparently, *tcf3* expression in the testis may be important for spermatogenesis progression.

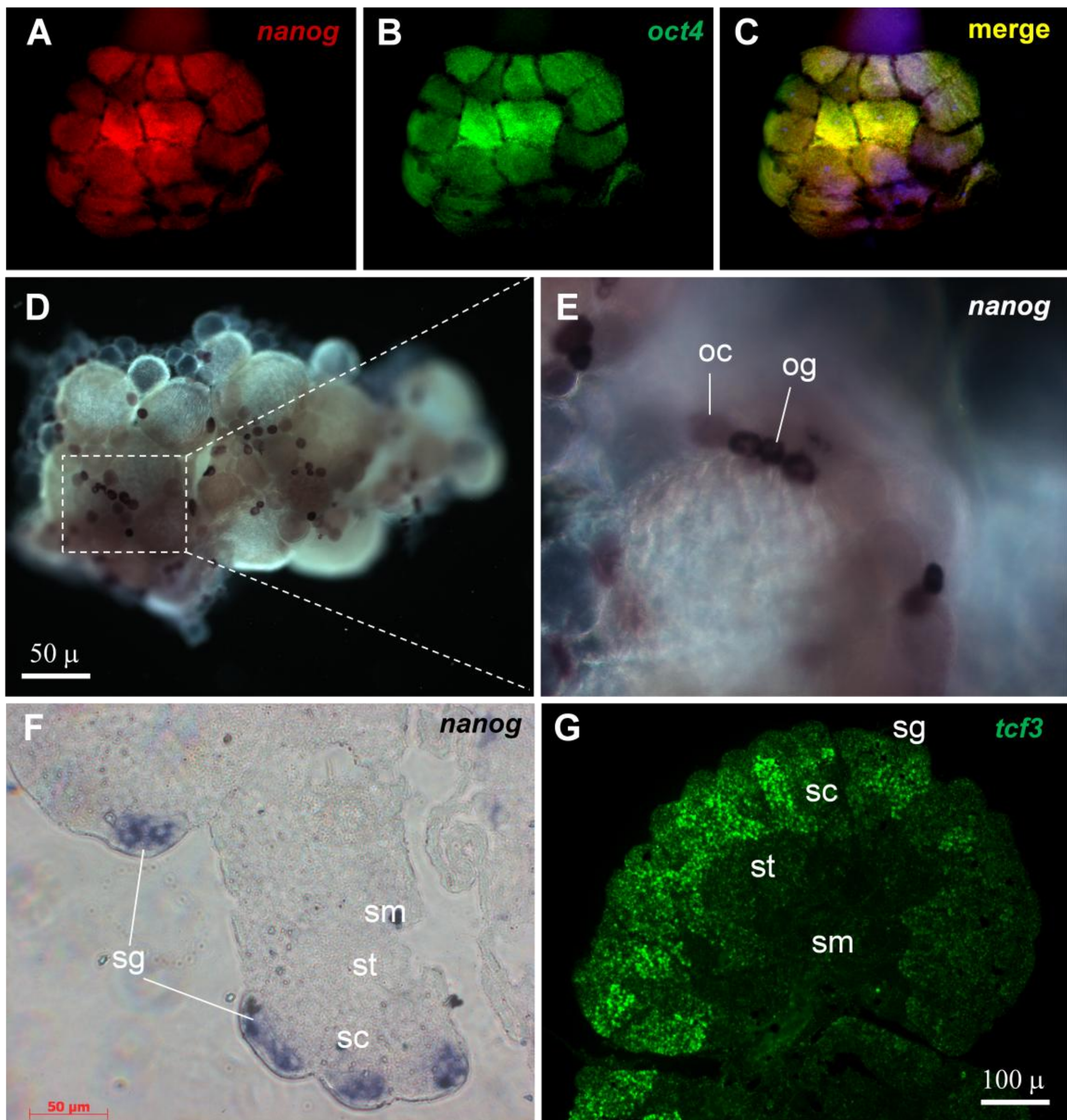

Figure 5. RNA expression by in situ hybridization. (A-C) 16-cell embryo after two-color in situ hybridization showing *nanog* (red) and *oct4* (green) RNA expression. (D-E) Adult ovary after chemical in situ hybridization with a *nanog* riboprobe. (F) Adult testicular section after chemical in situ hybridization with a *nanog* riboprobe. (G) Adult testicular section after fluorescent in situ hybridization with a *tcf3* probe.

og, oogonia; oc, oocytes; sg, spermatogonia; sc, spermatocytes; st, spermatids; sm, sperm.

3.2.4 Expression in ES cell culture

Finally, I made use of ES cell cultures to validate the candidacy of the seven medaka genes as pluripotency markers. I used well characterized medaka ES cell lines. MES1 is a diploid ES cell line from fertilization blastulae(Hong et al., 1996) and capable of chimera formation at high efficacy (Hong et al., 2010; Hong et al., 1998). HX1 is a haploid ES cell line from gynogenetic blastulae and capable of whole animal production by semicloning(Yi et al., 2009; Yi et al., 2010a). Both cell lines were maintained under undifferentiated conditions or subjected to suspension culture for embryoid body formation to induce differentiation(Hong et al., 1996; Yi et al., 2009). Undifferentiated ES cells and differentiated samples were analyzed for gene expression by RT-PCR. The expression of all the seven genes was high in undifferentiated cultures of both ES cell lines but dramatically reduced after 7 days of induced differentiation in suspension culture (Figure 6).

Successful induction of differentiation was evaluated by the expression of *no tail* (ntl), a mesendodermal marker encoding the T-box containing protein Brachyury(Araki et al., 2001). In this study, *ntl* expression was low in ES cells before differentiation but high upon differentiation (Figure 6). This result is consistent with what has been previously reported (Yi et al., 2009). Therefore, all the seven genes appear to be markers for pluripotency in undifferentiated ES cell cultures.

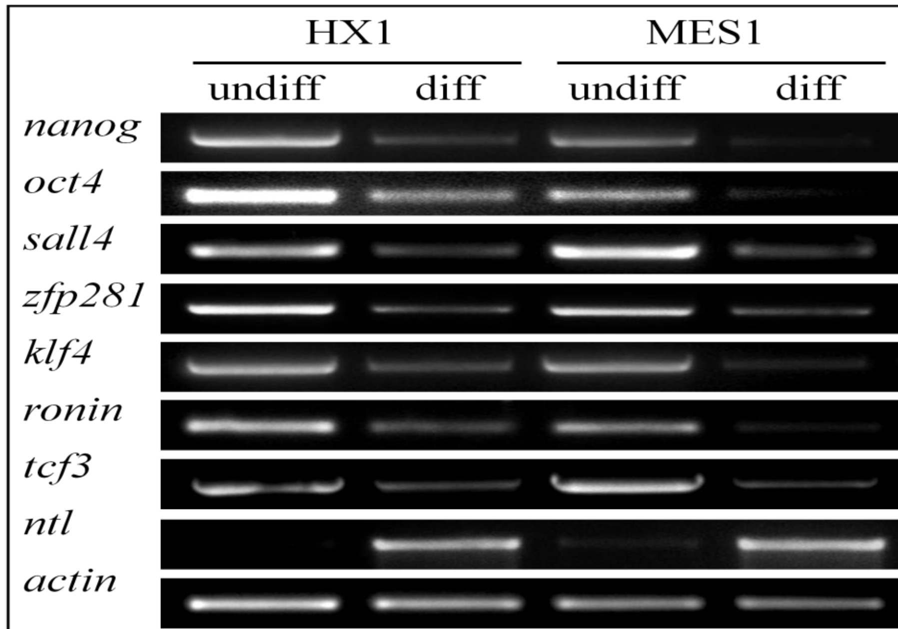

Figure 6. RNA expression in ES cell culture(RT-PCR). ES cells were maintained in adherent culture for undifferentiated (undiff) growth or in suspension culture for 10 days for induced differentiation (diff) by formation of embryoid bodies. (A) RT-PCR analysis in haploid (HX1) and diploid ES cell line (MES1). The mesendodermal marker ntl was used as differentiation marker β-actin was used as a loading control.

3.2.5 Quantitative-PCR analysis of expression in ES cell culture

To investigate gene expression before and after ES cell differentiation in more detail, I analyzed expression of *nanog* and *oct4* by real-time PCR in MES1 cells. We found that the transcripts of both genes were reduced by ≥4-fold (Figure 7), reinforcing that the the seven genes, particularly *nanog* and *oct4*, are useful for monitoring pluripotency in fish stem cell culture.

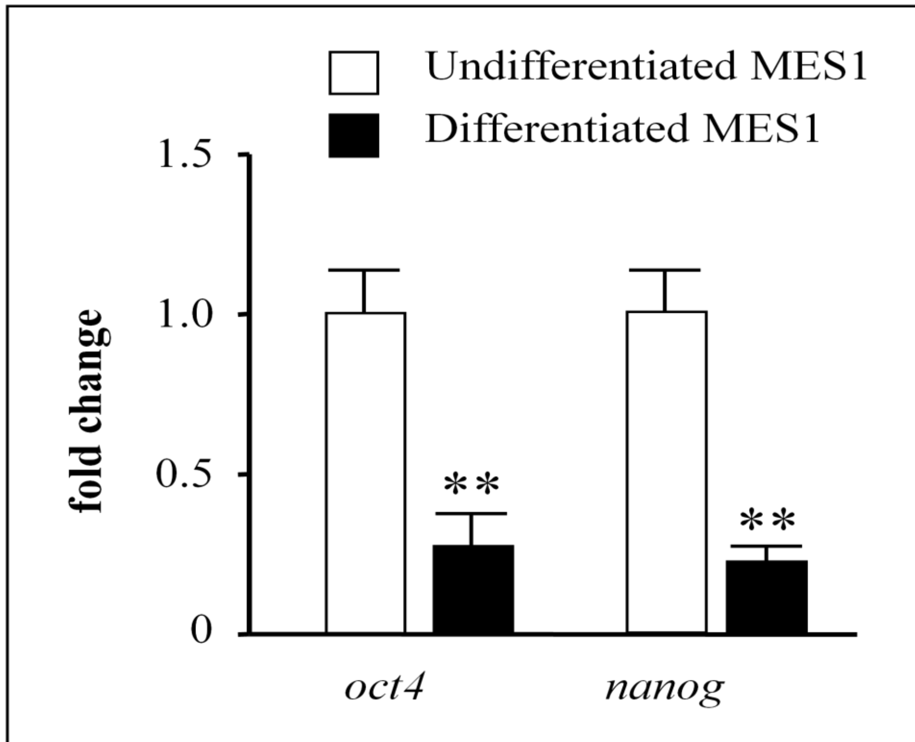

Figure 7. RNA expression in ES cell culture (qPCR). Real-time RT-PCR analysis of Oct4 and Nanog RNA expression. Data are means ± s.d (bars above columns) of three samples; **, p ≤ 0.01.

3.3 Zygotic expression examination

3.3.1 Gene Cloning in *O. celebensis*

In this study, in order to monitor the onset of the zygotic expression of the pluripotency genes, I planned to make use of the sequence differences between two related species: *O. latipes* and *O. celebensis*. Since there is no enough information about the sequence of certain genes in the species *O. celebensis*, the first job is to clone certain genes in *O. celebensis*. Normally, degenerate primer should be used in cloning genes in new species. However, in order to get enough information about the sequence of certain genes in *O. celebensis*, CDS full length primers from *O. latipes* were first used. PCR results showed positive band in *klf4, zfp281a, tcf3, oct4* and *nanog*. For the other two genes, *sall4* and *ronin*, primers at conserved regions were designed. The DNA fragments from both species were also successfully amplified (Figure. 9). The amplified PCR fragments were cloned into pGEM T vector for further sequencing.

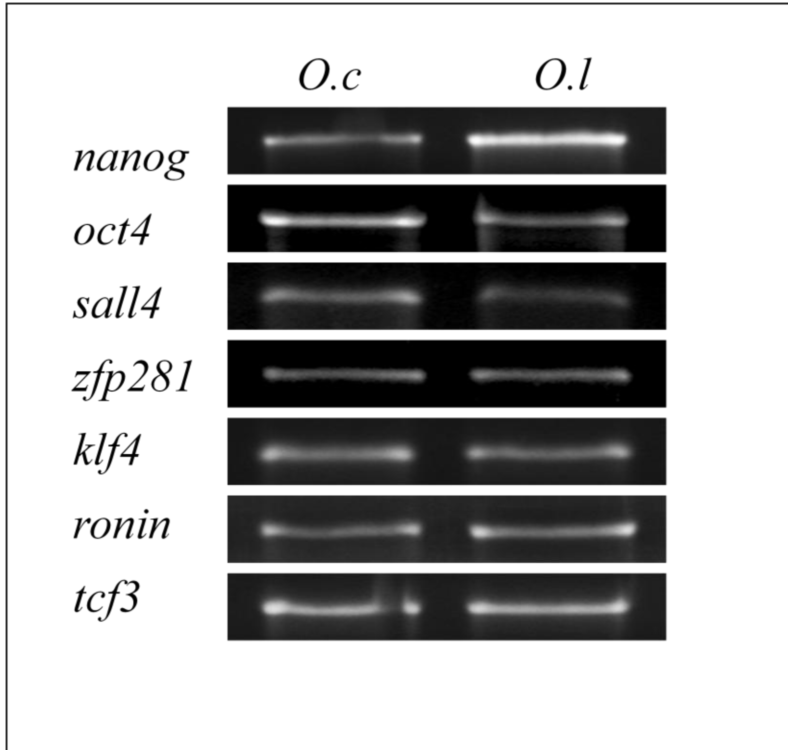

Figure 8. RT-PCR results of seven genes in *O. latipes* and *O. celebensis*

3.3.2 Sequence analysis

The sequence of those pluripotency genes was analyzed. The sequence alignment between *O. latipes* and *O. celebensis* was performed (Figure 9-16). Among those seven genes, the sequences of *nanog, oct4, sall4, klf4, ronin* and *tcf3* show obvious sequence differences which made it possible to design the species-specific primer. The regions chose to design species-specific primers were marked in Figure 9-16. Only the sequences of the genes *zfp281* from both species are so conserved that it is difficult to design species-spesific primer.

```
oc   (628) GAGAGCCAGATGAACACTCTGGTGCAGCGCTTCAACATGCAGAGGTACCTCACTNCCAGC
ol   (652) GAGAGCCAGATGAGCACTCTGGTGCAGCGCTTCAGCGTGCAGAGGTACCTCGCC-CCAGC

oc   (688) TGAGATGAAGANCCTGGCAGATGTGACTGGACTCACCTACAAGCAGATAAAAACATGGTT
ol   (711) TGAGATGAAGAACCTGGCAGATGTGACTGGACTCACCTACAAGCAGGTGAAAACATGGTT

oc   (748) TCAGAACCGTAGGATGAAGCTGAGGAGGCATCAGAAAGACACCAGCTGGGTTTCAGAGCG
ol   (771) TCAGAACCGTAGGATGAAGCTTAGGAGGCATCAGAAGGACACCAGCTGGGTTTCAGAGCG

oc   (808) ATATGCAATCAAAAAAGATGACACTGCTACTAACGCTGCATTCACAAACATGGCTCCACA
ol   (831) ATATACAATCAACAAGGACAACACGGCTGCTGACACTGTATTCTCAAACGTGGCTCCACA

oc   (868) TGTACCCTCTTATCAGGGGGATGGGACGTCCCATCTTCGGGATCACTACAACCAGCACAT
ol   (891) TGTCCCCCCTTATCAGGGGGATGGGATGTCCCATCTGCGGCATCACTACAACCAGCACAT

oc   (928) GATGAGTACAACTTTTAAGAACGCCCCGCAGAACTTGGCGTTCTACCTGGCTGCCATGGG
ol   (951) GATGGGGCAGCTTTCAAGAATACCCCACACAACCTGGCCTTCTATCTGGCTGCCATGGG

oc   (988) CAGCCCCACTGGAACTGCTGCTTACCCCCCGTGGTCTTCCAGCCCACCCCAAGCTCCTGT
ol  (1011) TAACCCCCCTGGAACTGCTGGTTACCCGCCATGGTCTTCCAGCCCCCCCCAGGCTGCGGT

oc  (1048) GACCAGCAGGCCCCAGGTACCAGGATGGCCGCTGCCGCCCGCCTGCAGTCAGTTTGAATT
ol  (1071) GCCCAGCAGACCCCAGGTACCAGGATGGCCCCTGCCGCCAGGCCGCAGTCAGTTTGGATT

oc  (1108) TCGCCCGCTTTCATACAACTCCGCCGGCGCTGCCTCTTTAAACACCTTTGTGCGCAACGC
ol  (1131) CTGCCCAATTCCATACGACCCCTCCGACGCCGCCTCTTTGAACAACTTTGAGCGCAACGC

oc  (1168) AAGCCTCGACGGCAAAGAGGGGGAGTCTGTCGGAGGTCTCAATGCAGCCATTTTGCANNN
ol  (1191) AATCCCCGACAGCAAAGACGGGGAGTCTGCTGGGGGTGCGAATGCAGCCATTTTGCACAA
```

Figure 9. Alignment of *nanog* cDNAs between *O. celebensis* and *O. latipes*.

```
1     -----------CCCGTNCATGTCCCNNTCACCCCCGANNNNNGNGNCCCCGGGGNCCCCT
721   TCCATACCCTCCCCATCCATGTCTCCTTCGCCCCCAAACAACGCGGCCCCAGGGGCCCCG

50    GCATATTTCGGCGCAAACGCAGCCCAAGGAGNCGNGGAGCCGCAAGCGCAG------ACC
781   GCTTTCTTTGGCGTAAACTCGTCTCAAGGAGCCGCGGAGCCGCAAGCGCAGAACCCGACC

104   TNGACGNGGAGCAGCGGNTCATCNAGCGGNGGCTGCAGCGANTCTGAGGAGGAGAACCTG
841   TCCACGCGGAGCAGCGGGTCATCCAGTGGGGGCTGCAGCGACTCTGAGGAGGAGAACCTT

164   TCANNTGAGGAGTNGGAGCAGTTNGCCAAGGAGCTGAAGCACAAAAGNATCACTTTGGGT
901   TCAACAGAGGAGCTGGAGCAGTTCGCGAAGGAGCTGAAACACAAAAGGATCACTTTGGGT

224   TTTACACAGGCAGATGTTGNCCTTGCGTTGGGGAATCTCTNCGGCAAGATGTTCAGCCAG
961   TTTACGCAAGCAGATGTTGGCCTTGCATTGGGGAATCTATATGGTAAGATGTTCAGCCAG

284   ACAACAATTTGTCGCTTTGAGGCCCTCCAGCTGAGCTTCTAGAACATGTGCAAGCTGAAG
1021  ACAACAATTTGCCGCTTTGAGGCTTTACAGCTGAGCTTCAAGAACATGTGCAAGCTGAAG

344   CCTCTTCTGCAAAGATGGCTGGATGAAGCAGAGACCTCAGAAAATCCCCAGGATATGTAC
1081  CCCCTTCTCCAGAGATGGCTAGATGAAGCAGAGACTTCAGAAAATCCCCAGGACATGTAC

404   AAAATTGAGCGCGTATTTGCTGACACCAGGAAGAGAANNCNNNGGACCAGTCTGGAAGGA
1141  AAAATTGAGCGCGTATTTGCTGACACCAGGAAGAGGAAGCGGAGGACCAGTCTGGAGGGA

464   GCGGTGCGTTCTGCTCTAGANNNNNANNNNNTCAAGTGCCCTAAACCAAATACCCAGGAG
1201  GCGGTGCGTTCTGCTCTGGAGGCGTACTTCATCAAGTGCCCTAAACCAAATACTCAGGAG

524   ATCACGCACATATCAGACGATCTGGGGTTGGAGAGAGACGTGGTGCGTGTTTGGTTCTGC
1261  ATCACACACATATCAGACGATCTGGGTTTGGAGAGAGACGTGGTGCGTGTTTGGTTCTGC

584   AACCGGAGACAAAAAGGAAAGCGTTTAGCCTTGCCACTAGATGAGGAGGGTGATATCCAG
1321  AACCGGAGACAAAAAGGAAAGCGTTTGGCCTTGCCCCTAGATGAGGAGGGCGATATCCAG

644   TACTATGAGCAGAGTGCTTCTCCACTCAACCTGGCGCCTTCTCCCATTTCCAGTCAAGGC
1381  TACTATGAGCAGAGTGCGTCTCCACTCAACCTGGCACATTCTCCCATTACCACTCAAGGC

704   TACCCGCCCTCGGGCTATCCTGGAGCCCCTCCACCACTCTACATGCCCCCGCTTCACCGA
1441  TACCCACCGTCTGGATATCCCGGAGCCCCCCCACAACTCTACATGCCCCAGCTTCACCGA

764   CCTGATGTCATGAAACCAGGCCTGCACCCTGGACTGGTNNNNNNCCTGACAGGATGA
1501  CCTGATGTAATGAAACCAGGCCTGCATCCTGGACTGGTAGGTCACCTGACAGGATGA
```

Figure 10. Alignment of *oct4* cDNAs between *O. celebensis* and *O. latipes*.

```
37   GGGCGTTNNCCGANCTGGCCAGCCGTTTACCGGNCCCTGCTGCCNCCAGTCNNTCAGGTG
900  GGGCGTTC-CCGAACTGGCCAGCCGTTTACCAG-CACTGCTGCC-CCAGTCTC-CAAATG

97   TCCTGTCTTTCCCCGGCTCCTTCAANGGAATGCAAA-CAGGGCTAGATTCCTCAAAGAAG
956  TCTTAGCTTTCCCCGGCTCCTTCAATGGTATGCAAA-CGGGGCTAGATTCCTCCAAGAAG

156  TTAAAGTCTGAGATGTTGAGCCTCCCGCCAGAATCAAAGAATGTGGAGTCATTGTACAAG
1015 TTGAAGTCCAAAATGCTGAACCTTCCCGCAGAATCAAAGAATGTGGAGTCATTGTACAAG

216  CACAAGTGTAAATACTGTGGAAAGACCTTTGGAAATGACAGTGCCCTCCAGATACACTTG
1075 CATAAGTGTAAATACTGTGGGAAGACCTTTGGGAACGACAGTGCCCTCCAGATTCACCTT

276  CGCTCTCACACTGGAGAGAGGCCTTTCAAGTGCAACATCTGTGGAAACCGCTTCACCACC
1135 CGCTCTCACACCGGAGAGAGACCCTTCAAGTGCAACATCTGTGGAAACCGCTTCACAACC

336  AAAGGAAACCTCAAAGTGCATTTCCAGAGACATAAAGACAAATACCCAAACATCAGCATG
1195 AAAGGAAACCTGAAAGTGCATTTCCAGAGGCATAAAGACAAATACCCAAACATCAGCATG

396  AATCCTCATCCTGTACCGGAGCACCTCGACAATATTCCAACTAGCAGTGGGATTCCATTT
1255 AACCCTCATCCTGTTCCGGAGCACCTTGACAATATTCCCACCAGCAGTGGGATTCCTTTT

456  GGTATGTCTGTACCCCTGGAGGAATCAAACCTGGCTGAAATGAAACCTATCCTCAGCCAT
1315 GGTATGTCCGTACCCCTGGAGGAATCGAACCTGACTGAAATGAAGCCTGTCCTTAGTCAC

516  CCTGCTGCTCAATTCAACCCGTCCTCCATACCTGGTTTCAAAACCTTTGATAGTTTTGGG
1375 CCTGCTGCTGGATTCAACCCACCCTCCATACATGGTTTCAAAACCTTTGATAGTTTTGGG

576  GGGCCATCGCCATCGGCAAATGACTCCCCTACACCTG------TGGCCTCTATGTTTGGC
1435 GGGCCGTCGCCATCGCCAAACGATGGCGCTACACCTGCGGCCCTGGCCTCTGTGTTCAGC

630  AAAGAGATGGGACTGGATCCAAGTCCGAAAGATGCCAAAGAGTTGTTGGGAGCGCTGCAT
1495 AAAGAGACGGGACTGGATCCAAGTCCGAAGGATGCCAAAGAGTTCCTGGGAGCGCTGCAT

690  CACGCCAACCCGCTTCCTGGAGAGCAGAGCTCTGGAACAGCAAAACTGCAGCAGATGGTG
1555 CACCCCAACCCGCTTCCTGCAGAGCAGAGCTCCGGAACAGCTAAACTGCAGCAGATGGTG

750  GACTGCCTGGAGAA-ATCACTAGTGA
1615 GACTGCCTGGAGAAGAGAACCAATGA
```

Figure 11. Alignment of *sall4* cDNAs between *O. celebensis* and *O. latipes*.

```
oc 119 TGCCGTCCATATCCACGTTCGCCAGCGGCCCGGTCGACAAGAACAGAACCATCGGAAGCG
ol  23 TGCCGTCCATATCCACGTTCGCCAGCGGCCCGGTCGACAAGAGCAGAACCATCGGAAGCG

oc 179 CCTACATTAAATGGAAGGAGGAGCTGTTCCACCTTAAGAGATCCAGTCTTCCTGCAGGAA
ol  83 CCTACATTAAATGGAAGGAGGAGCTGTTCCATGTTAAGAGATCCAGTGTTCCTGCAGGAA

oc 239 GTAGCTGCCCTGACCTGGACCTTCCCACCGCCACCAGTGCAACCAGCACCAGCATGTCCA
ol 143 GTAGCTGCTCTGACCTGGACCTTCCCACCGCCACCAG---------------CATGTCCA

oc 299 AGAAAGACCCAGAACCCGATCTGGACTTAGACTATGACTTTATCCTGTCAAACTCCATCC
ol 188 AGAAAGACCCAGAACCTGATCTGGACTTAGACTATGACTTTATTTTGTCCAACTCCATCC

oc 359 TGCAGCAGCAGCAGCAGCAGCAACGGCAGCAGGAGGACGGCGGCATGGCATGCGGCTCTG
ol 248 ---------------------------AGGAGGATGGCGGTGGCCTGGCGTGCGGCTCTG

oc 419 CCCGGGCACTGTCTCCCTCCCNNNNNTCTTTCTCCTCATCTTACCAGCTTCCCTCCCCTC
ol 281 CCCGGGTGCTGTCTCCCTCCCCCGGCTCCTTCTCCTCGTCTTTCCAGCTCCCTTCCCCTC

oc 479 AGGACGCCGCCAGCGAGCTTCTGTACACCATCCCGGACATCAGCGATGTTTCTCCCTCCG
ol 341 AGGACGCCGCCAGTGAGCTCTTGTACACCATCCCAGACATCAGCGACGTTTCTCCCTCGG

oc 539 GAGGATTCGTGGCTGAACTCATGAGACCTGAGCTGGACCCTGCATACCTCCACCCCACCG
ol 401 GAGGATTCGTGGCTGAACTCATGAGACCCGAGCTGGACCCCGCGTATCTCCATCCCACCG

oc 599 GCGTCCACGGCAAGTTTGTTGTCAAGGCGGCAATGGATGCAGGAGAGTACGGCAAGGCTT
ol 461 GGGTCCACGGCAAGTTTGTCGTCAAGGCAACAATGGAGGCGGGAGAGTACGGCAAGGCCT

oc 659 GCGCCAATGCTGGGTCGGACTCCCGAGTGTCTCCGTCCTTCCC-TGCCACAGAATAAAGC
ol 521 GCGCCAACGCCGGGTCGGACTCGCGAGTATCTCCGTCCTTCCCCTGCCACAGGATAAAGC

oc 718 AGGAGAACCCTAGTAACTGCACCATCTCCCAGTCCGNGGATCTCCACCTGGGNCGAGGAN
ol 581 AGGAGAACCCCAGTAACTGCACCATCTCCCAGACCGTGGATCTCCACCTAAGCCGGGGAG
```

Figure 12. Alignment of *klf4* cDNAs between *O. celebensis* and *O. latipes*.

```
oc   115  TGGCTGCTACAA-CAACTCGCACCGGGACCGCGAGCTGCGCTTCTACACATTTCCAAAGG
ol   360  TGGCTGCTACAA-CAACTCGCACCGGGACCGGGAGCTGCGCTTCTACACATTTCCAAAGG

oc   174  ACACCACGCTAAGAGAGCAGTGGCTGCGGAATATTTCCCGGGCCGGGGTCAGCGGCTGCT
ol   419  ACACCACGCTAAGAGAGCAGTGGCTGAGGAATATTTCCCGGGCCGGGGTCAGCGGCTGCT

oc   234  TCAGCACCTTCCAGCCCACCACAGGTCACCGAGTCTGCAGTGTGCATTTCGCCGGAGGGA
ol   479  TCAGCACCTTTCAGCCCACCACAGGACACCGAGTCTGCAGTGTGCATTTCGCCGGAGGGC

oc   294  GGAAAACCTACTCGGTTCGAGTGCCGTCCCTGTTCCCCCTGCGCGGGGTAAATGAACGCA
ol   539  GGAAGACCTACTCCGTTCGAGTGCCGTCCCTGTTTCCCCTGCGTGGAGTGAATGAACGCA

oc   354  GGTGTCGGCGGGGCAGAGGGAGAAAGGTTTCCGCTGCAGTGCTCGCCCCCGCCNNAAAAN
ol   599  GGTGTCGGCGGGGCAGAGGGAGAAAGGTTTACGCTGCAGTCCTCGCCCCCGCCGCCGCAG

oc   414  NGACCTCTGGGATAGTGGCCACTAGCGTGCTTGGTAACACGGCTGGAGGAGCCGAGGGCA
ol   659  CGACCTCTGGGATTGTGGCCACTAGCGTGCTTGGGAACACGGCCGGAGGAGCCGAGGGCA

oc   474  GCATCGGAGGCGAGGCCGACGACGACAGCATTACCGTGGTTCAAATAGGCCAGAACGGGG
ol   719  GCATTGGAGGCGAGGCGGGCGACGACAGCATTACCGTGGTTCAAATAGGCCAGAACGGGG

oc   534  AGTACCTGGGCACCGCGCGGCTACCCGCTCAGTCAGAGGCGACCTGTTACGCCGCGCCAA
ol   779  AGTACCTGGGCACCGCGCGGCTACCGGCTCAGTCAGAGGCGACTTGTTACTCCGCGCCAT

oc   594  TCAGCATCGAGGAGGAACTCAAAGGCAACGAGTTC------GAATCCGCGACTAACCAGC
ol   839  GCGGCAGCGCGGATGAACTGAGAAGCGACGAGTACTCGTTGGAAACCGGCACTAACCAGC

oc   648  CCGCCGTGCAGTACGTTAGCGTGACCAGCAGCCCCTTGGACCACTCTTACTCCCTGACGA
ol   899  CCACAGTGCATTACGTGAGTGTGACCAGCAGCCCTCTGGACCACTCGTACTCCCTGACGA

oc   708  CTGGAACCACATCAGCCGAACTCCTNNNGNAGCTGAACGAGCAGCGGGA
ol   959  CTGGAACCACGTCAGCCGAACTACTGCGGAAACTGAACGAGCAGCGGGA
```

Figure 13. Alignment of *ronin* cDNAs between *O. celebensis* and *O. latipes*.

```
oc   121 NNTGGNGGGGATGACTTGGGGGCCAACGATGAAATGATCTCCTTCAAAGACGAAGGGGAG
ol    19 GGCGGCGGGGATGACTTAGGGGCCAACGATGAAATGATCCCCTTCAAAGACGAAGGGGAG

oc   181 CAGGAGGAAAAGATCTCCGAAAACGTTTCCGCGGAGAGGGACTTGGATGACGTGAAGTCG
ol    79 CAGGAGGAAAAGATATCCGAAAACGTTTCCGCGGAGAGGGACTTGGATGATGTGAAGTCG

oc   241 TCTCTCGTGAACGAGTCGGAAAACAACAGCAGTTCGTCCGACTCCGAGCAAGCAGAGAGG
ol   139 TCCCTCGTGAACGAGTCGGAAAACAACAGCAGCTCGTCTGACTCCGAGCACGCAGAGAGG

oc   301 GGCCCCCAAACCAGACCAGACTCAGAAAGTTATGAGAAAACAAGAGAGTACTTCAGTGAA
ol   199 CGCCCCCAAACCAGACCAGACTCAGAAAGTTATGAGAAAACAAGAGAGTACTTCAGTGAA

oc   361 GCCTTGAGGAGACAGCAAGATGGTGGCTTTTTTAAGAGTCCCCACTACCCTGGCTACCCG
ol   259 GCGCTGAGGAGACAGCAAGATGGTGGCTTCTTTAAGAGTCCCCACTACCCTGGCTATCCG

oc   421 TTCCTCATGATTNNNNNNNTNNNNNATCCCTACCTATCCAACGGCTCACTGTCTCCTGGT
ol   319 TTCCTCATGATCCCAGACCTCACCAATCCATACCTATCCAACGGCTCACTGTCTCCCGGT

oc   481 GCAAGAACATATCTTCACATGAAGTGGCCTCTCCTGGATGTCCCAGGCTCTGCAGGAATT
ol   379 GCAAGAACATATCTGCATATGAAGTGGCCCCTCCTGGATGTTCCCAGCTCTGCAGGAATA

oc   541 AAAGACTCCCGTTCTCCAACACCAGGACATTTATCCAATAAAGTCCCTGTGGTACAGCAT
ol   439 AAAGACTCCCGTTCTCCAACACCAGGACATTTATCCAATAAAGTCCCTGTGGTACAGCAT

oc   601 GCCCACCATGTGCACCCACTGACACCTCTCATCACCTACAGCAGTAATGAACACTTTTCC
ol   499 GGCCACCATGTGCACCCGCTGACACCTCTCATCACCTACAGCAACAATGAACACTTTTCC

oc   661 CCCGGCACGCCGCCATCGCACCTCTCGCCAGAGATCCTTGACCCAAAGACAGGTATCCCT
ol   559 CCCGGCACGCCGCCTTCACATCTCTCACCAGAGATCCTCGACCCAAAGACAGGTATCCCT

oc   721 CGGACGCCTCACCCATCAGAGCTTTCTCCATACTACCCCCCTGTCTCCTGGNGCTGTCGG
ol   619 CGGACGCCTCACCCATCAGAGCTTTCTCCATACTACCCCC-TGTCTCCCGGTGCTGTCGG
```

Figure 14. Alignment of *tcf3* cDNAs between *O. celebensis* and *O. latipes*.

```
oc   62  TTTCAAGAACAGACCGGTTACTGAAGCACAAACGGACTTGTGGAGAAGCCATAAAGAAGG
ol  900  TT-CAAGAACAGACCGGTTACTGAAGCACAAACGGACTTGTGGAGAAGCCATAAAGAAGG

oc  122  GTCTAGACCCAAACATGCTGGAGCTCAGTGAGGCGGAGCTTGGCCAAGGCAGCTATTCAG
ol  959  GTCTAGACCCAAACATGCTGGAGCTTAGTGAAGCGGAGCTTGGTCAAGGCAGCTATTCAG

oc  182  TCACTCAGGGAAACACCAGCAGCTCTGGACGCAAGAGGGGGAAGTCTAAAAACAGCGAGG
ol 1019  TCACTCAGGGAAACGCCAGCAGCTCTGGACGCAAGAGGGGGAAGTCCAAAAACAGCGAGG

oc  242  GCGGTGAGCGCAAGAGGAAGAAGAATGCCTCGGCAACAGCAGCAGCGTCTTGTTCTGAGG
ol 1079  GCGGTGAACGCAAGAGGAAGAAGAATGCCT------CCGCAGCAGCGTCTTCTTCTGAGG

oc  302  GGATGGCCCGTGAGCTGGGCCTGCATGACTTCAACATGGAGCACCCCTCTGGCTCTGATT
ol 1133  GGATGGCCCGTGAGCTGGGCCTGCATGACTTCAACATGGAGCACCCCTCCGGCTCTGATT

oc  362  CTGCAATGCAGGGGCGTACCCCCAAATTGGTATTTAAGAAATCTGCCCGGAAAGGACTTG
ol 1193  CTGCAATGCAGGGGCGTACCCCCAAATTGGTATTTAAAAAATCGGCCCGGAAAGGGCTTG

oc  422  ACAAAGNCCTCCTGTCCCTGGACGATGGTGCTGATGAACAAAAGCTGTTAGACCAGAAGT
ol 1253  ACAAAGGCCTCCTGTCCCTGGACGATGGTGCTGATGAACAAAAACTGTTAGACCAGAAAT

oc  482  CCGTCTCCATGGATCACGTGGAAGCTTCTGGCCTTGACAGCATGGGTCTTCTCCAGGGAC
ol 1313  CCGTCTCCATGGATCACGTGGAAGCTTCTGGCCTTGATAGCATGGGTCTTCTCCAGGGAC

oc  542  CTGGGGCCAACAAACCTGGGCCCACCACCAGCAGCAACTACGATGATGCAATGCAGTTTG
ol 1373  CTGGGGCCAACAAACCTGGGCCCACCACCAGCAGCAACTACGATGATGCAATGCAGTTTG

oc  602  TCAAAAAGCGGCGCTACCTCCATGCAGTTAATAATGACTATGGAGCTGGATCCCTGCACA
ol 1433  TCAAAAAGCGGCGCTACCTCCATGCAGTTAATAATGACTATGGAGCTGGATCCCTGCACA

oc  662  TGGCATCCCAGGGTAACAGTGTGATCCAGGGTTCCCTTGGCCGGAGCCCACACTGGCCA
ol 1493  TGGCATCCCAGGGTAACAGCGTGATCCAGGGTTCCCTTGGCCGGAGCCCACGCTGGCCA
```

Figure 15. Alignment of *zfp281* cDNAs between *O. celebensis* and *O. latipes*.

3.3.3 RT-PCR analysis using species-specific primer

To establish embryonic stages at which transcription activation occurs in medaka, F1 embryos were obtained by crossing *O. celebensis* male with *O. latipes* female. By performing RT-PCR using species specific primer, I detected the paternal transcripts of the six pluripotency genes (Figure 16). Generally, zygotic expression of all these pluripotency genes does not start until early blastula, reinforcing that in medaka zygotic expression commences around midblastula stage. As can be seen in the zygotic expression pattern of these genes, obviously six genes fall into two groups, namely Group A and Group B, representing two patterns. In the first pattern including *oct4* and *nanog* (Group A), paternal expression could be detected from early blastula. Intensity of transcription increased at later stage. In the second pattern (Group B), all other five genes showed zygotic expression at late blastula stage, later than group A genes.

On the other hand, another set of F1 embryos were obtained by crossing *O. latipes* male with *O. celebensis* female. Only the paternal transcripts of Group A genes, *oct4* and *nanog*, were examined in those embryos. Similar results were obtained. The zygotic expression of both *oct4* and *nanog* started at early blastula (Figure 17).

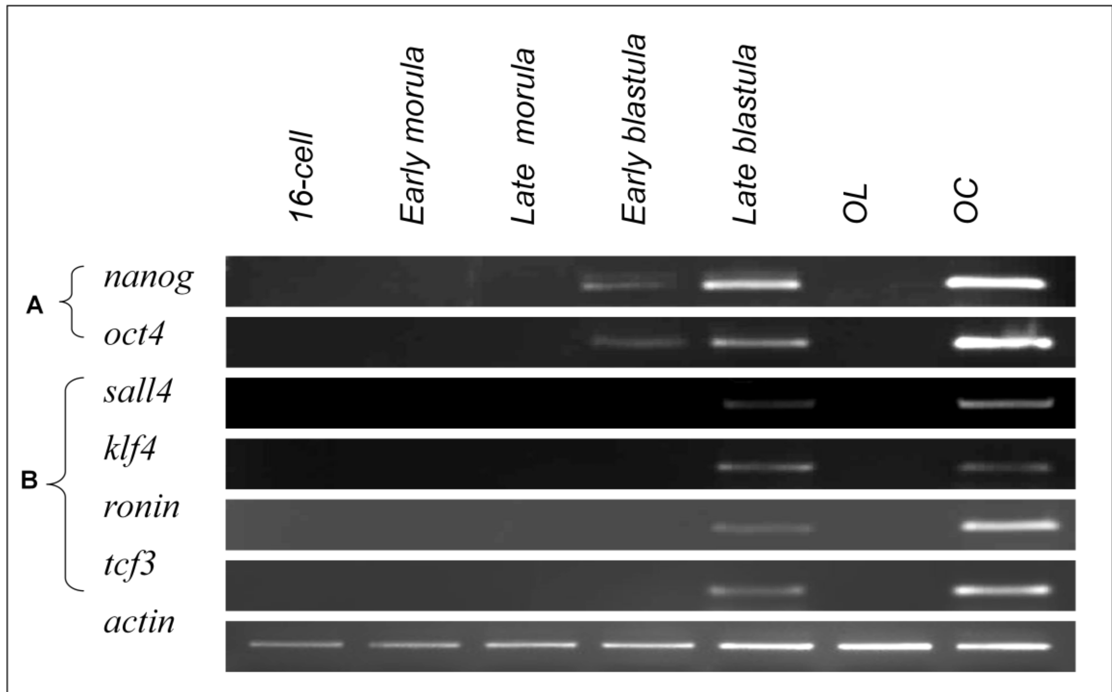

Figure 16. RT-PCR analysis of zygotic RNA expression in embryos (*O. celebensis* male X *O. latipes* female). OL, *O. latipes*; OC, *O. celebensis*

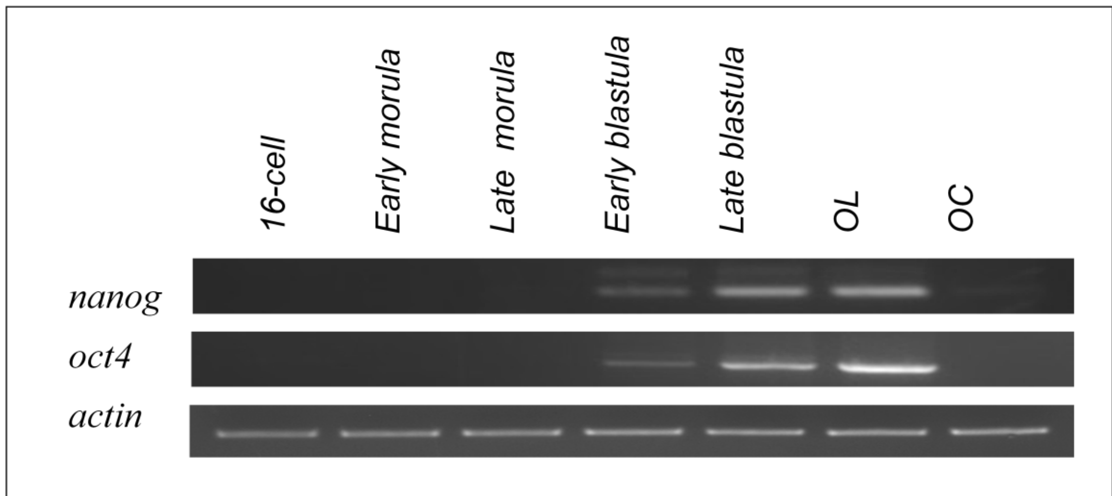

Figure 17. RT-PCR analysis of zygotic RNA expression in embryos (*O. latipes* male X *O. celebensis* female). OL, *O. latipes*; OC, *O. celebensis*

CHAPTER 4: DISCUSSION

4.1 Gene identification

In this study I identified seven medaka homologs/orthologs of mammalian pluripotency genes and examined their expression pattern. It has generally been accepted that fish has more genes than mammals because of one additional genome duplication event in the common fish lineage after its separation ~450 million years ago from the tetrapod lineage leading to birds and mammals (Amores et al., 1998). Duplicated fish genes may have different functions compared to their single copy gene counterparts in tetrapod vertebrates: one gene remains as homolog/ortholog, whereas the other may have got lost, or become co-homolog or co-ortholog (by loss of the chromosome synteny and original function). Many fish genes are said to be mammalian homologs on the basis of sequence comparisons and expression patterns. Among these genes in this study, medaka *nanog* and *oct4* have been reported to be mammalian homologs (Sánchez Sánchez et al., 2010). In this study, a phylogenetic sequence comparison and a conserved syntenic relationship along with the expression pattern suggest that the medaka *tcf3* identified in this study is homologous to the mammalian *tcf3*. However, in the absence of syntenic data, it cannot be determined whether the other medaka genes are homologous or orthologous to the mammalian counterparts.

According to the RT-PCR result in adult tissues and embryos, some of these

medaka genes show similar RNA expression patterns to those of their mouse counterparts, suggesting their homology in expression pattern.

The gene *oct4* is distinguished by exclusive expression in blastomeres, pluripotent early embryo cells, and the germ cell lineage (Nichols et al., 1998). In mouse blastocyst, *oct4* mRNA and protein are present in the ICM but not in the trophoblast (Palmieri et al., 1994). In vitro Oct4 is found only in undifferentiated embryonal carcinoma (EC), embryonic stem (ES), and embryonic germ(EG) cells (Nichols et al., 1998). According to the PCR analysis of *oct4* in medaka, *oct4* showed gonad-specific and early embryo expression pattern which is similar to that of mouse *oct4*.

In mice, trace amounts of *sall4* mRNA were detected in several somatic tissues such as the brain, heart and muscle (Tsubooka et al., 2009a), while in medaka, the transcripts of *sall4* were detected weakly in somatic tissues. In both mice and medaka, *sall4* is expressed highly in the testis and ovary, showing gonad-preferential expression. Meanwhile, *sall4* has been found to be expressed during early developments in mouse (Yang et al., 2010). Such expression pattern is similar to that in medaka early embryos.

In addition, the expression of *tcf3* in medaka is detectable in five somatic tissues. Such expression pattern corresponds to the finding that the *tcf3* expression is also not limted to stem cells in mouse (Cole et al., 2008a), because Tcf3 is a terminal component of the Wnt pathway which plays multiple functions such as body axis specification and morphogenic signaling.

On the other hand, difference in the expression pattern of these genes between medaka and mouse also exists.

The mRNA of *nanog* is present in pluripotent mouse and human cell lines, and absent in differentiated cells (Chambers et al., 2003). In mouse, *nanog* is not maternally deposited. Instead its expression commences at the morula stage and later on occurs in the ICM, epiblast, PGCs and ES cell cultures, albeit its expression is absent in spermatogonia of the adult testis(Cavaleri and Scholer, 2003; Chambers et al., 2003). This expression pattern is quite different from the results in this study, in which *nanog* is maternally supplied and *nanog* signal is detectable in spermatogonia. So it deserves to note that medaka *nanog* possesses a salient difference in expression *in vivo* from the mouse *nanog*.

Second, in this study, it was observed that the medaka *klf4* has maternal inheritance and early embryonic expression, while in zebrafish expression of *klf4* in early embryos has also been reported (Luo D, 2011). In mouse, however, *klf4* lacks expression in early embryos until the blastocyst stage, leading to a notion that *klf5* might substitute *klf4* for blastocyst development (Ema et al., 2008). Work is needed to determine whether this salient difference in embryonic *klf4* expression is due to the species difference or independent divergence between the fish and tetrapod lineages.

Furthermore, *ronin* in mouse was abundantly expressed in two adult tissues: ovary and some areas of the brain (Dejosez et al., 2008a). In medaka, *ronin*

shows wide expression pattern in all the adult tissue chosen in this study. During early embryonic development, *ronin* was found to first appear at the 2-cell stage, intensify during the 8-cell and compact morula stages (Dejosez et al., 2008a), suggesting that *ronin*, like *nanog*, is also not maternally supplied. However, according to the RT-PCR result, *ronin* is maternally deposited in early embryos of medaka.

4.2 Identification of pluripotency markers

4.2.1 *nanog* and *oct4*

In the study of mouse, the gene *nanog* and *oct4* are the best studied pluripotency genes in ES cell culture. Oct4 and Nanog are thought to be central to the transcriptional regulatory hierarchy that specifies ES cell identity because of their unique expression patterns and their essential roles during early development (Boyer et al., 2005a).

The PT-PCR analysis shows that the expression patterns of *nanog* and *oct4* in adult tissues and embryos are quite similar. Both of them are categorized in Group I. According to their RNA expression patterns in adult tissues, *nanog* and *oct4* show gonad-specific expression. Then the gene *nanog* was chosen as a representative of group-I genes for *in situ* hybridization in the ovary and testis. From the results in both ovary and testis, *nanog* expression is high in germ stem cells, and absent in somatic cells. Therefore, *nanog* expression is a pluripotency marker in the adult gonads. As for the weak signal of *nanog* in small oocyte, it can be explained by the fact that the oocyte is an exception to

differentiated cell, because it expresses pluripotency genes such as *oct4* at a high level for maternal supply to early embryogenesis(Pesce et al., 1998a).

The RT-PCR analysis of expression in embryos provides more information. The RNAs of *nanog* and *oct4* are maternally supplied, as can been seen from the RT-PCR result of the detection of both gene in 16-cell, morula, early blastula. The zygotic transcription occurs in MBT. After MBT, the expression of both *nanog* and *oct4* sharply declines and ultimately becomes undetectable. The midblastula stage in lower vertebrates such as fish is a critical stage for pluripotency. At this stage, the embryo is thought to consist of developmentally indeterminate cells allowing for ES cell derivation (Hong et al., 1996; Yi et al., 2009). After the blastula stage, cell lineage restriction and differentiation occurs during and after gastrulation. The expression patterns suggest that *nanog* and *oct4* appear to be expressed mainly in pluripotent cells of developing embryos.

I also chose 16-cell staged embryo to perform *in situ* hybridization to detect the *in vivo* expression of *nanog* and *oct4*. Both *nanog* and *oct4* were found to be present in blastomeres of the 16-cell staged embryo (Figure 5A-C), in accordance with a high level of their RNAs by RT-PCR analysis (Figure 4). This result is supported by the findings of two other studies. First, the *in vivo* RNA expression pattern of medaka *oct4* is similar to what was reported (Sánchez Sánchez et al., 2010). Meanwhile, the protein expression of medaka *nanog* detected by immunostaining (Camp et al., 2009) conforms in general to

its *in vivo* RNA expression pattern observed in this study. Since the blastomeres at this stage are thought to be totipotent cells in which zygotic transcription has not yet commenced, it is possible that both *nanog* and *oct4* are expressed in totipotent cells.

Furthermore, in mouse, *oct4* and *nanog* were reported to be down-regulated upon ES cells differentiation (Boyer et al., 2005a; Heng and Ng, 2010). In this study, RT-PCR analysis and quantitative PCR analysis both provide more convincing evidence that the expression level of *nanog* and *oct4* in both ES cell line HX1 and MES1 reduced dramatically after differentiation. Once again, *nanog* and *oct4* appear to be markers for pluripotency in undifferentiated ES cell cultures.

To sum up, it is revealed in this study that medaka *nanog* and *oct4* expression delineates with pluripotency in early embryos, adult gonads and ES cell culture, an expression pattern similar to that of the mouse *oct4*, which shows expression throughout the totipotent cycle and ES cells (Pesce et al., 1998a).

4.2.2 Other genes

On the other hand, the expression of other five genes, namely *klf4, ronin, sall4, tcf3* and *zfp281* occurs in one or more somatic tissues besides the adult gonads and persists beyond gastrulation. These genes *in vitro* closely resemble *nanog* and *oct4*, in that they exhibit a high level of expression in undifferentiated ES cells and dramatic down-regulation upon ES cell differentiation. Similar results have been reported in the studies in mouse.

Expression of *klf4, ronin, sall4, tcf3* and *zfp281* is strongly positive in undifferentiated ES cells and less so in differentiated ES cells (Chan et al., 2009; Dejosez et al., 2008a; Tsubooka et al., 2009b; Wang et al., 2008b; Yi et al., 2008). Accordingly, these five genes appear to be also pluripotency markers and perhaps even regulators for fish ES cell culture. Obviously, in medaka, the lack of pluripotency-specific expression *in vitro* is not exclusive for a gene in question as a pluripotency marker *in vivo*. A similar situation is also seen in mouse, where *klf4* exhibits wide expression but is a integral component of the four factors capable of reprogramming differentiated cells into iPS cells (Takahashi and Yamanaka, 2006).

Six of the seven medaka genes have previously been found to be expressed highly in undifferentiated cultures of two haploid ES cell lines (HX1 and HX2) and one diploid MES1 line, and two of them, *nanog* and *oct4* exhibited dramatic down-regulation upon haploid ES cell differentiation (Yi et al., 2009). The present study corroborates and extends the previous work by demonstrating the same pluripotency-specific expression pattern in ES cell culture for a set of seven genes including *tcf3*, a new member identified.

4.2.3 Somatic expression of pluripotency genes

By identifying candidate pluripotency genes through the homology approach and by examining their expression *in vivo* and *in vitro* in this study, I got two important findings in medaka. The first is that pluripotency-specific expression *in vivo* is strongly indicative of pluripotency-specific expression also *in vitro*, as

is the case for *nanog* and *oct4*. The other is that pluripotency genes do not necessarily show pluripotency-specific expression *in vivo*, as is the case for the other five genes.

According to the RT-PCR analysis of these five genes, which shows that they are expressed in one or more somatic tissues besides the adult gonads and beyond the gastrulation in embryo development, expression of these genes is not restricted to pluripotency *in vivo* but extended for other processes.

The finding in the *in situ* hybridization of *tcf3* in the testis provides example for such statement. The wide *tcf3* expression in various stages of spermatogenesis suggests that *tcf3* expression in the testis is not limited to stem cells but may be important also for additional functions, such as spermatogenesis progression. This finding in medaka can be supported by other studies in zebrafish wherein both *tcf3* genes are required for brain patterning (Dorsky et al., 2003; Kim et al., 2000),

Besides *tcf3*, other genes can also be found to play function other than pluripotency. *Klfs* family are important in blood vessel development in addition to hematopoiesis and epidermal development (Kawahara and Dawid, 2000; Suzuki et al., 2005). Klf4 is a pathologically induced factor in endothelial cells (ECs) as well as in vascular smooth muscle cells (SMCs) to regulate vascular cell function (Suzuki et al., 2005). In addition, the other functions played by *sall4* are implied by the *sall4* mutation disorder, known as Okihiro syndrome, which is characterized by limb deformity, eye movement deficits and, less

commonly, anorectal, ear, heart, cranial midline and kidney anomalies (Tsubooka et al., 2009b). Therefore, many pluripotency genes, such as those of groups II and III, appear to be pleiotropic *in vivo*.

In this regard, the candidacy for pluripotency genes cannot be determined directly according to their *in vivo* expression pattern. On the other hand, only test in ES cell culture showed obvious difference in expression level of these genes before and after ES cell differentiation, providing direct evidence that these genes are pluripotency markers in undifferentiated ES cell culture. Taken together, this study demonstrates that established ES cell culture is a unique *in vitro* system to ultimately test putative pluripotency genes for practical use in monitoring the pluripotency of putative ES cells and other stem cell cultures.

4.2.4 Conserved expression of pluripotency genes in vertebrates

According to the protein alignment result of those genes, protein products of the five genes in expression groups II and III showed substantially high sequence conservation while Nanog and Oct4 both exhibited considerable divergence in protein sequence. In this study, it is shown that like their mammalian counterparts, all the seven medaka genes identified by homology search are molecular markers for undifferentiated ES cell culture. These data demonstrate that pluripotency genes, in spite of sequence divergence, have conserved their pluripotency-specific expression *in vitro* from mammals to lower vertebrates. This notion is supported by previous demonstration that the mouse *oct4* promoter is active in medaka to drive transgene expression of

reporters in early embryos and of drug selectable markers in undifferentiated ES cells (Hong et al., 2004b).

4.3 Zygotic expression pattern

4.3.1 Zygotic expression at midblastula

As can be seen from the RT-PCR analysis in embryo, those seven genes in this study are all maternally supplied. The activation of zygotic expression of those genes cannot be monitored.

In medaka, by crossing two parental inbred strains, Aizawa made use of polymorphism between two strains, to examine the zygotic transcription. For gene whose activation begins by the 2 cell stage in mouse embryo, such as *eif-4c*, the zygotic expression of its medaka homolog is activated after stage 11, (Aizawa et al., 2003). According to the manifestation of onset of zygotic gene expression by Azaiwa, there is no evidence of zygotic expression of EST markers before stage 10 (Aizawa et al., 2003). It was concluded that medaka MBT in terms of first time transcription of paternal genes begins at stage 11 (Aizawa et al., 2003).

In this study, the same method was carried out between two available species, *O. latipes* and *O. celebensis*. The zygotic expression of several pluripotency candidate genes was shown. According to the result, the onset of zygotic expression of the gene *nanog* and *oct4* is earlier than stage 11 (late blastula), at early blastula stage (stage 10). Meanwhile, it is shown that the other five genes are activated at late blastula stage, corresponding to the conclusion in

Aizawa's study.

The up-regulation of the expression of genes at the MBT has been regarded as a sign of the potential developmental roles for certain genes (O'Boyle et al., 2007). Based on this hypothesis, several important genes which play critical roles in embryonic patterning and general cellular metabolism were isolated. For example, the activation of *apaf1*, which is responsible for cell death, and of *hsp27*, which can prevent apoptosis, is key to the balance of apoptosis with proliferation. Such balance is crucial in the early embryo for normal development to ensue (O'Boyle et al., 2007). Furthermore, in zebrafish, the gene *klf4*, which was also found to be expressed for the first time after the MBT, regulates cell growth, proliferation, differentiation and embryogenesis (Dang et al., 2000). *Klf4* also contributes to blood and vascular development and was shown to have a critical role in erythroid cell differentiation in zebrafish (Kawahara and Dawid, 2000). In this study, the zygotic expressions of six pluripotency genes in this study are activated around MBT, suggesting that these genes will play critical roles in later events of embryo development, such as gastrulation and segmentation (Bree et al., 2005). Such notion corresponded to their roles in maintaining pluripotency and self-renew in ES cells.

Those six genes fall into two groups according to the zygotic expression pattern. The group I genes *nanog* and *oct4,* which shows pluripotency-specific expression, commence their zygotic expression earlier than group II and III

genes, which are expressed in other tissues besides gonad and are also detectable beyond gastulation. Since the maternal transcript will be exhausted at the onset of paternal expression (Aizawa et al., 2003), combining the RT-PCR results of both embryos and hybrid embryos, the zygotic expression patterns of those genes are inferred (Table 6).

Table 6 Summary of zygotic expression in medaka embryos

Gene	16-cell	EM	LM	EB	LB	PMG	34som	Pre
nanog	-	-	-	+	++	++	-	-
oct4	-	-	-	+	++	++	-	-
sall4	-	-	-	-	+	++	++	++
klf4	-	-	-	-	+	++	++	++
ronin	-	-	-	-	+	+++	++	++
tcf3	-	-	-	-	+	++	++	++

Relative levels of expression are indicated by -, +, ++ and +++ for barely detectable, faintly detectable, easily detectable and high expression. EM, early morula; LM, late morula; EB, early blastula; LB, late blastula; PMG, pre-mid gastrula; 34som, 34 somite; Pre, prehatch.

4.3.2 *O. celebensis* in building the zygotic expression examining model

Oryzias celebensis is a medaka related species which inhabits in a tropical zone (Hirayama et al., 2006). In the study of molecular phylogeny of the medaka fishes genus oryzias, Takehana divided the oryzias species into three major species group, namely the latipes, javanicus, and celebensis groups (Takehana et al., 2005).

Previously, the study concerning the zygotic expression of gene has focused on two inbred strains within *O. latipes* (Aizawa et al., 2003). The basis of the

study was the polymorphism between different strains. However, if the genes do not have obvious polymorphism, the proposed experiment cannot be done. *O. celebensis* is different species from *O. latipes*, therefore the sequences of certain genes show obvious sequence differences between those two species. Such hypothesis was confirmed by the sequence obtained in this study.

O. celebensis is not well studied. Only several studies involving *O. celebensis* were reported, which focused on taxonomic. Although there is no enough sequence information about *O. celebensis*, I can still get the useful data through cloning certain genes. According to the sequence analysis of those genes between *O. latipes* and *O. celebensis*, the sequence differences in certain region of the gene sequence makes it possible to design species-specific primer.

The method in this study to examine zygotic expression of certain genes is also applicable in other studies.

4.3.3 Hierarchical expression among pluripotency genes

ES cells, possess tightly controlled transcriptional regulatory networks that maintain the cells in their self-renewing and pluripotent state and poise them for differentiation. By DNA microarray and Chromatin Immunoprecipitation assay, the transcriptional regulatory network in embryonic stem cells has been dissected in mammals (Heng and Ng, 2010). However, in medaka it has not been done. By the method in this study, the hierarchical expression pattern which inferred the relationship between those pluripotency genes was

obtained （Table 7）.

Table 7. The hierarchy of zygotic activation

Hierarchy	Gene
Tier I	*nanog, oct4*
Tier II	*sall4, klf4, ronin, tcf3*

According to the hierarchy among the pluripotency genes, several facts can be related. First, Oct4 and Nanog, along with sox2 form the transcriptional core that preserves ES cells self-renewal and pluripotency, regulates and supports the pluripotent framework of ES cells by employing regulatory mechanisms such as feedforward and autoregulatory loops (Boyer et al., 2005a; Heng and Ng, 2010). Particularly, Nanog and Oct4 have been suggested to be important for mediating a pluripotent “ground state” in ES cells (Heng and Ng, 2010). Furthermore, Oct4 and Nanog were identified to occupy a large number of target genes (Boyer et al., 2005a). Their downstream targets are also important to keep ES cells from differentiating (Loh et al., 2006a). Taken together, the emerging picture is that Oct4 and Nanog control a cascade of pathways that are intricately connected to govern pluripotency. Thus, the higher hierarchy of zygotic expression of Oct4 and Nanog corresponds to their core character within the transcriptional regulatory network in maintaining pluripotency of ES cells.

In addition, it was stated that the expression of genes that are activated at the MBT may promote, or be a result of, the dramatic changes in gene expression

that occur with the activation of the zygotic genome (O'Boyle et al., 2007). Therefore, the hierarchical expression of those genes in this study provides clue for their relativity. Since the zygotic expression of *nanog* and *oct4* is activated earlier, it is possible that once the decrease and exhaustion of maternal transcript occurs, *nanog* and *oct4* are responsible for the zygotic activation of other pluripotency genes, including the genes in this study. However, those transcription factors do not necessary act on target gene independently but work in concert with other factors. The intricate interplay and partnership of transcription factors established a complex transcriptional network (Heng and Ng, 2010). Anyhow, given the hypothesis based on the result in this study, more researches involving the networking of pluripotency genes are needed.

CHAPTER 5: CONCLUSION AND FUTURE WORK

5.1 Conclusion

In this study, I identified seven pluripotency candidate genes by using homology approach. By examining their expression in vivo and in vitro in this study, those genes were identified as pluripotency markers. This indicates the applicability of the homology approach to identify pluripotency genes. The seven pluripotency genes and the approach for identifying and testing them as pluripotency markers in vitro provide a general tool for testing stem cell derivation in other fish species and even other lower vertebrates.

In addition, a good model was built to examine the timing of zygotic expression of the pluripotency genes in developmental embryos. By using this method, I examined the paternal transcript of six pluripotency genes. Their activation at midblastula transition suggested important roles in embryogenesis. Meanwhile, the model to study the zygotic expression of pluripotency genes can be applied to the study of other genes, such as germ cell markers namely *vasa*, *boule*, *dazl* and *dnd*.

5.2 Future work

Following the investigations described in this thesis, a number of projects could be taken up:

- Further investigation and identification of more potential genes that can be used as medaka pluripotency markers, such as *sox2* and *lin28* which are used in generation of iPS cells.
- The detailed roles of the seven genes in maintaining pluripotency are not studied. Therefore, studies of gene function are needed, such as gene knockout and knockdown.
- According to the hierarchical expression of the seven genes, it would be interesting to examine their interaction between each other.

CHAPTER 6: REFERENCES

Aizawa, K., Shimada, A., Naruse, K., Mitani, H., Shima, A., 2003. The medaka midblastula transition as revealed by the expression of the paternal genome. Gene expression patterns. 3**,** 43-47.

Amores, A., Force, A., Yan, Y. L., Joly, L., Amemiya, C., Fritz, A., Ho, R. K., Langeland, J., Prince, V., Wang, Y. L., Westerfield, M., Ekker, M., Postlethwait, J. H., 1998. Zebrafish hox clusters and vertebrate genome evolution. Science. 282**,** 1711-4.

Araki, K., Okamoto, H., Graveson, A. C., Nakayama, I., Nagoya, H., 2001. Analysis of haploid development based on expression patterns of developmental genes in the medaka Oryzias latipes. Dev Growth Differ. 43**,** 591-9.

Botquin, V., Hess, H., Fuhrmann, G., Anastassiadis, C., Gross, M. K., Vriend, G., Schöler, H. R., 1998. New POU dimer configuration mediates antagonistic control of an osteopontin preimplantation enhancer by Oct-4 and Sox-2. Genes & development. 12**,** 2073.

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., 2005a. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell. 122**,** 947-956.

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., Gifford, D. K., Melton, D. A., Jaenisch, R., Young, R. A., 2005b. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell. 122**,** 947-56.

Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J., Li, Y., Xu, C., Fang, R., Guegler, K., 2004. Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation. Nature biotechnology. 22**,** 707-716.

Bree, R. T., McLoughlin, S., Jin, S. W., McMeel, O. M., Stainier, D. Y. R., Grealy, M., Byrnes, L., 2005. Nanor, a novel zygotic gene, is expressed initially at the midblastula transition in zebrafish. Biochemical and biophysical research communications. 333**,** 722-728.

Camp, E., Sanchez-Sanchez, A. V., Garcia-Espana, A., Desalle, R., Odqvist, L., Enrique O'Connor, J., Mullor, J. L., 2009. Nanog regulates proliferation during early fish development. Stem Cells. 27**,** 2081-91.

Cavaleri, F., Scholer, H. R., 2003. Nanog: a new recruit to the embryonic stem cell orchestra. Cell. 113**,** 551-2.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., Smith, A., 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell. 113**,** 643-655.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., Smith, A., 2007. Nanog safeguards pluripotency and mediates germline development. Nature. 450**,** 1230-1234.

Chan, K. K. K., Zhang, J., Chia, N. Y., Chan, Y. S., Sim, H. S., Tan, K. S., Oh, S. K. W., Ng, H. H., Choo, A. B. H., 2009. KLF4 and PBX1 directly regulate NANOG expression in human embryonic stem cells. Stem cells. 27**,** 2114-2125.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V. B., Wong, E., Orlov, Y. L., Zhang, W., Jiang, J., 2008. Integration of external signaling pathways with the core transcriptional network in embryonic

stem cells. Cell. 133, 1106-1117.

Cole, M. F., Johnstone, S. E., Newman, J. J., Kagey, M. H., Young, R. A., 2008a. Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes Dev. 22, 746-55.

Cole, M. F., Johnstone, S. E., Newman, J. J., Kagey, M. H., Young, R. A., 2008b. Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes & development. 22, 746.

Dang, D. T., Pevsner, J., Yang, V. W., 2000. The biology of the mammalian Krüppel-like family of transcription factors. The international journal of biochemistry & cell biology. 32, 1103-1121.

Dejosez, M., Krumenacker, J. S., Zitur, L. J., Passeri, M., Chu, L. F., Songyang, Z., Thomson, J. A., Zwaka, T. P., 2008a. Ronin is essential for embryogenesis and the pluripotency of mouse embryonic stem cells. Cell. 133, 1162-1174.

Dejosez, M., Krumenacker, J. S., Zitur, L. J., Passeri, M., Chu, L. F., Songyang, Z., Thomson, J. A., Zwaka, T. P., 2008b. Ronin is essential for embryogenesis and the pluripotency of mouse embryonic stem cells. Cell. 133, 1162-74.

Dorsky, R. I., Itoh, M., Moon, R. T., Chitnis, A., 2003. Two tcf3 genes cooperate to pattern the zebrafish brain. Development. 130, 1937-47.

Ema, M., Mori, D., Niwa, H., Hasegawa, Y., Yamanaka, Y., Hitoshi, S., Mimura, J., Kawabe, Y., Hosoya, T., Morita, M., Shimosato, D., Uchida, K., Suzuki, N., Yanagisawa, J., Sogawa, K., Rossant, J., Yamamoto, M., Takahashi, S., Fujii-Kuriyama, Y., 2008. Kruppel-like factor 5 is essential for blastocyst development and the normal self-renewal of mouse ESCs. Cell Stem Cell. 3, 555-67.

Evans, M., Kaufman, M., 1981. Establishment in culture of pluripotential cells from mouse embryos. Nature. 292, 154-156.

Evans, P. M., Zhang, W., Chen, X., Yang, J., Bhakat, K. K., Liu, C., 2007. Krüppel-like factor 4 is acetylated by p300 and regulates gene transcription via modulation of histone acetylation. Journal of Biological Chemistry. 282, 33994.

Fujita, J., Crane, A. M., Souza, M. K., Dejosez, M., Kyba, M., Flavell, R. A., Thomson, J. A., Zwaka, T. P., 2008. Caspase activity mediates the differentiation of embryonic stem cells. Cell stem cell. 2, 595-601.

Guan, K., Nayernia, K., Maier, L. S., Wagner, S., Dressel, R., Lee, J. H., Nolte, J., Wolf, F., Li, M., Engel, W., Hasenfuss, G., 2006. Pluripotency of spermatogonial stem cells from adult mouse testis. Nature. 440, 1199-203.

Guo, G., Yang, J., Nichols, J., Hall, J. S., Eyres, I., Mansfield, W., Smith, A., 2009. Klf4 reverts developmentally programmed restriction of ground state pluripotency. Development. 136, 1063.

Hall, J., Guo, G., Wray, J., Eyres, I., Nichols, J., Grotewold, L., Morfopoulou, S., Humphreys, P., Mansfield, W., Walker, R., 2009. Oct4 and LIF/Stat3 additively induce Krüppel factors to sustain embryonic stem cell self-renewal. Cell stem cell. 5, 597-609.

Hamatani, T., Carter, M. G., Sharov, A. A., Ko, M. S. H., 2004. Dynamics of global gene expression changes during mouse preimplantation development. Developmental Cell. 6, 117-131.

Heng, J. C. D., Ng, H. H., 2010. Transcriptional Regulation in Embryonic Stem Cells. Advances in experimental medicine and biology. 695, 76-91.

Hirayama, M., Mitani, H., Watabe, S., 2006. Temperature-dependent growth rates and gene expression patterns of various medaka Oryzias latipes cell lines derived from different

populations. Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology. 176, 311-320.

Hong, N., Li, M., Zeng, Z., Yi, M., Deng, J., Gui, J., Winkler, C., Schartl, M., Hong, Y., 2010. Accessibility of host cell lineages to medaka stem cells depends on genetic background and irradiation of recipient embryos. Cell Mol Life Sci.

Hong N, L. Z., Hong Y, 2011. Fish stem cell cultures. I*nternational Journal of Biological Sciences*.

Hong, Y., Liu, T., Zhao, H., Xu, H., Wang, W., Liu, R., Chen, T., Deng, J., Gui, J., 2004a. Establishment of a normal medakafish spermatogonial cell line capable of sperm production in vitro. Proc Natl Acad Sci U S A. 101, 8011-6.

Hong, Y., Winkler, C., Liu, T., Chai, G., Schartl, M., 2004b. Activation of the mouse Oct4 promoter in medaka embryonic stem cells and its use for ablation of spontaneous differentiation. Mech Dev. 121, 933-43.

Hong, Y., Winkler, C., Schartl, M., 1996. Pluripotency and differentiation of embryonic stem cell lines from the medakafish (Oryzias latipes). Mech Dev. 60, 33-44.

Hong, Y., Winkler, C., Schartl, M., 1998. Production of medakafish chimeras from a stable embryonic stem cell line. Proc Natl Acad Sci U S A. 95, 3679-84.

Iwamatsu, T., 2004. Stages of normal development in the medaka Oryzias latipes. Mech Dev. 121, 605-18.

Iwamatsu, T., Ohta, T., Oshima, E., N, S., 1988. Oogeneis in the Medaka *Oryzias latipes* -- Stage of Oocyte Development. Zoological science. 5, 353-373.

Jiang, J., Chan, Y. S., Loh, Y. H., Cai, J., Tong, G. Q., Lim, C. A., Robson, P., Zhong, S., Ng, H. H., 2008. A core Klf circuitry regulates self-renewal of embryonic stem cells. Nature cell biology. 10, 353-360.

Kawahara, A., Dawid, I. B., 2000. Expression of the Krüppel-like zinc finger gene biklf during zebrafish development. Mechanisms of Development. 97, 173-176.

Kim, C. H., Oda, T., Itoh, M., Jiang, D., Artinger, K. B., Chandrasekharappa, S. C., Driever, W., Chitnis, A. B., 2000. Repressor activity of Headless/Tcf3 is essential for vertebrate head formation. Nature. 407, 913-6.

Ko, M., Kitchen, J. R., Wang, X., Threat, T. A., Hasegawa, A., Sun, T., Grahovac, M., Kargul, G., Lim, M., Cui, Y., 2000. Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development. Development. 127, 1737.

Li, M., Hong, N., Xu, H., Yi, M., Li, C., Gui, J., Hong, Y., 2009. Medaka vasa is required for migration but not survival of primordial germ cells. Mech Dev. 126, 366-81.

Li, Z., Bhat, N., Manali, D., Wang, D., Hong, N., Yi, M., Ge, R., Hong, Y., 2011. Medaka cleavage embryos are capable of generating ES-like cell cultures. Int J Biol Sci. 7, 418-25.

Lim, C. Y., Tam, W. L., Zhang, J., Ang, H. S., Jia, H., Lipovich, L., Ng, H. H., Wei, C. L., Sung, W. K., Robson, P., Yang, H., Lim, B., 2008. Sall4 regulates distinct transcription circuitries in different blastocyst-derived stem cell lineages. Cell Stem Cell. 3, 543-54.

Liu, L., Hong, N., Xu, H., Li, M., Yan, Y., Purwanti, Y., Yi, M., Li, Z., Wang, L., Hong, Y., 2009. Medaka dead end encodes a cytoplasmic protein and identifies embryonic and adult germ cells. Gene Expr Patterns.

Loh, Y. H., Wu, Q., Chew, J. L., Vega, V. B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., 2006a. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nature genetics. 38, 431-440.

Loh, Y. H., Wu, Q., Chew, J. L., Vega, V. B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., Wong, K. Y., Sung, K. W., Lee, C. W., Zhao, X. D., Chiu, K. P., Lipovich, L., Kuznetsov, V. A., Robson, P., Stanton, L. W., Wei, C. L., Ruan, Y., Lim, B., Ng, H. H., 2006b. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet. 38, 431-40.

Luo D, H. W., Chen S, Zhu Z, 2011. Critical developmental stages for the efficiency of somatic cell nuclear transfer In zebrafish International Journal of Biological Sciences.

Ma, C., Fan, L., Ganassin, R., Bols, N., Collodi, P., 2001a. Production of zebrafish germ-line chimeras from embryo cell cultures. Proc Natl Acad Sci U S A. 98, 2461-6.

Ma, J., Svoboda, P., Schultz, R. M., Stein, P., 2001b. Regulation of zygotic gene activation in the preimplantation mouse embryo: global activation and repression of gene expression. Biology of reproduction. 64, 1713.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A. A., 2007. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nature cell biology. 9, 625-635.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., Yamanaka, S., 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell. 113, 631-642.

Munoz, M., Trigal, B., Molina, I., Diez, C., Caamano, J. N., Gomez, E., 2009. Constraints to progress in embryonic stem cells from domestic species. Stem Cell Rev. 5, 6-9.

Nakatake, Y., Fukui, N., Iwamatsu, Y., Masui, S., Takahashi, K., Yagi, R., Yagi, K., Miyazaki, J., Matoba, R., Ko, M. S. H., 2006. Klf4 cooperates with Oct3/4 and Sox2 to activate the Lefty1 core promoter in embryonic stem cells. Molecular and cellular biology. 26, 7772.

Newport, J., Kirschner, M., 1982a. A major developmental transition in early Xenopus embryos: I. characterization and timing of cellular changes at the midblastula stage. Cell. 30, 675-686.

Newport, J., Kirschner, M., 1982b. A major developmental transition in early Xenopus embryos: II. Control of the onset of transcription. Cell. 30, 687-696.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., Smith, A., 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell. 95, 379-391.

Niwa, H., Miyazaki, J., Smith, A. G., 2000a. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nature genetics. 24, 372-376.

Niwa, H., Miyazaki, J., Smith, A. G., 2000b. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet. 24, 372-6.

O'Boyle, S., Bree, R. T., McLoughlin, S., Grealy, M., Byrnes, L., 2007. Identification of zygotic genes expressed at the midblastula transition in zebrafish. Biochemical and biophysical research communications. 358, 462-468.

Palmieri, S. L., Peter, W., Hess, H., Scholer, H. R., 1994. Oct-4 transcription factor is differentially expressed in the mouse embryo during establishment of the first two extraembryonic cell lineages involved in implantation. Developmental biology. 166, 259-267.

Pereira, L., Yi, F., Merrill, B. J., 2006. Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. Molecular and cellular biology. 26, 7479.

Pesce, M., Gross, M. K., Scholer, H. R., 1998a. In line with our ancestors: Oct-4 and the mammalian germ. Bioessays. 20, 722-32.

Pesce, M., Schöler, H. R., 2001. Oct 4: gatekeeper in the beginnings of mammalian development. Stem cells. 19, 271-278.

Pesce, M., Wang, X., Wolgemuth, D. J., Scholer, H., 1998b. Differential expression of the Oct-4 transcription factor during mouse germ cell differentiation. Mech Dev. 71, 89-98.

Saijoh, Y., Fujii, H., Meno, C., Sato, M., Hirota, Y., Nagamatsu, S., Ikeda, M., Hamada, H., 1996. Identification of putative downstream genes of Oct-3, a pluripotent cell-specific transcription factor. Genes to cells: devoted to molecular & cellular mechanisms. 1, 239.

Sánchez Sánchez, A. V., Camp, E., García España, A., Leal Tassias, A., Mullor, J. L., 2010. Medaka Oct4 is expressed during early embryo development, and in primordial germ cells and adult gonads. Developmental Dynamics. 239, 672-679.

Sridharan, R., Tchieu, J., Mason, M. J., Yachechko, R., 2009. Role of the murine reprogramming factors in the induction of pluripotency. Cell. 136, 364-377.

Suzuki, T., Aizawa, K., Matsumura, T., Nagai, R., 2005. Vascular implications of the Kruppel-like family of transcription factors. Arteriosclerosis, thrombosis, and vascular biology. 25, 1135.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., Yamanaka, S., 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell. 131, 861-872.

Takahashi, K., Yamanaka, S., 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 126, 663-676.

Takehana, Y., Naruse, K., Sakaizumi, M., 2005. Molecular phylogeny of the medaka fishes genus Oryzias (Beloniformes: Adrianichthyidae) based on nuclear and mitochondrial DNA sequences. Molecular phylogenetics and evolution. 36, 417-428.

Tam, W. L., Lim, C. Y., Han, J., Zhang, J., Ang, Y. S., Ng, H. H., Yang, H., Lim, B., 2008a. T-cell factor 3 regulates embryonic stem cell pluripotency and self-renewal by the transcriptional control of multiple lineage pathways. Stem Cells. 26, 2019-31.

Tam, W. L., Lim, C. Y., Han, J., Zhang, J., Ang, Y. S., Ng, H. H., Yang, H., Lim, B., 2008b. T Cell Factor 3 Regulates Embryonic Stem Cell Pluripotency and Self Renewal by the Transcriptional Control of Multiple Lineage Pathways. Stem cells. 26, 2019-2031.

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., Jones, J. M., 1998. Embryonic stem cell lines derived from human blastocysts. Science. 282, 1145-7.

Tsubooka, N., Ichisaka, T., Okita, K., Takahashi, K., Nakagawa, M., Yamanaka, S., 2009a. Roles of Sall4 in the generation of pluripotent stem cells from blastocysts and fibroblasts. Genes Cells. 14, 683-94.

Tsubooka, N., Ichisaka, T., Okita, K., Takahashi, K., Nakagawa, M., Yamanaka, S., 2009b. Roles of Sall4 in the generation of pluripotent stem cells from blastocysts and fibroblasts. Genes to Cells. 14, 683-694.

Verfaillie, C., 2009. Pluripotent stem cells. Transfusion clinique et biologique. 16, 65-69.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D. N., Theunissen, T. W., Orkin, S. H., 2006. A protein interaction network for pluripotency of embryonic stem cells. Nature. 444, 364-368.

Wang, Z. X., Teh, C. H., Chan, C. M., Chu, C., Rossbach, M., Kunarso, G., Allapitchay, T. B., Wong, K. Y., Stanton, L. W., 2008a. The transcription factor Zfp281 controls embryonic stem cell pluripotency by direct activation and repression of target genes. Stem Cells. 26, 2791-9.

Wang, Z. X., Teh, C. H. L., Chan, C. M. Y., Chu, C., Rossbach, M., Kunarso, G., Allapitchay, T. B., Wong, K. Y., Stanton, L. W., 2008b. The transcription factor Zfp281 controls embryonic stem cell

pluripotency by direct activation and repression of target genes. Stem cells. 26, 2791-2799.

Wei, Z., Yang, Y., Zhang, P., Andrianakos, R., Hasegawa, K., Lyu, J., Chen, X., Bai, G., Liu, C., Pera, M., 2009. Klf4 interacts directly with Oct4 and Sox2 to promote reprogramming. Stem cells. 27, 2969-2978.

Wittbrodt, J., Shima, A., Schartl, M., 2002. Medaka--a model organism from the far East. Nat Rev Genet. 3, 53-64.

Wobus, A. M., Boheler, K. R., 2005. Embryonic stem cells: prospects for developmental biology and cell therapy. Physiol Rev. 85, 635-78.

Xu, H., Li, Z., Li, M., Wang, L., Hong, Y., 2009. Boule is present in fish and bisexually expressed in adult and embryonic germ cells of medaka. PLoS One. 4, e6097.

Yang, J., Chai, L., Fowles, T. C., Alipio, Z., Xu, D., Fink, L. M., Ward, D. C., Ma, Y., 2008. Genome-wide analysis reveals Sall4 to be a major regulator of pluripotency in murine-embryonic stem cells. Proceedings of the National Academy of Sciences. 105, 19756.

Yang, J., Gao, C., Chai, L., Ma, Y., Bridger, J. M., 2010. A novel SALL4/OCT4 transcriptional feedback network for pluripotency of embryonic stem cells. PloS one. 5, e10766.

Yao, Y. Q., Xu, J. S., Lee, W. M., Yeung, W. S. B., Lee, K. F., 2003. Identification of mRNAs that are up-regulated after fertilization in the murine zygote by suppression subtractive hybridization. Biochemical and biophysical research communications. 304, 60-66.

Yi, F., Pereira, L., Merrill, B. J., 2008. Tcf3 Functions as a Steady State Limiter of Transcriptional Programs of Mouse Embryonic Stem Cell Self Renewal. Stem cells. 26, 1951-1960.

Yi, M., Hong, N., Hong, Y., 2009. Generation of medaka fish haploid embryonic stem cells. Science. 326, 430-3.

Yi, M., Hong, N., Hong, Y., 2010a. Derivation and characterization of haploid embryonic stem cell cultures in medaka fish. Nat Protoc. 5, 1418-30.

Yi, M., Hong, N., Li, Z., Yan, Y., Wang, D., Zhao, H., Hong, Y., 2010b. Medaka fish stem cells and their applications. Sci China Life Sci. 53, 426-34.

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R., 2007. Induced pluripotent stem cell lines derived from human somatic cells. Science. 318, 1917.

Yuri, S., Fujimura, S., Nimura, K., Takeda, N., Toyooka, Y., Fujimura, Y. I., Aburatani, H., Ura, K., Koseki, H., Niwa, H., 2009. Sall4 is essential for stabilization, but not for pluripotency, of embryonic stem cells by repressing aberrant trophectoderm gene expression. Stem cells. 27, 796-805.

Zhang, J., Tam, W. L., Tong, G. Q., Wu, Q., Chan, H. Y., Soh, B. S., Lou, Y., Yang, J., Ma, Y., Chai, L., 2006. Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. Nature cell biology. 8, 1114-1123.