NORTHWESTERN UNIVERSITY

Reprogramming Pluripotent Stem Cell Towards Totipotency

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Northwestern University Interdepartmental Neuroscience Program

By

Sanders Oh

EVANSTON, ILLINOIS

December 2017

Abstract

Totipotent cells have the highest developmental potential and can only be created by nuclear transfer into oocytes. Identities of maternal factors that can induce this reprogramming remain a mystery. In this report, we demonstrate induction of totipotency on mouse embryonic stem cells by introducing six factors, Hist1h2aa, H3f3b, H1foo, p-Npm2, Zscan4d, and Ubtfl1. We observed dose-dependent increases in the MuERV-L endogenous retrovirus expression, typically seen in totipotent 2-cell stage blastomere, and adding p150 siRNA and trichostatin A further increased the expression. These cells, which we designated iTLCs (induced totipotent-like cells), had upregulation of totipotent genes but downregulation of pluripotent and differentiation genes, suggesting a distinct shift towards the totipotent state. Furthermore, iTLCs displayed unusually large nuclei, a characteristic of zygotic genome activation (ZGA). Also, iTLCs showed telomere lengthening and were able to be cultured in totipotent condition. Meanwhile, iTLCs did not show malignant transformations as indicated by normal karyotypes, inability to grow in nutrient-deprived condition, and sensitivity to contact inhibition. iTLCs demonstrated expanded cell fate potential by differentiating into all three distinct lineages of the pre-implantation embryo and expressed markers for both embryonic and extraembryonic lineages. RNAseq data showed remarkable similarities between iTLCs and totipotent cells. Early ZGA genes were strongly upregulated in iTLCs, indicating active totipotent state. When reprogrammed with factors only for an extended period, we observed cells resembling various stages of embryogenesis. These data suggest that pluripotent stem cells can be reprogrammed toward totipotent state without the need of oocytes and raise the tantalizing possibility of creating totipotent cells.

This is more than a dissertation

This is a journey

To reach what appears to be
Unreachable

To achieve what seems
Impossible

To create what many call a

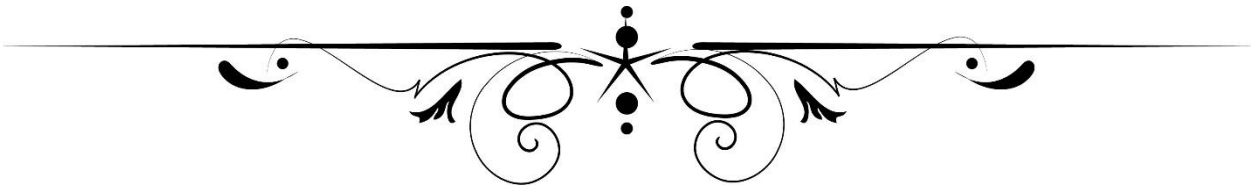
Miracle

Milestones

I.

The Beginning

"As scientists, we step on the shoulders of science, building on the work that has come before us – aiming to inspire a new generation of young scientists to continue once we are gone." – Stephen Hawking



The beginning of this journey serendipitously started with the observation of a dead cell under a microscope by Robert Hooke in 1665.[1,2] About a decade later, in 1674, a live cell was observed by Anton van Leeuwenhoek.[2-4] These discoveries opened a new world; Brown described cell nucleus in 1833, Roelliker showed that sperm and egg are also cells in 1840, and Pringsheim witnessed how a sperm penetrated an egg in 1856.[5-7] Many conclusions about a cell have been made during this time—some correct and some wildly wrong—but one that remains to be true and perhaps most influential came from Rudolf Virchow: *omnis cellula e cellula*, all cells arise from pre-existing cells.[7,8]

The implication of Virchow's conclusion meant that cells could be used to renew all cell types. This renewable source would not be discovered until 1981 when Gail Martin derived embryonic stem cells (ESCs) from the embryos of mice.[9] In 1998, Jeffrey Jones and his team derived the first ESC lines from

human blastocysts.[10] While the term “stem cell” was used much earlier, first appearing in scientific literature in 1868, many authors misused and applied the term incorrectly.[11-13] The meaning of the term stem cell is defined as having the capacity to both self-renew and give rise to differentiated cells.[14,15] There are many types of stem cells—embryonic, hematopoietic, adult, and others—with embryonic having higher cell potency.

Cell potency is a cell’s capacity to differentiate into other cell types. In short, the more different types a cell can differentiate into, the higher its potency. A cell with the highest potency is referred to as totipotent. Examples are zygotes and 2-cell stage embryos. The unique properties of totipotent cells are that they can give rise to both embryonic and extra-embryonic cells, in essence to a whole organism. Therefore all animal reproduction and cloning require achieving totipotency.[16,17] A more differentiated cells are referred to as pluripotent. A classic example of this is ESC, which can give rise to all three germ layers of an embryo but cannot give rise to extra-embryonic tissues.[18,19] Sometimes the term “stem cell” is used to mean pluripotency, however, this not always correct and can lead to confusion. A more differentiated from pluripotency is multipotent. A multipotent cell can give rise to all cell types of a given tissue within a germ layer.[20,21] An example is hematopoietic stem cell which can give rise to all cell types of blood tissue within the mesoderm. Further down the cell potency leads to oligopotency, sometimes called progenitor cells.[22] An example is a lymphoid cell, which can give rise to B and T cells but not to red blood cell. Next, unipotency, also called precursor cell, refers to a cell that can give rise to only one cell type.[23] An example of this is hepatoblast which can only differentiate into hepatocyte. Lastly, a cell with no potency is referred to as terminally differentiated cell.[24]

A cell with higher potency has strong predisposition to differentiate down its lineage spontaneously. This spontaneous differentiation can be blocked with exogenous inhibitors, for example, Leukemia Inhibitory Factor (LIF) for pluripotent stem cells.[25-29] This seemingly unidirectional path of cell differentiation was

once believed to be irreversible with germ cells possessing unique properties to reset this potency.[30-32] However, in 2006, Shinya Yamanaka's group was able to reverse this process by restoring pluripotency in terminally differentiated mouse fibroblasts.[33] These embryonic-like stem cells, called induced pluripotent stem cells (iPSCs), were generated by overexpressing four transcription factors—Oct3/4, Sox2, Klf4, and c-Myc—that play known important roles in establishing pluripotent pathway. Interestingly, the same homologous factors were able to reprogram human fibroblasts to iPSCs.[34] This meant that human pluripotent stem cells could be acquired without the ethical issues of destroying human embryos. Furthermore, iPSCs would be genetically identical to the host in which they were reprogrammed from, avoiding the concerns of immune rejection. Since then others have improved this reprogramming technique by eliminating the need to use c-Myc, a strong proto-oncogene, and delivering the genes without integrating into the genome, overcoming the safety concerns of iPSC therapy.[35-37]

Despite the improvements in generating iPSCs, problems remain that raise important safety concerns. DNA methylation signatures of the previous somatic cell of origin persist in iPSCs.[38-42] This epigenetic memory limits the differentiation potential and efficiency.[42-45] Furthermore, cells derived from iPSCs show premature senescence and apoptosis than the ones derived from ESCs.[46,47] Subpopulations of iPSCs show abnormal 3D genome architecture and folding patterns.[48] Also, iPSCs have higher mutational load than ESCs.[40,49-51] In contrast, when genetic mutations were exogenously introduced to early human preimplantation embryos, the subsequent embryogenesis process led to the promotion of normal cells and elimination by apoptosis of cells with mutations, resulting in blastocysts with normal development potential.[52] This "self-correcting" process is absent in iPSCs and may partly explain why ESCs derived from blastocysts display superior pluripotent stem cell qualities.

Not surprisingly, iPSC-based therapy, so far, has been unsuccessful. The first and only approved iPSC-based clinical trial which aimed at treating a degenerative eye condition, that Yamanaka's team was

involved in, had to be stopped due to unexpected mutations. In contrast, many ESC-based clinical trials are currently active for wide range of diseases, including the macular degeneration disease for which iPSC trial had attempted but failed.[53]

Despite the discovery of iPSCs, at present, ESCs remain indispensable and serve as the "gold standard" for pluripotent stem cells. Similarly, the results achieved from ESC-based clinical trials will likely set the standards for future iPSC-based therapies. Then, instead of generating patient-derived iPSCs, the paragon is in generating patient-derived ESCs which will be named induced totipotent cells (iTCs).

ESCs can only be obtained from blastocysts.[9,10] The only cells capable of generating blastocysts are totipotent cells. There are only two ways to generate totipotent cells; union of sperm and egg, also known as reproduction, or somatic cell nuclear transfer (SCNT), also known as cloning.[54-61] The one constant variable here is oocyte. Therefore, oocytes most likely hold the necessary key factors for establishing totipotency. Unfortunately, the paucity of the cells generated by the female creates difficulty, and, in the case of human, raises ethical concerns. Furthermore, ESCs derived from reproduction are not patient-derived therefore can only be therapeutically used in immunopriviledged areas such as the eye, and generation of ESCs from human SCNT has been extremely difficult for reasons unknown.[62,63] An ideal approach is to reprogram patient-derived cells toward totipotency using defined factors.

Unlike pluripotency, the molecular mechanism underlying totipotency remains mostly unknown. The volume of oocyte rapidly increases during oogenesis presumably due to accumulation of maternal factors such as Nucleoplasmin (Npm) 2 and subcortical maternal complex (Nlrp5, Tle6, Floped, Padi6, Filia) that are required for establishing totipotency.[64-66] Once the mature oocyte is fertilized by a sperm or activated by SCNT, the stored maternal factors reset epigenome of the parent DNA, in part, by active demethylation that is driven by Tet3 genome-wide oxidation of 5-methylcytosine to 5-hydroxymethylcytosine.[67-71] This

process has also been found in other mammalian cells and even present as DNA code in genomes of bacteriophages, indicating evolutionarily conserved importance.[72] In addition to epigenome reset, histones undergo extensive modifications, in particular, lysine acetylation and lysine trimethylation, driven by maternal Brg1 which is a component of the ATP-dependent chromatin remodeling SWI/SNF complex.[73,74] The net effect of histone modifications result in euchromatin structure. This loosely folded chromatin combined with hypomethylation of the genome allows for hyper-accessibility of transcription factors which ultimately result in whole-genome activation and totipotency. This sudden burst of transcription is known as zygotic genome activation (ZGA) which starts at the late one-cell stage and peaks at the two-cell stage in the mouse.[75] Interestingly, this period perfectly overlaps with totipotency. Perhaps then the underlying mechanism of totipotency may be found within the ZGA.

Unfortunately, even if key activators of totipotency are present within the ZGA, large numbers of genes that are expressed during this period cause too much background noise and reduce the probability of selecting the correct factors. Furthermore, the molecular mechanisms of ZGA are not well understood. During ZGA, transcripts that lack TATA promoters which are normally silenced are activated.[76] This includes repetitive elements such as endogenous retrotransposons.[77] A specific example is murine endogenous retrovirus with a leucine tRNA primer binding site (MERVL) at the two-cell stage which has been established as a marker for totipotent cells by a number of different groups.[78,79] These repeat elements may be expressed during this stage only as byproducts due to the hyper-permissive state of the genome. Alternatively, temporal overlap with the narrow window of totipotent state may indicate that some of these repeat elements may be involved in establishing totipotency.

Based on the current limited understanding of totipotency and its culture condition, successful reprogramming of iTCs require serendipitous chance. Therefore, we have made several choices that may increase the odds of success. First, mouse cell line was used instead of human. Most of the current

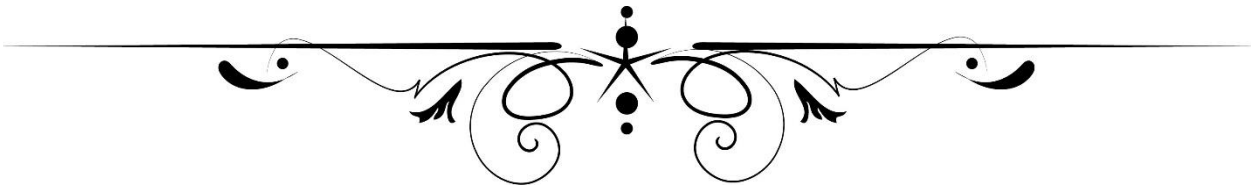
understanding about totipotency come from mouse studies. Also, generating human totipotent cells, even in the absence of human oocyte, may pose an ethical conundrum. Second, instead of somatic cells, pluripotent stem cells were used for reprogramming. Totipotency and pluripotency are closer in development, therefore, the barrier for successful reprogramming would be lower. Third, we did not limit ourselves to only transcription factors. Exogenous chemicals when relevant were used to enhance our chance towards the totipotency.

These choices do not lessen the potential application if successful. First, the same homologous four factors that generated mouse iPSCs were also able to generate human iPSCs, indicating that cell potency pathways are highly conserved across species.[33,34] Second, when combined with iPSC method, reprogramming pluripotency to totipotency would, by a two-step process, allow for induction of totipotency in somatic cells. Third, recent iPSC methods have shown that small inhibitors and chemicals can enhance the efficiency and replace pluripotent transcription factors.[80-82] Despite all these precautions, so much uncertainty continues to plague this journey. Much more is needed to arrive at totipotency. We had to build the most advanced vehicle for reprogramming technology; synthetic modified mRNA.

II.

Synthetic Modified mRNA

"Technological progress is like an axe in the hands of a pathological criminal." – Albert Einstein



Technological progress in cell reprogramming enabled the advancement of existing and the development of new vehicles in which to deliver factors into a cell. DNA-based viral vectors are the most commonly used method for expressing proteins in a cell. However, this approach has some significant drawbacks including the possibility of viral integration into the host genome, immunogenicity, and gene silencing. Non-viral transfection of plasmid DNA (pDNA) is associated with reduced immunogenicity and less chance of genome integration,[83,84] but it has much lower efficiency and reduced protein production.[36,85,86]

RNA-based protein expression provides several key advantages when compared to DNA-based approaches. First, RNA-based vectors do not enter the cell nucleus, eliminating potential complications due to integration into the host genome. Second, the strength of protein expression is independent of promoter activity and thus avoids the gene silencing that is common with DNA-based techniques for gene expression. Third, RNA-mediated manipulation of cellular phenotype may have a higher efficiency than

DNA-based methods.[37,87] Similar to DNA-based methods, RNA-based methods have both viral and non-viral approaches. RNA viruses such as Sendai virus function exclusively in the cytoplasm and are capable of producing large amounts of proteins.[88] However, they are more challenging to construct and produce, the cessation of gene expression is difficult to control accurately, and there are safety concerns related to the persistent virus.

Synthetic modified mRNA is a non-viral RNA-based technology that solves many safety concerns present in other gene expression techniques. Being both non-integrative and non-viral, mRNA technology allows for direct clinical applications without the need for laborious and time-consuming screening of integrated oncogenes or viral components.[87,89] In addition, unlike plasmid DNA-mediated expression, mRNA-mediated expression is effective for both mitotic and post-mitotic cells.[89] In contrast to retro- and lentiviral vectors which often drive multiple genes under a single promoter, each mRNA-mediated gene can be dosed independently from one another, resulting in more sophisticated stoichiometric expression of multiple genes.[90] Furthermore, unlike other inducible systems that may have leaky background expression, the short half-life of mRNA allows good temporal control of silencing of exogenous expression. Despite these advantages, concerns with stability and immunogenicity of earlier versions of synthetic mRNAs prevented widespread use.[91] The protocol presented here improves upon these shortcomings.

The stability and immunogenicity of mRNA can be changed through transcript engineering.[92-95] Five components make up the eukaryotic mature mRNA: 5’ cap, 5’ untranslated region (UTR), open reading frame (ORF), 3’ UTR, and 3’ poly adenosine (poly(A)) tail. Each component plays a vital role in stability and immunogenicity.

Standard *in vitro* transcription (IVT) of RNAs results in 5’ triphosphate which is vulnerable to exonuclease degradation.[96,97] In contrast, eukaryotic cells post-transcriptionally modify mRNA in the nucleus with a

methylated cap structure at the 5’ end. In addition to stabilizing the mRNA, the 5’ cap increases protein production through recognition by eukaryotic initiation factor 4E (eIF4E) which leads to ribosome recruitment and translation.[94,96,98,99] The 5’ cap is also involved in other mRNA processes such as cellular transport, splicing into substrate RNAs, and translational repression via microRNA.[100-102] The protocol in this paper acknowledges this importance and uses the RNA cap structure analog, m7G(5’)ppp(5’)G, which shares the uncommon 7-methylguanosine and 5’-5’ triphosphate bridge found in eukaryotic mRNA.

Eukaryotic UTR sequences have a variety of regulatory functions and influence the degree of stability of the mRNA transcript. For example, AU-rich sequences in the 3’ UTR induce degradation by promoting the removal of poly(A) tail while iron response elements present in both 5’ and 3’ UTR regulate stability and translation.[95,103,104] UTRs of the globin gene are known to be stable and therefore are used in this protocol.[105]

The poly(A) tail also plays an important role in enhancing mRNA stability and translation.[106] Most actively translated eukaryotic mRNAs have a poly(A) tail of 100-250 bases.[107] While exogenous protein synthesis can occur with a much shorter tail, stability and translation efficiency are positively correlated with length.[108,109] The poly(A) tail binds to polyadenosyl-binding proteins (PABP) which interact with the N-terminal region of eukaryotic initiation factor 4G (eIF4G) which then interacts with the 5’ cap and eIF4E.[110] This cap-eIF4E-eIF4G-PABP-poly(A) interaction creates a circularized closed loop mRNA structure, which may increase stability by protection from nucleolytic degradation and increase translation efficiency by ribosome reutilization.[111,112] Since the poly(A) tail plays an important role in stability and translation, the reproducible results require a consistent and standardized number of poly(A) bases. This protocol attaches the same length of 120 bases of poly(A) to the 3’ UTR of each mRNA vector.

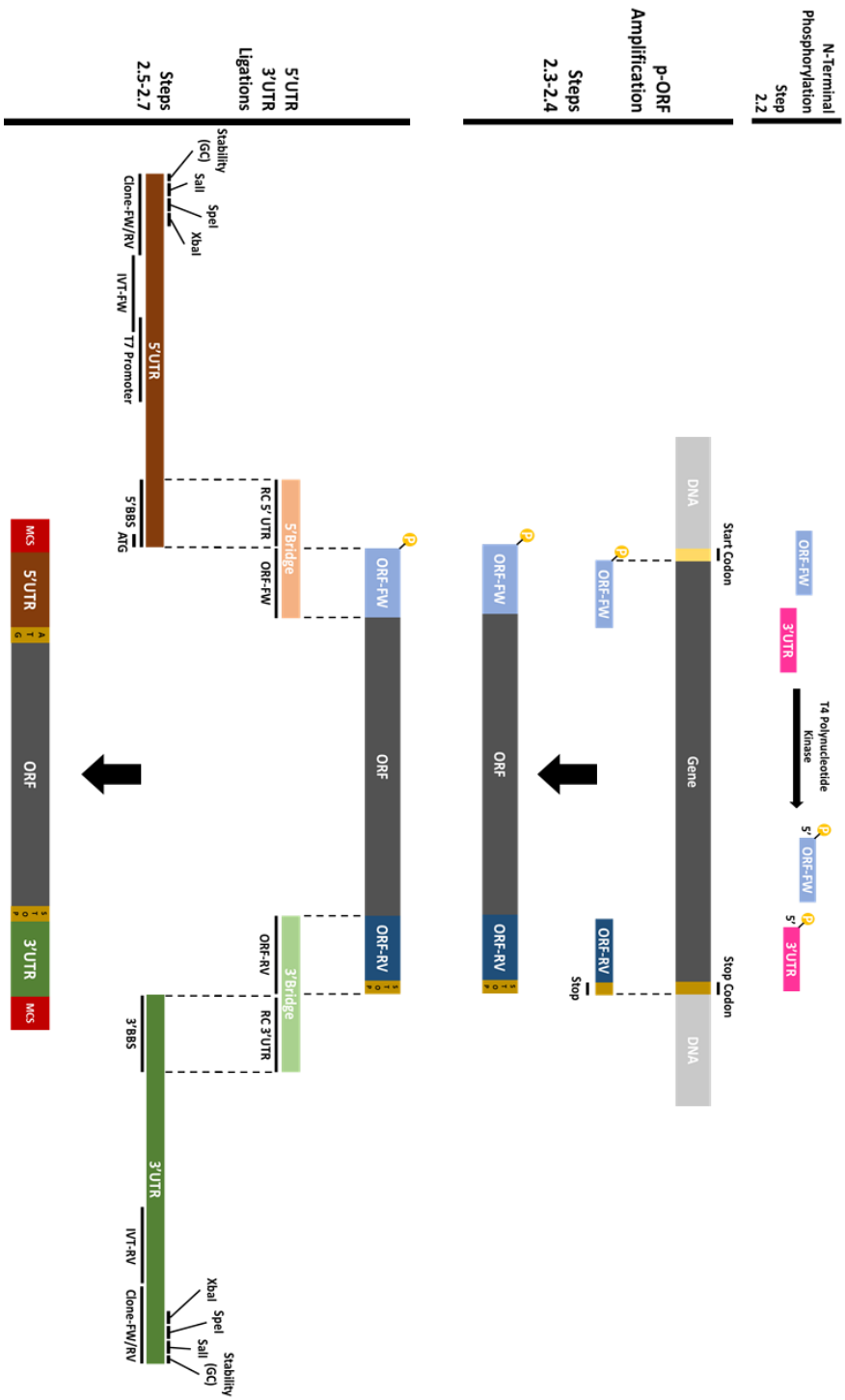
Immune responses also can be reduced by modifying the bases of mRNA.[37,89,113] Exposure to non-self

nucleic acid activates the mammalian innate immune system by activating Toll-like receptor (TLRs).[114] Different TLRs recognize different types of nucleic acids.[113,115-118] TLR3 in particular recognizes mRNA and induces a strong cascading immune response.[118] The use of modified nucleosides in place of naturally occurring ones in mammalian cells dramatically reduces immunogenicity.[37,89,113] Thus, this protocol substitutes cytidine and uridine with 5-methylcytidine and pseudouridine, respectively. Further reduction in immunogenicity is achieved by removing 5' triphosphates and adding interferon inhibitor B18R into the cell culture media.[37]

The use of synthetic modified mRNA for gene expression is very promising because of its superior safety profile and higher efficiency.[37,87] When compared to other methods, mRNA-mediated reprogramming of somatic cells to pluripotent stem cells resulted in 10-100 times more efficient and faster emergence of colonies while still maintaining the lowest aneuploidy of the treated cells.[87,119,120] Despite many advantages, mRNA stability and immunogenicity have been the most significant barriers to widespread use.[91] This protocol addresses these shortcomings. The protocol is meticulously detailed so that any user, experienced or inexperienced, can readily apply this technique.

The overall schematic of this protocol and its major steps are depicted in Figure 1. The ORF of the gene of interest is amplified with N-terminal phosphorylated forward primer, which provides the necessary energy for the subsequent 5'UTR ligation. The energy necessary for the 3'UTR ligation is provided by N-terminal phosphorylation of 3'UTR. The 5'UTR and 3'UTR oligonucleotides are ligated to ORF with the assistance of 5' and 3'Bridge, respectively. The resulting 5'UTR-ORF-3'UTR construct is cloned into a vector which then transformed into ultra-competent *E. coli*. After subsequent bacterial amplification, the extracted clones are linearized to avoid producing run-on transcripts during the IVT reaction. The linearized templates, along with 5' cap analog and modified base pairs, are combined to generate the synthetic modified mRNA using *In Vitro* Transcription Kit. The newly synthesized mRNAs are transfected

into the cells using one of three kits recommended in this protocol.



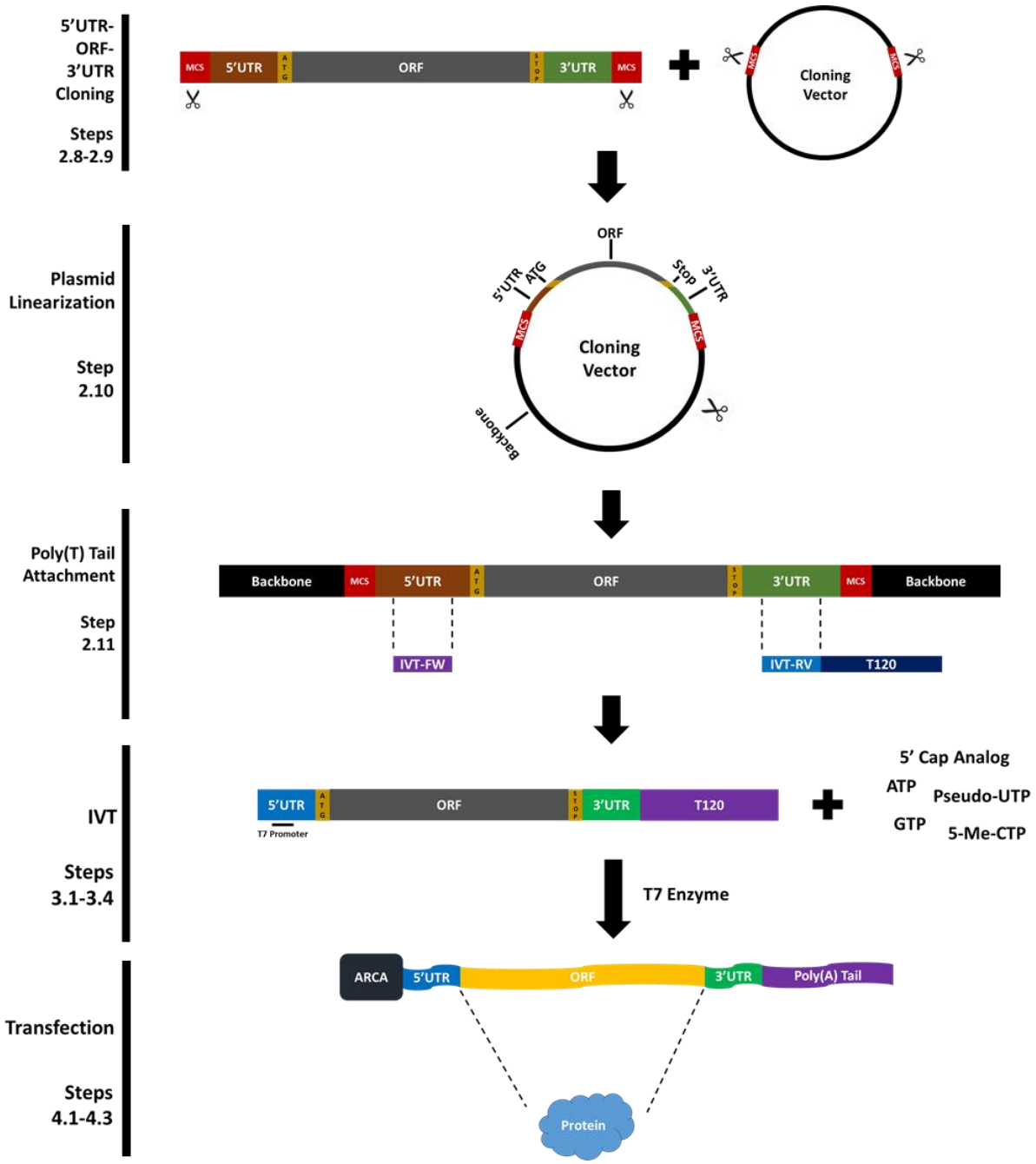


Figure 1: Schematic overview of the method. Major steps are depicted with corresponding Steps on the left-hand side. Gaps between major steps, as indicated by the breaks in vertical lines, indicate possible pause points. MCS refers to SalI, SpeI, and XbaI restriction sites.

The details of 5'UTR and 3'UTR regions are shown in Figure 2. Particular attention should be given to the

three available restriction sites that are flanking both UTRs. Ensure that the restriction enzymes chosen are appropriate for both the UTRs and the gene of interest. If none of these three restriction sites can be used for a particular gene, the user can replace with the appropriate restriction sequence when synthesizing the UTRs. Ensure that the GC stability is present when making changes to the restriction sites. Other regions to note are the actual UTR sequences. The UTRs shown here are derived from alpha globin gene of *Mus. musculus* which have shown to have high stability and expression. UTRs are common sites of transcript regulation. Therefore other UTRs may be more desirable depending on the need for particular experiment. In such situations, ensure that both BBS oligonucleotides are appropriately changed to match the new UTR sequences. Similarly, the length of poly(T)-tail, not shown in this figure, may be lengthened or shortened depending on the desired degree of stability. Lastly, if users choose to perform IVT using other than the T7 RNA polymerase, the noted T7 Promoter sequence in this figure needs to be changed appropriately.

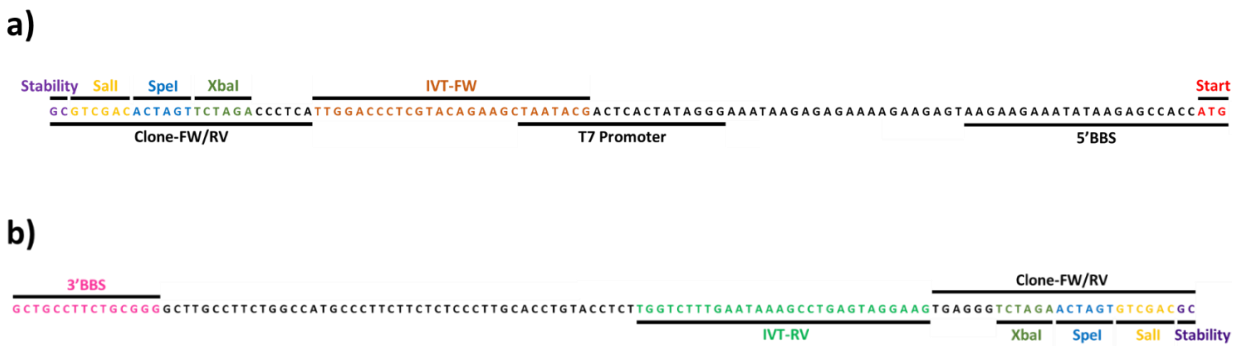


Figure 2: UTR maps and sequences. (A) 5'UTR and its restriction, binding, and promoter sites are shown. Note the ATG on the 3' end which obviates the need for an ORF-FW start codon. (B) 3'UTR and its restriction and binding sites are shown. Other UTR sequences are compatible with this protocol as long as the various corresponding sequences noted here also are changed appropriately.

Key molecules used for IVT are shown in Figure 3. The four RNA bases used in this protocol are shown in

Figure 3a. Other modified bases that have shown to be present in mammalian cells, therefore, may theoretically be used as a substitute are shown in Figure 3b. Depending on the target mammalian cell types, other modified bases may be more appropriate. Please note that the substitution of other modified bases needs to be validated. The red dashed circles note the modifications of each base pair in comparison to the generic ATCG. The Figure 3c shows the 5' anti-reverse cap analog (ARCA) used in this protocol and Figure 3d shows an interesting molecular modification called m^6A_m that has shown to increase the half-life of mRNAs.[121,122] Substituting ARCA with m^6A_m, therefore, may change the decay rates of mRNAs which may provide for an interesting approach to investigating temporal dependent events. Modifications to this protocol need to be carefully validated for protein expression.

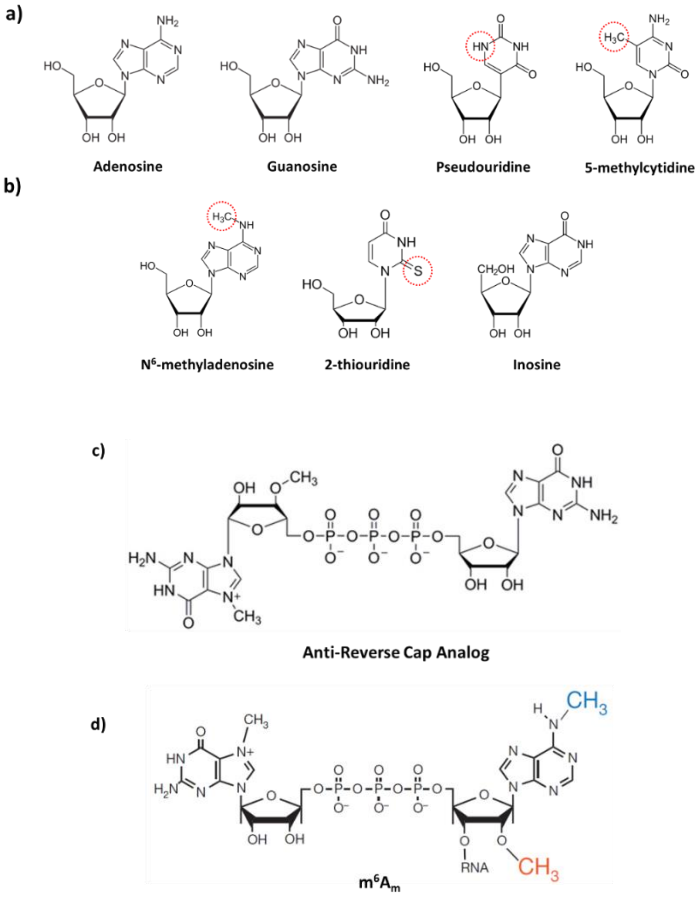


Figure 3: Nucleosides and 5’ caps of mRNA. (A) Canonical and modified base pairs used in this protocol. (B) Other nucleoside modifications for immune escape. Dashed red circles note base modifications. (C) The 5’ cap used in this protocol. (D) Alternative 5’ cap that has been shown to increase the half-life of eukaryotic mRNA.[121] Blue and orange indicate modifications.

Gel electrophoresis of the major steps is shown in Figure 4. Ten ORFs (A to J) of varying sizes, 384 to 3336 nucleotides, and functions—histone variants (Th2b, Th2a, H1foo), embryogenesis (Npm2, Zar1), maternal-effect factors (Filia, Nlrp5), arginine-modifying enzyme (Padi6), oogenesis (Nobox)—have been successfully tested. The initial PCR products for each phosphorylated ORFs are shown in Figure 4a. The UTR double ligation of each sample shows expected upshift in the gel electrophoresis as shown in Figure 4b. The double ligated samples are subsequently cloned into vectors, which are then extracted and purified. The clones are linearized as seen in Figure 4d. Complete linearization is important for successful downstream IVT reaction, therefore, uncut clones should be run in parallel for comparison as shown in Figure 4c. The final DNA portion of the protocol concludes with attaching poly(T)-tail to 5’-ORF-3’ template. As seen in Figure 4e, each sample shows the expected upshift—297 nucleotide increase to be exact—from their original phosphorylated ORFs (Figure 4a) which would likely indicate successful 5’UTR, 3’UTR, and Poly(T) attachments.

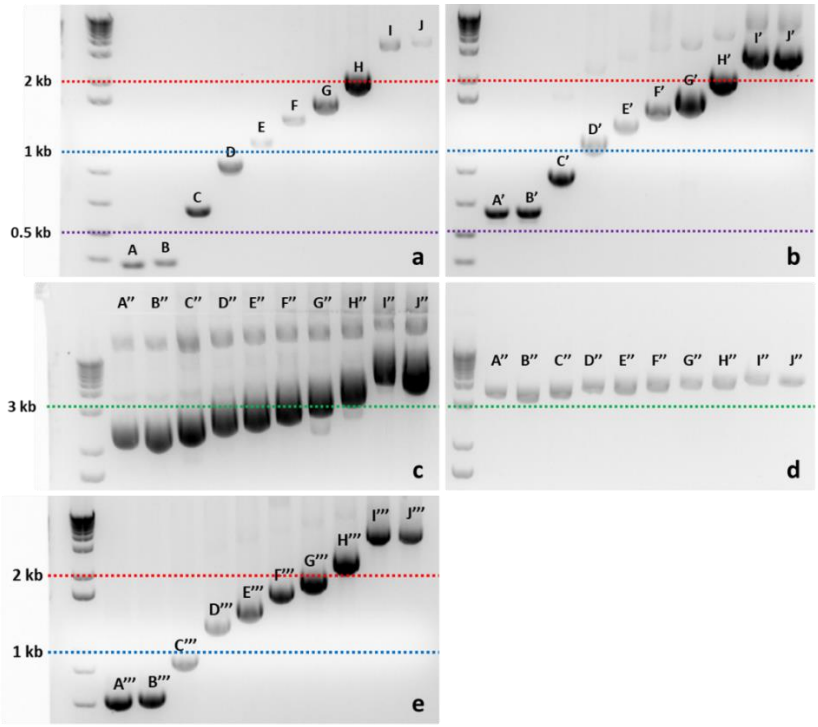


Figure 4: Gel electrophoresis of all intermediate and final DNA products. (A) The beginning PCR products of the N-terminal phosphorylated ORF. (B) Double ligation of 5' and 3'UTRs. (C) Undigested cloned vectors. (D) Linearized cloned vectors. (E) Poly(T)-tail attached DNA templates for the IVT reaction. Note the corresponding increase in size at each step. A = Th2b (384 bp), B = Th2a (390 bp), C = Npm2 (624 bp), D = H1foo (915 bp), E = Zar1 (1086 bp), F = Filia (1323 bp), G = Nobox (1584 bp), H = Padi6 (2049 bp), I = Nlrp5 isoform b (3288 bp), J = Nlrp5 isoform a (3336 bp). Th2b, Th2a, H1foo (histone variants), Npm2, Zar1 (embryogenesis), Filia, Nlrp5 (maternal-effect factors), Padi6 (arginine-modifying enzyme), Nobox (oogenesis).

Measure the concentration and purity of each major step of the protocol using a NanoDrop and a Bioanalyzer as shown in Figure 5. The poor result (left) and the expected result (right) for DNA in Figure 5a correspond to post-PCR steps and plasmid extraction. For gel extraction steps, the yield will be less than what is shown. While the yield and purity of DNA can be suboptimal and still be successful, the yield

and purity of RNA require higher standards. Pay particular attention to the yield and purity of mRNA products as shown on the right in Figure 5b. Poor yield or purity as shown on the left in Figure 5b requires repeat IVT reaction. In addition to NanoDrop, mRNA products require additional analysis by a Bioanalyzer. Ensure that the mRNA products have clean and discrete expected bands as shown in electrophoresis and electropherogram. Representative successful (Figure 5c) and poor (Figure 5d) Bioanalyzer results are shown.

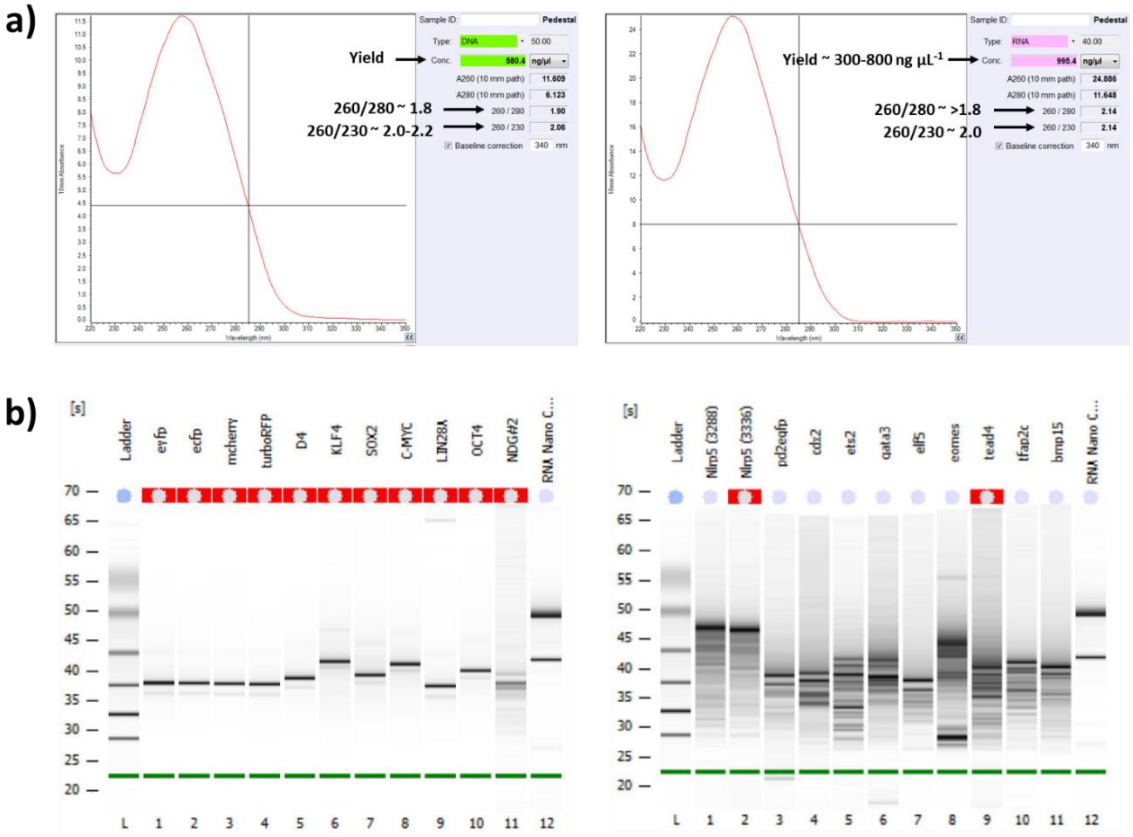


Figure 5: Quality control of DNA and RNA products. (A) Representative NanoDrop microvolume spectrophotometer measurement of an expected cloned DNA extraction (left) and of an expected IVT RNA result (right). (B) Analysis of successfully generated (left) and of unsuccessfully generated (right) modified mRNAs by Bioanalyzer automated electrophoresis instrument.

The results of modified mRNA transfections are shown in Figure 6. Protein expression was observed in transfected cells with modified mRNA encoding a number of different fluorophores in Figure 6a. A cocktail of modified mRNA composed of OCT4, SOX2, KLF4, c-MYC, and LIN28A was transfected daily into human fibroblasts for 10 days. As shown in Figure 6b, human fibroblasts started to undergo morphological changes over time during this reprogramming process. The transfection was stopped and two days later, the colonies that emerged show immunostaining for SSEA-4 and OCT-3/4.

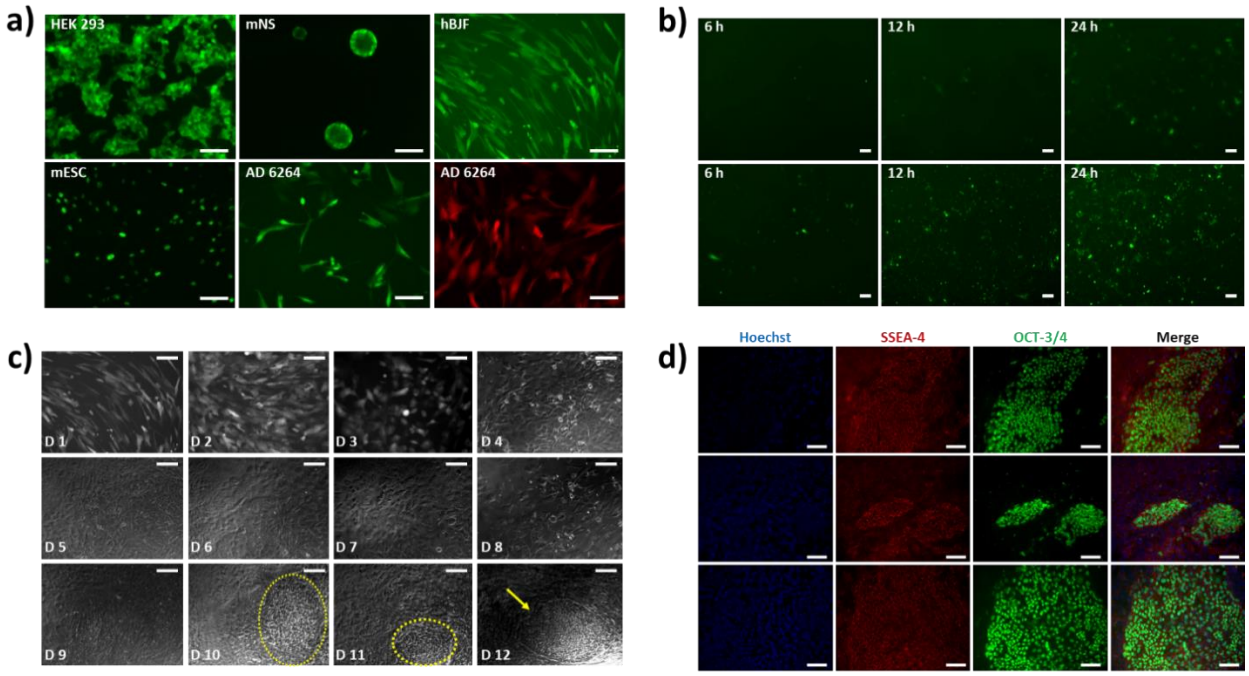


Figure 6: Modified mRNA protein expression and reprogramming of transfected cells. (A) A number of different cell lines were transfected with modified mRNAs encoding fluorophore variants: stabilized eGFP, destabilized eGFP, nuclear-localizing eGFP, or mCherry. HEK = human embryonic kidney, mNS = mouse neurosphere, hBJF = human foreskin fibroblasts, mESC = mouse embryonic stem cells, AD = human Alzheimer's fibroblasts. (B) Time-course comparison of eGFP expression on mESC by lentivirus (top) or by modified mRNA (bottom). Multiplicity of infection of 5 in the presence of polybrene at 8 μg mL^{-1} was used for lentiviral transduction. Lipofectamine RNAiMax was used for modified mRNA transfection. (C)

Modified mRNA cocktails comprised of OCT4, SOX2, KLF4, c-MYC, and LIN28A were transfected into human fibroblasts every day. Daily morphological changes are shown. Characteristic mesenchymal-to-epithelial transitions typically observed in cellular reprogramming are marked with yellow dashed circles. An emerging colony with sharp boundary is marked with a solid yellow arrow. (D) Emerging colonies after several days post-transfection showing immunostaining of SSEA-4 and OCT-3/4. Scale bars, 100 μm.

The comprehensive and detailed protocol presented here provides a non-viral, integration-free method for expressing proteins using modified mRNA.[89] This protocol can be applied to any gene of interest and therefore can replace most other gene expression methods. The inherent properties of synthetic modified mRNAs provide some unique advantages such as the ability to express varying stoichiometric ratios of vectors, to adjust the dose of each gene accurately, and to rapidly and completely silence expression.[37]

Furthermore, as shown in Figure 6b, modified mRNA method results in faster protein expression and higher efficiency than the commonly used DNA-based viral approach for hard-to-transduce cell types. Others have well characterized transfection efficiency, expression kinetics, and degradation rates of mRNA for in vitro cell culture and in vivo animal model.[123-126]

To fully realize all of the advantages of this method, some general precautions and optimizations are needed. Modified mRNA must be handled carefully and often requires new synthesis before experiments. All cell types must be optimized for transfection. This includes the dose of mRNA, types of transfection reagents, the ratios of mRNA to transfection reagents, the health and cell density, types of media, and other necessary supplements such as B18R or serum. All such optimizations should be monitored with modified mRNA encoding a fluorescent reporter to provide fast and convenient feedback. Such reporters should always be included in assays.

One of the challenges of this method is the labor-intensive and time-consuming work associated with it.

In contrast to an integrating viral method, which relies on the host cells to continuously generate the desired proteins, modified mRNAs must be repetitively synthesized and transfected into cells to maintain gene expression. The modified mRNA approach thus is not ideal for exploratory experiments or early stages of projects where large sets of genes are tested on a wide variety of cell types. One solution to this limitation is the construction of a plasmid that bears all parts of the IVT template minus the ORF. This would streamline some of the labor-intensive work of generating modified mRNAs encoding large numbers of different genes. However, this would limit the ability to change or customize UTR regions and potentially increase the chance of mutations in the amplification steps.

Another caveat when applying this method is that, despite careful UTR designs and incorporation of modified bases, synthetic modified mRNAs are still potentially recognized as non-self nucleic acids by mammalian cells.[114] This could potentially activate an innate immune response and cytokine signaling that alters the intended effects of the modified mRNAs. If such immune responses are encountered, B18R must be included in the culture medium to inhibit interferon secretion caused by the modified mRNA transfection, and stringent purification procedures by HPLC, instead of the standard silica-based spin column purification, would be necessary to further reduce immunogenic responses.[127]

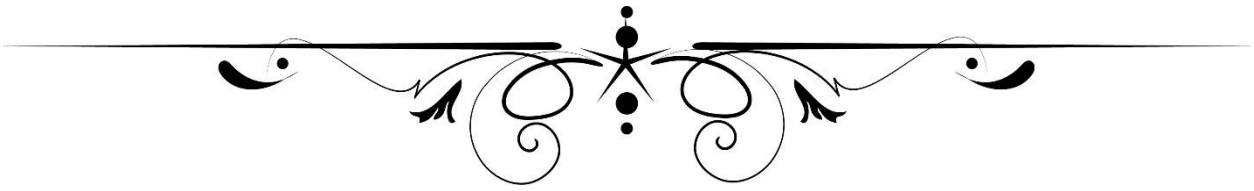
Despite the above limitations, synthetic modified mRNAs provide a safe and clinically applicable method for delivering genetic information into a mammalian cell. This protocol largely addresses previous concerns about mRNA stability and translation efficiency. Other nucleotide modifications can be incorporated (Figure 3) to optimize for the cellular background and increase the stability of mRNA. Importantly, synthetic modified mRNAs have been used successfully to reprogram and differentiate cells at a higher efficiency and faster kinetic than commonly used viral approaches.[37,87] This suggests that modified mRNA can be uniquely applied in areas of cell biology that may have been previously inaccessible by other methods.

Now that we have built the most advanced vehicle, the next thing we need for this journey is the map; MuERV-L reporter.

III.

Protocol

"This paper will no doubt be found interesting by those who take an interest in it." – John Dalton



This is the blueprint for building one of the most powerful tools in cellular reprogramming technology.

1. Design

1.1. *ORF-specific Primers and Oligonucleotides*

1.1.1. Synthesize the forward primer (ORF-FW) for the gene of interest.

Note: Do not include the start codon. The 5'UTR contains ATG.

1.1.2. Synthesize the reverse primer (ORF-RV) for the same ORF of interest.

Note: Ensure that the ORF-RV contains the stop codon.

1.1.3. Synthesize the 5'Bridge as follows: reverse complement of ORF-FW + CATGGTGGCTCTTATATTTCTTCTT

1.1.4. Synthesize the 3'Bridge as follows: CCCGCAGAAGGCAGC + ORF-RV

1.2. *Common Primers and Oligonucleotides*

1.2.1. Synthesize the Clone-FW/RV as follows: GCGTCGACACTAGTTCTAGACCCTCA (26 bp)

1.2.2. Synthesize the IVT-FW as follows: TTGGACCCTCGTACAGAAGCTAATACG (27 bp)

1.2.3. Synthesize the IVT-RV as follows: T_{120} CTTCCTACTCAGGCTTTATTCAAAGACCA (149 bp)

1.2.4. Synthesize the 5'UTR as follows:

GCGTCGACACTAGTTCTAGACCCTCATTGGACCCTCGTACAGAAGCTAATACGACTCACTATAGGGAAATAAGAGA
GAAAAGAAGAGTAAGAAGAAATATAAGAGCCACCATG (113 bp)

1.2.5. Synthesize the 3'UTR as follows:

GCTGCCTTCTGCGGGGCTTGCCTTCTGGCCATGCCCTTCTTCTCTCCCTTGCACCTGTACCTCTTGGTCTTTGAATAA
AGCCTGAGTAGGAAGTGAGGGTCTAGAACTAGTGTCGACGC (119 bp)

Note: Fluorescent reporter should be included for internal control and transfection optimization. Destabilized and nuclear-localizing fluorescent protein with excitation and emission ranges that will not interfere with the downstream immunohistochemistry is highly recommended. Starting with just the open reading frame (ORF) of fluorescent protein may be beneficial when going over this protocol for the first time.

2. Assembly

2.1. *Stock Solutions and Working Concentrations*

2.1.1. Add TE buffer (10 mM Tris·Cl, 1 mM EDTA, pH 8.0) to make 100 µM stock solution for each primer and oligonucleotide.

2.1.2. Further dilute each ORF-RV, 5'UTR, Clone-FW/RV, 5'Bridge, and 3'Bridge to achieve 1 µM working concentration.

2.2. *ORF-FW and 3'UTR N-Terminal Phosphorylation*

2.2.1. Prepare phosphorylation reaction for ORF-FW and 3'UTR as shown below:

Reagent	Amount (µL)	Final Concentration
ORF-FW or 3'UTR (100 µM)	3	6 µM
ATP (from MEGAscript T7 Kit, 75 mM)	0.7	1 mM
Polynucleotide Kinase Buffer (10X)	5	1X
Polynucleotide Kinase (10 U)	1	0.2 U μL^{-1}
Nuclease-free Water	40.3	50 µL total volume

Mix and centrifuge briefly.

2.2.2. Incubate at 37 °C for 30 min.

2.2.3. Denature polynucleotide kinase by incubating at 60 °C for 25 min.

2.2.4. Add 250 µL TE buffer to make 1 µM 5'phosphorylated (p-) ORF-FW and 3'UTR.

2.3. *p-ORF Amplification*

2.3.1. Prepare PCR reaction as shown below:

Reagent	Amount (µL)	Final Concentration
Template DNA	1	50-500 pg μL^{-1} (simple

(1-10 ng μL^{-1} for a simple template such as a plasmid or 10-100 ng μL^{-1} for a more complex template such as genomic DNA)		template) 0.5-5 ng μL^{-1} (complex template)
p-ORF-FW (1 µM)	4	200 nM
ORF-RV (1 µM)	4	200 nM
KAPA HiFi HotStart ReadyMix (2X)	10	1X
Nuclease-free Water	1	20 µL total volume

Mix and centrifuge briefly.

2.3.2. Amplify the ORF using PCR program as shown below:

Step	Temperature	Time	Cycles
Initial Denaturation	95.0 °C	5 min	1x
Denaturation	98.0 °C	20 sec	25x
Annealing	60.0-75.0 °C	15 sec	
Extension	72.0 °C	15-60 sec per kb	
Final Extension	72.0 °C	5 min	1x

2.4. *Isolation and Purification of p-ORF Product*

2.4.1. Run the products on gel electrophoresis.

2.4.2. Cut and extract only the product bands and elute in 30 µL Elution Buffer (EB).

Note: If there is no discrete band of expected size then perform gradient PCR from Step 2.3.2. Even though gel extraction is performed, obtaining clean ORF amplification needs to be prioritized in order to minimize potential non-specific downstream reactions.

2.4.3. Use Nanodrop microvolume spectrophotometer to verify reasonable amplification. Adjust to 24 ng μL^{-1} per sample.

Note: If large amounts of p-ORF products are lost during the extraction step, decrease the EB to 10 µL. Minimum of 5 µL at 24 ng μL^{-1} for each sample of amplified products are required for the next step.

Pause Point: Extracted and purified p-ORF can be stored at -20 °C for a prolonged time.

2.5. *5'UTR-ORF-3'UTR Double Ligation*

2.5.1. Prepare double ligation reaction as shown below:

Reagent	Amount (µL)	Final Concentration
p-ORF (24 ng μL^{-1})	1	1.2 ng μL^{-1}
5'UTR (1 µM)	4	200 nM
5'Bridge (1 µM)	2	100 nM
p-3'UTR (1 µM)	4	200 nM
3'Bridge (1 µM)	2	100 nM
Ampligase Reaction Buffer (10X)	2	1X
Ampligase DNA Ligase (5 U μL^{-1})	1	0.25 U μL^{-1}
Nuclease-free Water	4	20 µL total volume

Mix and centrifuge briefly.

2.5.2. Run PCR double ligation program as shown below:

Step	Temperature	Time	Cycles
1st Incubation	95.0 °C	10 sec	5x
2nd Incubation	45.0 °C	1 min	
3rd Incubation	50.0 °C	1 min	
4th Incubation	55.0 °C	1 min	
5th Incubation	60.0 °C	1 min	

2.5.3. Purify the products using QIAquick PCR purification kit and elute with 10 µL elution buffer (EB).

2.6. *5'UTR-ORF-3'UTR Amplification*

2.6.1. Prepare PCR reaction as shown below:

Reagent	Amount (µL)	Final Concentration
5'UTR-ORF-3'UTR (unamplified)	5	---
Clone-FW/RV (1 µM)	10	200 mM
KAPA HiFi HotStart ReadyMix (2X)	25	1X
Nuclease-free Water	10	50 µL total volume

Mix and centrifuge briefly.

2.6.2. Amplify using PCR program as shown below:

Step	Temperature	Time	Cycles
Initial Denaturation	95.0 °C	5 min	1x
Denaturation	98.0 °C	5 sec	25x
Annealing	60.0-75.0 °C	15 sec	
Extension	72.0 °C	15-60 sec per kb	
Final Extension	72.0 °C	5 min	1x

2.6.3. Run the products on gel electrophoresis.

2.6.4. Cut and extract only the product bands and elute in 40 µL EB.

Note: If the products are absent, ligate each UTR one at a time. Refer to Step 2.7.

Pause Point: Amplified 5'UTR-ORF-3'UTR can be stored at -20 °C for a prolonged time.

2.7. Optional: *5'UTR-ORF-3'UTR Single Ligation*

Note: This Step should be performed only when double ligation has failed.

2.7.1. Prepare the first single-ligation reaction as shown below:

Reagent	Amount (µL)	Final Concentration
p-ORF (24 ng μL^{-1})	1	1.2 ng μL^{-1}
5'UTR (1 µM)	4	200 nM
5'Bridge (1 µM)	2	100 nM
Ampligase Reaction Buffer (10X)	2	1X

Ampligase DNA Ligase (5 U μL^{-1})	1	0.25 U μL^{-1}
Nuclease-free Water	10	20 µL total volume

Mix and centrifuge briefly.

2.7.2. Run the PCR program as shown in Step 2.6.2.

2.7.3. Purify the products using PCR purification kit and elute in 12 µL EB.

2.7.4. Prepare the second single-ligation reaction as shown below:

Reagent	Amount (µL)	Final Concentration
5'UTR-ORF	10	
p-3'UTR (1 µM)	4	200 nM
3'Bridge (1 µM)	2	100 nM
Ampligase Reaction Buffer (10X)	2	1X
Ampligase DNA Ligase (5 U μL^{-1})	1	0.25 U μL^{-1}
Nuclease-free Water	1	20 µL total volume

Mix and centrifuge briefly.

2.7.5. Run the PCR program as shown in Step 2.6.2.

2.7.6. Purify the products using PCR purification kit and elute in 10 µL EB.

2.7.7. Amplify the purified ligated 5'UTR-ORF-3'UTR by following Step 2.6.

2.8. *5'UTR-ORF-3'UTR Cloning*

Note: The following Steps use TOPO TA cloning method that is sensitive to the size of the inserts. If inserts are large, or TOPO TA cloning was unsuccessful, perform the traditional cloning method of restriction endonucleases followed by T4 ligation instead.

2.8.1. Prepare dA-Tail reaction as shown below:

Reagent	Amount (µL)	Final Concentration
5'UTR-ORF-3'UTR (amplified)	40	---
dA-Tailing Reaction Buffer (10X)	5	1X
Klenow Fragment (3'→5' exo$^-$)	3	---
Nuclease-free Water	2	50 µL total volume

Mix and centrifuge briefly.

2.8.2. Incubate at 37.0 °C for 30 min.

2.8.3. Purify the products using PCR purification kit and elute in 30 µL EB.

2.8.4. Prepare TOPO TA cloning reaction as shown below:

Reagent	Vector Only (Negative Control)	Vector + Insert (Positive Control)	Vector + Insert (Experimental)
Positive Control PCR Insert	---	1 µL	---
5'UTR-ORF-3'UTR dA-tailed	---	---	4 µL
TOPO Vector	1 µL	1 µL	1 µL

Salt Solution (from TA Cloning Kit)	1 µL	1 µL	1 µL
Nuclease-free Water	4 µL	3 µL	---
Final Volume	6 µL	6 µL	6 µL

Mix and centrifuge briefly.

Note: Ensure to include reactions for positive and negative controls.

2.8.5. Incubate at room temperature for 30 min.

Pause Point: The cloned reactions can be stored at -20 °C for a prolonged time.

2.9. *Cloned 5'UTR-ORF-3'UTR Transformation*

2.9.1. Set up multiple transformations with varying amounts of cloned reactions per sample. Transform into competent cells according to the manufacturer's instructions.

Note: Be sure to include positive and negative transformation controls.

2.9.2. Compare against the positive and negative controls to confirm that the transformation was successful. Pick 4-8 colonies that are grown overnight on a selective plate and inoculate each colony into a culture of 1-5 mL LB medium containing the appropriate selective antibiotic. Incubate at 37.0 °C for 12-16 hr with vigorous shaking.

Note: Pick larger colonies that tend to have correct inserts than the smaller colonies.

2.9.3. Extract and purify the plasmid DNA using a QIAprep Spin Miniprep kit according to the manufacturer's instructions.

Note: QIAprep Maxiprep kit is preferred for large-scale experiments.

2.9.4. Adjust the plasmid concentration to 100 ng μL^{-1}.

Pause Point: Extracted plasmids can be stored at -20 °C for a prolonged time.

2.10. *Plasmid Linearization and Extraction*

Note: RNA polymerases are very processive therefore even a small amount of undigested circular DNA templates can generate run-on heterogenous transcripts. Ensure that plasmids are completely linearized with appropriate restriction enzymes downstream of the insert.

2.10.1. Prepare digestion reaction as shown below:

Reagent	Amount (µL)	Final Concentration
Plasmid (100 ng μL^{-1})	40	20 ng μL^{-1}
Restriction Enzyme	4	0.1-1.0 U μL^{-1}
NEBuffer (10X)	20	1X
Nuclease-free Water	136	200 µL total volume

Mix and centrifuge briefly.

2.10.2. Incubate at 37 °C for 1 hr.

2.10.3. Run the products on gel electrophoresis.

Note: Ensure that the plasmids are completely digested. Run the undigested plasmid in another lane as a control.

2.10.4. Cut and extract only the linearized plasmid bands and elute in 50 µL EB.

Pause Point: Linearized plasmids can be stored at -20 °C for a prolonged time.

2.11. *Attach Poly(T)-tail and Sequence Validation*

2.11.1. Prepare the Poly(T) attachment master mix as shown below:

Reagent	Amount (µL)	Final Concentration
Digested Template DNA	16	---
IVT-FW (1 µM)	12	30 nM
IVT-RV (1 µM)	12	30 nM
KAPA HiFi HotStart ReadyMix (2X)	200	1X
Nuclease-free Water	160	400 µL total volume

Mix and centrifuge briefly.

2.11.2. Aliquot the master mix into 8 PCR tubes (50 µL each) and amplify by running the PCR program as shown below:

Step	Temperature	Time	Cycles
Initial Denaturation	95.0 °C	3 min	1x
Denaturation	98.0 °C	20 sec	15-25x
Annealing	60.0 °C	15 sec	
Extension	72.0 °C	15 sec per kb	
Final Extension	72.0 °C	3 min	1x

2.11.3. Run the products on gel electrophoresis. Cut and pool all product bands. Extract and elute in 50 µL EB.

2.11.4. Adjust the final concentration to 100 ng μL^{-1} using EB.

2.11.5 Run an aliquot on gel electrophoresis to verify proper amplification. Verify clean extraction.

Note: If non-specific amplification was observed then reduce the number of PCR cycles on Step 2.11.2.

2.11.6. Sequence the templates prior to the next step.

Pause Point: Templates can be stored at -20 °C for a prolonged time.

3. Production

Note: Decontaminate all working surfaces and pipettes with RNase*Zap*. Ensure that all microcentrifuge tubes and pipette tips are RNase-free.

3.1. *In Vitro Transcription*

3.1.1. Spin down and reconstitute one vial of 3'-O-Me-m^7G(5')ppp(5')G RNA cap analog with 16.7 µL nuclease-free water to obtain 60 mM stock solution. Use GTP and ATP from MEGAscript T7 kit. Prepare the custom nucleotide triphosphates (NTPs) mix as shown below:

Reagent	Stock Concentration (mM)	Amount (µL) per IVT reaction	NTP Mix Concentration (mM)	Final IVT Reaction Concentration (mM)
RNA cap analog	60	4.0	16.2	6.0
GTP	75	0.8	4.1	1.5

ATP	75	4.0	20.3	7.5
5-Me-CTP	100	3.0	20.3	7.5
Pseudo-UTP	100	3.0	20.3	7.5

Mix and centrifuge briefly.

Note: Scale up the amount of each reagent according to the desired number of IVT reactions.

3.1.2. Add the IVT reagents in a PCR tube at room temperature in the following order as shown below:

Reagent	Amount (µL)	Final Concentration
NTP Mix (from Step 3.1.1)	14.8	Refer to Step 3.1.1
Tailed Poly(T) PCR Product, 100 ng μL^{-1} (from Step 2.15)	16.0	40 ng μL^{-1}
T7 Buffer (10X)	4.0	1X
T7 Enzyme (10X)	4.0	1X
Nuclease-free Water	1.2	40 µL total volume

Mix and centrifuge briefly.

Note: Multiple reactions should be set up based on a scale of the experiment.

3.1.3. Incubate at 37 °C for 4-6 hrs in a thermocycler. Do not incubate in a water bath to avoid contamination with RNases.

3.2. *DNase Treatment*

3.2.1. Add 2 µL of Turbo DNase (from MEGAscript T7 kit) to each reaction. Mix gently and centrifuge briefly. Incubate at 37 °C for 15 min in a thermocycler.

3.2.2. Purify each reaction using MEGAclear kit. Follow the RNA elution option 2 by pre-heating 110 µL of EB to 95 °C and eluting twice with 50 µL for a total of 100 µL.

3.3. *Phosphatase Treatment*

3.3.1. Add 11 µL of 10X Antarctic Phosphatase Buffer and 2 µL of Antarctic Phosphatase to each reaction. Mix gently and centrifuge briefly.

3.3.2. Transfer each reaction to PCR tube and incubate at 37 °C for 1 hr in a thermocycler.

3.3.3. Purify using MEGAclear kit as described in Step 3.2.2.

3.4. *Quality Control*

3.4.1. Measure the concentration of each purified mRNA sample using a NanoDrop. The expected yield is between 300-800 ng μL^{-1} in 100 µL elution volume. The absorbance ratios of 260 nm / 280 nm should be greater than 1.8 and 260 nm / 230 nm should be close to 2.0. Absorbance ratios outside of the expected values require repetition of the reaction from Step 3.1.2.

3.4.2. Adjust each reaction to 100 ng μL^{-1} by adding EB. Reactions with yields too low to achieve 100 ng μL^{-1} concentration need to be repeated from Step 3.1.2.

Note: Higher concentrations may be necessary depending on the cell types, media, expression levels, and other experimental parameters. Save an aliquot of undiluted reaction in case a higher dose is preferred.

3.4.3. Analyze the integrity of the synthetic modified mRNA using a Bioanalyzer. Running a small aliquot on a denaturing agarose gel should be performed to check the integrity if a Bioanalyzer is inaccessible.

Note: If the integrity analysis shows non-specific amplification and/or degradation, repeat the Production from Step 3.1.2.

Pause Point: Purified synthetic modified mRNA can be stored at -80 °C for several months.

4. Transfection

4.1. *Synthetic Modified mRNA Cocktail Preparation*

Note: Spike each cocktail with an equimolar mRNA fluorescent reporter to monitor transfection efficiency. We recommend using destabilized fluorescent protein that has a half-life of hours. This allows for accurate monitoring of daily transfection. Addgene has pRVGP (plasmid 15377) that contains the gene for destabilized eGFP which can be used as a DNA template to generate mRNA fluorescent reporter. Thaw all synthetic modified mRNAs on ice.

4.1.1. Determine the nucleotide base length of each ORF. This value should be derived from the template DNA on which the ORF PCR was performed. If genomic DNA was used as the template, determine the base length via NCBI Gene database and confirm with the sequencing data.

4.1.2. To the determined value from the previous (Step 4.1.1), add 297 base pairs. This addition represents the Poly(T)-tail and parts of the 5'UTR and 3'UTR that are carried over to IVT.

4.1.3. Convert each determined nucleotide length from the previous (Step 4.1.2) to the mole. Refer to the table below for conversion:

Size (nt)	20	100	300	500	1000	1800
pmol	152.23	31.05	10.38	6.23	3.12	1.73

Note: Nucleotide lengths that are outside the range of this table can be converted to moles by scaling up

or down proportionately to the values in the table.

4.1.4. Alternatively, NEBioCalculator from New England BioLabs website can be used. Go to ssRNA: Moles to Mass Convertor. Input the RNA length as determined from Step 4.1.2 then input the arbitrary but same value for RNA moles, 1 for example, to obtain RNA mass for each sample.

4.1.5. Normalize the values of each mRNA product calculated from the previous steps by dividing by the smallest number (moles for Step 4.1.3 or RNA mass for Step 4.1.4) across all samples that will be mixed in a cocktail.

4.1.6. Determine the total volume of cocktail mRNA needed for a given experiment, then add the normalized volumetric ratios of each mRNA product to meet this demand. Mix gently but thoroughly and centrifuge briefly.

Note: For situations where different molar ratios are desired, make the appropriate calculation changes on Step 4.1.3 or 4.1.4.

4.1.7. Aliquot the cocktail for single-use in order to minimize freeze/thaw cycle.

Note: Cocktail aliquots can be stored at -80 °C for several months.

4.2. *Identifying the Optimal Transfection Reagent*

Note: Cell types, media, and other experimental designs can greatly affect the success of mRNA transfection. To address this, we have provided three transfection reagents that have previously shown success. Choose the reagent that results in most successful transfection for a given experimental design as determined by the fluorescent reporter. We find that 1 µg mRNA for one well of 12-well plate was optimal in most conditions. The optimal dose of mRNA for a given experiment may need to be determined by titration. The following reaction volumes should be scaled according to the number of wells to be

transfected.

4.2.1. Lipofectamine RNAiMax

4.2.1.1. Add 1 µg (10 µL of 100 ng μL^{-1}) mRNA cocktail into 40 µL Opti-MEM I Reduced-Serum Medium. Gently pipet and mix thoroughly.

4.2.1.2. In a separate tube, add 5 µL Lipofectamine RNAiMax into 45 µL Opti-MEM I Reduced-Serum Medium. Gently pipet and mix thoroughly.

4.2.1.3. Combine the 50 µL diluted mRNA and the 50 µL diluted Lipofectamine RNAiMax. Gently pipet and mix thoroughly.

4.2.1.4. Incubate at room temperature for 15 min.

4.2.1.5. Add the complexes drop-wise to one well of a 12-well plate. Gently rock the plate back-and-forth and from side-to-side. Incubate for 4 hrs at 37 °C. Longer incubation may increase cytotoxicity.

4.2.2. *Trans*IT-mRNA

4.2.2.1. Add 1 µg (10 µL of 100 ng μL^{-1}) mRNA cocktail into 100 µL Opti-MEM I Reduced-Serum Medium. Gently pipet and mix thoroughly.

4.2.2.2. Add 2 µL mRNA Boost Reagent to the above. Gently pipet and mix thoroughly.

4.2.2.3. Add 2 µL *Trans*IT-mRNA Reagent to the above. Gently pipet and mix thoroughly.

4.2.2.4. Incubate at room temperature for 3 min. Do not incubate more than 5 min.

4.2.2.5. Add the complexes drop-wise to one well of a 12-well plate. Gently rock the plate back-and-forth and from side-to-side. Incubate for overnight at 37 °C.

4.2.3. Stemfect

4.2.3.1. Add 1 µg (10 µL of 100 ng μL^{-1}) mRNA cocktail into 60 µL Stemfect buffer. Gently pipet and mix thoroughly.

4.2.3.2. In a separate tube, add 4 µL Stemfect transfection reagent into 60 µL Stemfect buffer. Gently pipet and mix thoroughly.

4.2.3.3. Combine the 70 µL diluted mRNA and the 64 µL diluted Stemfect transfection reagent. Gently pipet and mix thoroughly.

4.2.3.4. Incubate at room temperature for 15 min.

4.2.3.5. Add the complexes drop-wise to one well of a 12-well plate. Gently rock the plate back-and-forth and from side-to-side. Incubate for overnight at 37 °C.

4.3. After the indicated incubation time, change the medium. If desired, mRNA transfection can be repeated daily or once every two days for continuous expression. Observe for fluorescent reporter and cell toxicity over the next few days. Confirm the expression of encoded protein by western blot or immunocytochemistry.

Note: If high cytotoxicity was observed, supplement the media with B18R at 200 ng ml^{-1}. If low fluorescent reporter expression was observed, increase the dose of mRNA. Increasing the cell density and using low passage cells improve cell survival and transfection uptake. Additional optimization may be necessary.

Name of Material/ Equipment	Company	Catalog Number	Comments/Description
KAPA HiFi HotStart ReadyMix	KAPA Biosystems	KK2602	store at -20°C
Ampligase Enzyme and Buffer	Epicentre	A32750	store at -20°C

	Biotechnologies		
T4 Polynucleotide Kinase	New England Biolabs (NEB)	M0201S	store at -20°C
NEBNext dA-Tailing Module	NEB	E6053S	store at -20°C
TOPO TA Cloning Kit	Invitrogen	K4500-01	store at -20°C
One Shot TOP10 Chemically Competent E. coli	Invitrogen	C4040-03	store at -80°C
S.O.C. Medium	ThermoFisher Scientific	15544034	store at RT
Difco LB Broth	BD	240230	store at RT
Ampicillin, sodium salt, irradiated	ThermoFisher Scientific	11593027	store at 4°C
Falcon® 100 mm x 15 mm Not TC-Treated Bacteriological Petri Dish, Sterile	Corning	351029	store at RT
50x Tris/Acetic Acid/EDTA (TAE), Nucleic Acid Electrophoresis Buffer	Bio-Rad	1610743	store at RT
Agarose	Denville Scientific Inc	GR140-500	store at RT

QIAquick Gel Extraction Kit	QIAGEN	28704	store at RT
QIAquick PCR Purification Kit	QIAGEN	28106	store at RT
QIAprep Spin Miniprep Kit	QIAGEN	27106	store at 4°C for P1 Buffer, rest at RT
MEGAscript T7 Kit	Ambion	AM1334	store at -20°C
MEGAclear Kit	Ambion	AM1908	store at 4°C
Antarctic Phosphatase	NEB	M0289S	store at -20°C
3'-O-Me-m7G(5')ppp(5')G RNA Cap Structure Analog	NEB	S1411S	store at -20°C
Pseudouridine-5'-Triphosphate	TriLink BioTechnologies	N-1019	store at -20°C
5-Methylcytidine-5'-Triphosphate	TriLink BioTechnologies	N-1014	store at -20°C
SalI	NEB	R3138S	store at -20°C
SpeI	NEB	R0133S	store at -20°C
XbaI	NEB	R0145S	store at -20°C
Lipofectamine RNAiMAX	Invitrogen	13778-075	store at 4°C
TransIT-mRNA Transfection Kit	Mirus Bio	MIR 2225	store at 4°C

Stemfect™ RNA Transfection Kit	Stemgent	00-0069	store at 4°C
B18R Recombinant Protein Carrier-free	eBioscience	34-8185-85	aliquot and store at -80°C
StemfactorTM B18R Recombinant Protein, Carrier-free	Stemgent	03-0017	aliquot and store at -80°C
pRVGP (d2eGFP)	Addgene	Plasmid 15377	sequence verify and store at 4°C for short-term or -20°C for long-term
RNaseZap™ RNase Decontamination Solution	Life Technologies	AM9780	store at RT
TE, pH 7.0	ThermoFisher Scientific	AM9861	store at RT
Opti-MEM™ I Reduced Serum Medium	ThermoFisher Scientific	31985070	store at 4°C
Nucleases-free 0.5 mL Sterile Microcentrifuge Tubes	Capitol Scientific	022600001	store at RT
Nucleases-free 1.5 mL Sterile Microcentrifuge Tubes	Fisher Scientific	05-402-24B	store at RT

Falcon 15 mL Conical Centrifuge Tubes	Fisher Scientific	352097	store at RT
Falcon 50mL High Clarity PP Centrifuge Tube, Conical Bottom, Sterile	Corning	352098	store at RT
Falcon 24 Well Clear Flat Bottom TC-Treated Sterile	Corning	353047	store at RT
Falcon 12 Well Clear Flat Bottom TC-Treated Sterile	Corning	353043	store at RT
Falcon 6 Well Clear Flat Bottom TC-Treated Sterile	Corning	353046	store at RT
Corning 100mm TC-Treated Culture Dish	Corning	430167	store at RT
Nucleases-free Sterile Aerosol-barrier Tips	Denville Scientific Inc		10, 20, 200, 1000 µL
Nucleases-free Sterile Serological Pipettes	Fisher Scientific		1, 5, 10, 25 mL
Pipetman	Fisher Scientific		P2, P10, P20, P200, P1000
Pipet-Aid	Fisher Scientific		or equivalent
Gloves	Denville Scientific		or equivalent

	Inc		
Gel Electrophoresis Equipment	Bio-Rad		or equivalent
Vortex-Mixers	VWR		or equivalent
Microcentrifuges	Eppendorf		or equivalent
ThermoMixer	Eppendorf		or equivalent
PCR Thermocycler	Bio-Rad		or equivalent
NanoDrop	ThermoFisher Scientific		or equivalent
Bioanalyzer	Agilent Technologies		or equivalent
Cell Incubator	ThermoFisher Scientific		or equivalent
Fluorescence Microscope	Nikon		or equivalent

IV.

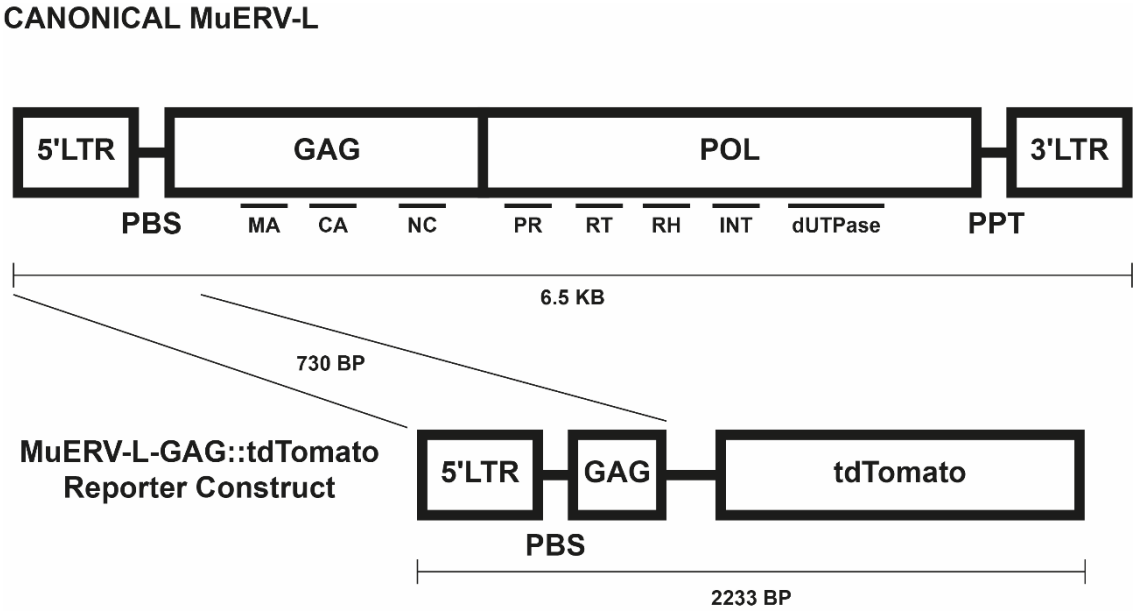
MuERV-L Reporter

"The farther the experiment is from theory, the closer it is to the Nobel Prize." – Irene Joliot-Curie

In vitro dedifferentiation is a very inefficient process and most cells reprogram unsuccessfully. A reliable totipotent reporter can provide information about the direction of reprogramming and can help identify potential iTCs. The reporter can also aid in selecting for candidate genes that may be necessary for inducing totipotency. However, the totipotent state is very transient that begins at one-cell stage embryo and peaks at two-cell then quickly disappears shortly after.[16,128] Majority of genes that are expressed during one- to two-cell embryos continue to express long after the disappearance of totipotency. Interestingly, a particular retrotransposon named murine endogenous retrovirus-like (MuERV-L or MERVL) is expressed in one- and two-cell embryos, but silenced after two-cell, making this an ideal marker of totipotency.[78,79,129,130] The main mechanism of retrotransposon silencing is DNA methylation which is largely absent during totipotency due to unusually hypomethylated genome state.[16,131] Other

retrotransposons are also expressed, but MuERV-L, in particular, has been well established as a marker of totipotency.[78,79,132]

As depicted in Figure 7, the canonical MuERV-L is 6.5 kb in size. Handling such large construct to establish a reporter system can cause numerous technical problems. To mitigate this, only the critical components of canonical MuERV-L were used. Fortunately, much of the expression activity resides in the 5' end. Therefore, the first 730 bp which include portions of GAG elements were amplified and combined with tdTomato. The choice of using tdTomato was deliberate. Since partial reprogramming and partial induction of totipotency will likely occur, resulting in weak expression of MuERV-L, we opted for an exceptionally bright red fluorescent protein to detect even the subtle changes.



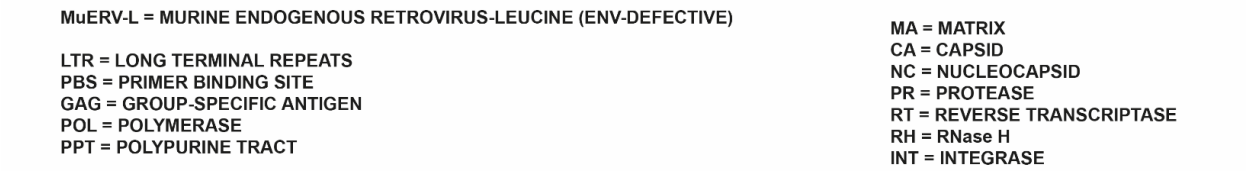


Figure 7: Schematic diagram of the reporter design. The canonical MuERV-L is shown. The 5' end is responsible for expression during totipotency. The 730 bp portion of 5' end of the canonical MuERV-L was combined with tdTomato, resulting in 2233 bp size totipotent reporter construct.

The reporter construct, MuERV-L-GAG::tdTomato, created from the previous step was cloned into a vector which was then transformed into ultracompetent *E. coli* for high-fidelity DNA amplification. After MaxiPrep, the purified reporter construct was isolated from the vector backbone and linearized. The linearized reporter was then injected into the pronuclei of the mouse zygotes in the picograms range, as shown in Figure 8. Control was also performed in which the reporter construct was digested between the MuERV-L-GAG and tdTomato then injected into the pronuclei. The number of embryos that survived the injection was comparable to the non-injection, indicating minimal toxicity from the pronuclear microinjection.

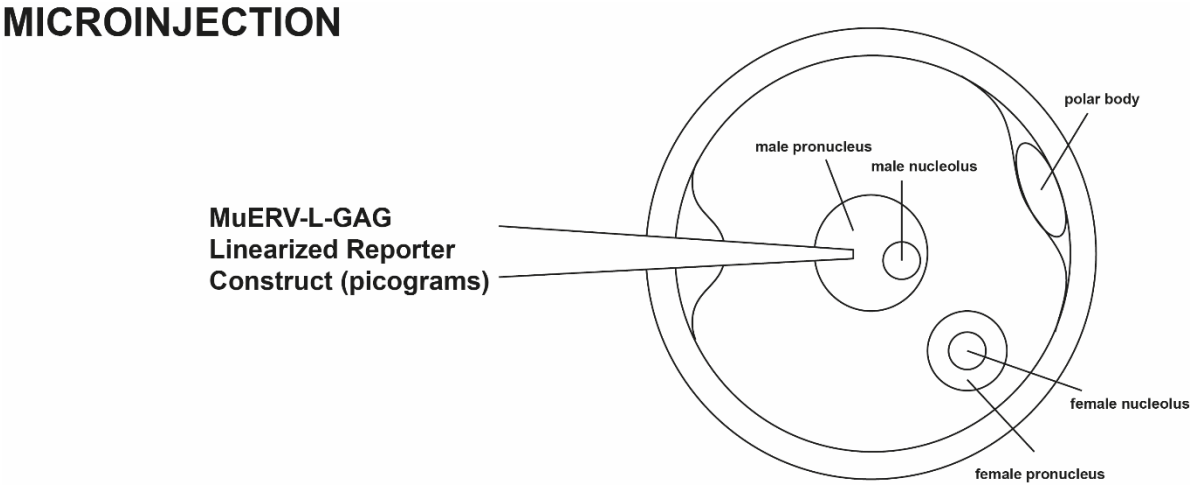


Figure 8: Reporter DNA microinjection. The linearized construct of MuERV-L-GAG::tdTomato and of digested MuERV-L-GAG::tdTomato were microinjected into the pronuclei of zygotes. The zygotes were derived from 12-hour post fertilized eggs that were verified by microscope based on one cell with two pronuclei morphology. The injection site occurred at the larger of the two pronuclei.

Microinjected zygotes were transferred to live-imaging 37 °C incubator. Expressions of tdTomato were observed, as shown in Figure 9, in MuERV-L-GAG::tdTomato but not in digested MuERV-L-GAG::tdTomato. Furthermore, expressions of tdTomato only occurred in subpopulations of one-cell and two-cell stage embryos.

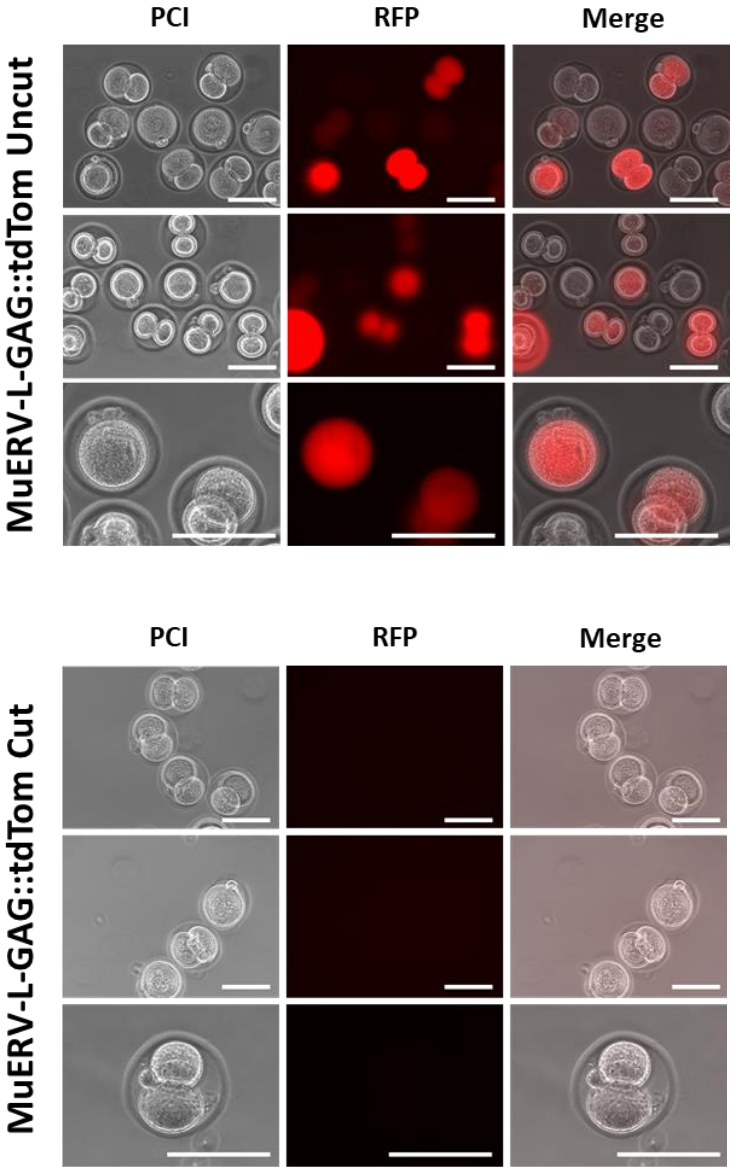


Figure 9: Reporter expression validation. Top shows the linearized MuERV-L-GAG::tdTomato reporter injected embryos. Bottom shows the negative control of the injected reporter that was cut between the MuERV-L-GAG and tdTomato. Embryos used in both experiments were derived from the same mouse. Scale bars, 100 µm.

Reporter microinjected zygotes were quantified for their tdTomato expressions at subsequent developmental stages, as shown in Figure 10. Quantifications revealed that tdTomato expressions were,

as expected, restricted to zygotes and 2-cell stage embryos. Interestingly, the quantification also revealed that not all of the zygotes and 2-cell stage embryos showed tdTomato expressions. Whether influenced by the microinjection or by the innate natural selection, these particular cells appear to have transcript landscapes that are insufficient to induce MuERV-L despite their morphological resemblance to totipotency. This suggests that the reporter is highly specific but is also very stringent for cells that are in the totipotent state.

Table 1

Early Embryos > Totipotent Cells > 1-cell (zygote) and 2-cell > MuERV-L-GAG::tdTomato^{+}

	ZYGOTES	2-CELLS	4-/8-CELLS	MORULA
tdTomato^{+}	14	21	0	0
tdTomato^{-}	26	39	5	6
	54%	54%	0%	0%

MuERV-L-GAG::tdTomato expression is specifically restricted and highly stringent to the most prime totipotent cells. The expression was only observed from the subset of the healthy totipotent zygotes and two-cells only.

Figure 10: Summary table of all embryos that were injected with MuERV-L-GAG::tdTomato reporter construct. Expressions of tdTomato were detected only in zygotes and 2-cell stage embryos. None were observed in later development. About half of zygotes and 2-cell stage embryos did not express tdTomato.

After validating MuERV-L-GAG::tdTomato as a totipotent marker, mouse ESC line bearing this reporter construct has been established. The reporter was cloned into a vector containing hygromycin b resistance which was then transfected into early passage wild-type mouse ESCs. The transfected cells were cultured in hygromycin b for two weeks then the surviving cells were transferred and expanded in 2i/L culture

media. The transfected ESCs maintained round-shaped colonies with clear margins and displayed proliferation rates comparable to that of untransfected ESCs. This totipotent reporter ESC line shows that a tiny percentage, about 0.5%, of the cells transiently express tdTomato, confirming the previous finding that a rare transient cell population within mouse ESCs can express transcripts typically found in 2-cell stage embryos.[78] Therefore, the reporter, MuERV-L-GAG::tdTomato, will also be referred to as "2C" to signify the similarities between the tdTomato+ cells and the 2-cell stage embryos.

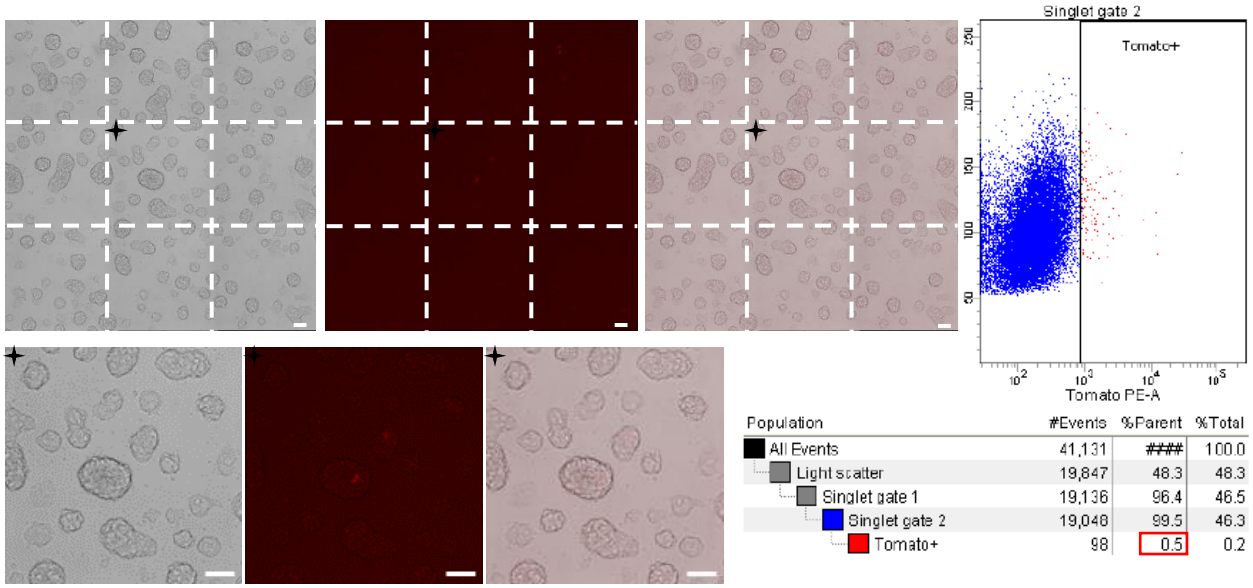


Figure 11: Establishing totipotent reporter mouse ESC cell line. MuERV-L-GAG::tdTomato DNA construct was transfected into early passage wild-type mouse ESCs and selected for stable clonal integrants. Top images show stitched nine viewing fields of low-magnification and bottom images show higher-magnification of the center viewing field. The black cross indicates the same viewing field and shows the faint tdTomato expressing cell. The FACS quantification on the right shows that 0.5% of cells express tdTomato. Scale bars, 100 µm.

Correct identification of cells that have successfully reprogrammed to totipotency poses an exceptional challenge since the state of totipotency is very brief and transient. Unlike culturing pluripotent stem cells

with LIF, there is currently no known way to continuously maintain totipotent cells without spontaneous differentiation. Continous expression of totipotent activating factors may be one approach to maintain a cell's totipotent state until a more suitable inhibitor molecule similar to LIF is discovered. This is assuming that the cell's tendency to achieve homeostasis by downregulating overactive genes does not occur.

One notable finding that has never been reported by others is the stringent nature of MuERV-L activation. While only the zygotes and 2-cell embryos expressed tdTomato, which have been shown previously, we also noticed that about half of these totipotent cells did not show tdTomato. This was in the context of post pronuclear injection which can be somewhat invasive to a cell and may have caused disturbances in the integrity of totipotency. Another possibility is that fertilized eggs are not all the same and often vary in their degree of totipotency and their subsequent ability to form successful embryos. This may partly explain relatively high spontaneous abortions seen in mammals. If true, this would mean that not all fertilized eggs, or zygotes, are totipotent or capable. Perhaps then the most accurate definition of totipotent cells are those subpopulations of zygotes which have subtle differences from others that allow them to be capable, in which case the new name may be needed to distinguish these particular cells further.

V.

Candidates

"If I have a thousand ideas and only one turns out to be good, I am satisfied." – Alfred Nobel

Thousands of genes have been associated with totipotency but current understanding of cell reprogramming, iPSC in particular, suggests that only a few essential genes are necessary.[33,34,133] Correctly identifying key factors that can induce totipotency relies on the existing knowledge about the earliest embryo development. Unfortunately, investigating the earliest embryo development has been a challenge, in part, due to the relative paucity of cells that are not easily available in large quantities. Not surprisingly, much of the molecular mechanisms of totipotency remain a mystery.

There are knockout gene studies that show developmental arrest at the early stages of embryogenesis, implying the possible importance of such genes in ZGA.[134-138] However, many of these knockout studies are one-offs and have not been repeated or further characterized—author bias may be a contributory

factor. The genes from these studies often have weak associations to totipotency. Many non-totipotent genes are embryonically lethal when completely absent but bear no role in inducing totipotency. More evidence and better characterization are needed before building upon this, potential but currently fragile, foundation.

Most unbiased studies use RNAseq and microarray data that compare each embryo stages throughout its early development.[133,139-141] However, these types of studies just provide a long list of differentially expressed genes with each being more-or-less equally probable due to complete lack of relevant information about them. The odds of randomly selecting the correct factors from such database is far too low.

If we follow the abundance of evidence and most well-characterized parts of ZGA, a rather different class of genes begin to stand out. Instead of picking specific transcription factors, an alternative approach to inducing totipotency is mimicking the known overall changes in chromosome and its epigenetic effects. In particular, maternal versions of histone variants are known to interact with chromatin differently than canonical variants and influence wide variety of genes, presumably including totipotent associated genes.[142-145] Interestingly, some maternal variants primarily exist during zygotes and 2-cell stages and have shown to improve the efficiency of iPSC reprogramming.[142,146,147] Also, exogenous molecules such as histone deacetylation inhibitors, which are chromosome modulating compounds, have shown to improve the efficiency of SCNT.[148-150] Perhaps then the candidates should come from those of maternal origin that activate genes indirectly via epigenome.

Candidate genes were initially chosen from the published literature based on several criteria. Some of these criteria include expression in totipotent cell but not in pluripotent stem cell, maternal-effect epigenome remodeling that corresponds to ZGA, quality of data and journal impact factors, maternal

variants that are only found in mature oocytes and early parts of one- to two-cell embryos, general consensus by the totipotent-related review papers, and reasonable ORF size with verified DNA sequence that would allow for successful generation of synthetic modified mRNA.[128,142,146,147,151-167] As shown in Figure 12, the initial list of 12 candidates and GFP ORFs were processed to generate synthetic modified mRNAs for reprogramming experiments.

Official Name	Common Names	NCBI ID	Summary
Hist1h2aa	Th2a	319163	histone variant
H3f3b	H3.3b; 9430068D06Rik	15081	histone variant
H1foo	H1.8; H1fo; C86609	171506	histone variant
p-Npm2	---	328440	nucleophosmin
Zscan4d	EG545913	545913	zinc finger and SCAN domain
Ubtfl1	Hmgpi; B020006M18Rik	546118	upstream binding transcription factor
Zar1	---	317755	oocyte-to-embryo transition
Nobox	OG2; Og2x	18291	oogenesis homeobox
Padi6	Pad6; ePAD; Padi5	242726	peptidyl arginine deiminase
Khdc3	FILIA; OEEP48; AI467128; 2410004A20Rik	66991	subcortical maternal complex
Sebox	OG9; Og9x	18292	homeobox protein
Nlrp5	Op1; Mater; Nalp5; PAN11	23968	NACHT and pyrin domain

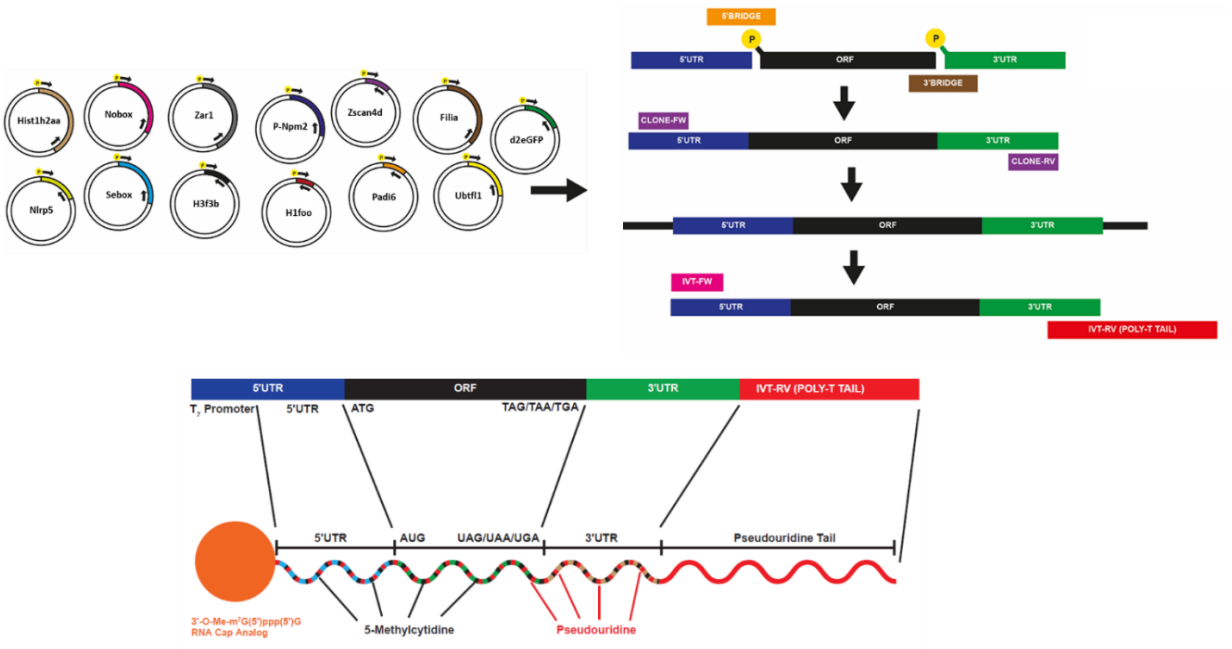


Figure 12: Diagram showing the simplified overview of candidate genes' processes to become reprogrammable synthetic modified mRNAs. All candidate ORFs were first sequence verified and confirmed by the NCBI gene database. All candidate genes have the same 5' cap, UTR sequences, poly (T) lengths, and mRNA nucleotide base variants. The 12 candidates were Hist1h2aa, Nlrp5, Nobox, Sebox, Zar1, H3f3b, p-Npm2, H1foo, Zscan4d, Padi6, Filia, and Ubtfl1.

Every synthetic modified mRNAs are confirmed by Bioanalyzer prior to transfection. Protein expressions from mRNAs of select candidate genes are shown in Figure 13. Others could not be stained due to unavailability of antibodies or lack of antibodies that can distinguish the variants from the canonicals. All synthetic modified mRNAs have the identical parts except the ORF, therefore, the regions that are involved in protein expressions are same for all.

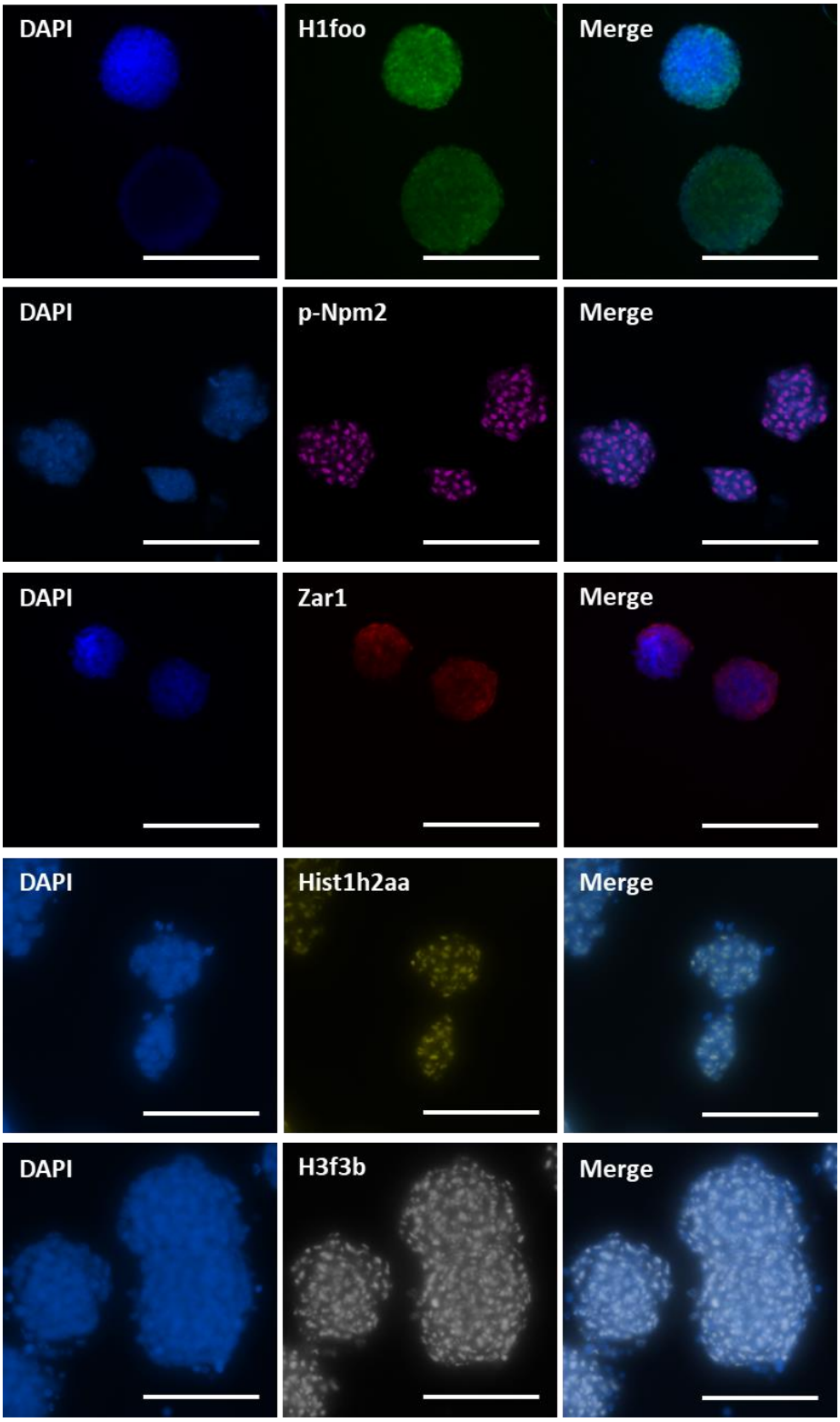

DAPI
H1foo
Merge
DAPI
p-Npm2
Merge
DAPI
Zar1
Merge
DAPI
Hist1h2aa
Merge
DAPI
H3f3b
Merge

Figure 13: Representative images of protein expressions from candidate genes' mRNAs. Synthetic modified mRNAs were derived from candidate genes and were transfected daily. Protein expressions were confirmed by immunohistochemistry staining. Scale bars, 100 µm.

Culture condition is an essential component to the success of cell reprogramming. Culture optimization was performed by culturing ESCs for an extended period using commonly used ESC media and 2C+ were quantified. As shown in Figure 14a, two culture conditions showed statistically higher 2C+; serum with LIF (serum/L) and knockout serum-replacement with LIF (KOSR/L). Despite KOSR/L culture condition resulting in the highest 2C+, there was no statistical difference between KOSR/L and serum/L. Despite the statistically significant differences, the absolute 2C+ changes observed were in the ranges of a tenth of a percent, indicating no meaningful differences.

Next, we began testing the 12 candidate genes. We prepared modified-mRNA cocktail by gently mixing each of the factors in equimolar ratio and transfected them daily. By day five we observed noticeable separation amongst different culture conditions for 2C+ as shown in Figure 14b. Continuously generating synthetic modified mRNAs from 12 genes require an immense amount of time and labor. Therefore, we decided to narrow down our list of candidates. We have prepared various combinations of mRNA cocktails which we named cocktail 1 to 9 also known as C1-9. Out of all combinations, we found C1 and C2 to induce that highest 2C+. Not surprisingly, these two combinations also had the second largest numbers of factors with six in each. The further reduction resulted in a rapid decrease in 2C+ as shown in Figure 14c. One thing to note is that we found C3 group to be comparable in its ability to induce 2C+. However, as seen in Figure 14c, there was even greater spread in 2C responses from each experiment. Considering that C3 has four genes, which require comparable amount of work as six genes, we opted to go with C1 instead. We have noticed that C1 to be the most consistent in contrast to C2 or C3 in inducing 2C+, therefore, C1 has been used for all future experiments.

We wanted to verify that the changes in 2C+ are truly due to our candidate genes so we transfected at varying doses of C1 which showed dose-dependent 2C+ as shown in Figure 14d. Furthermore, this showed that the optimal dose to be around 1 µg since there were no statistical differences in 2C responses beyond 1 µg range.

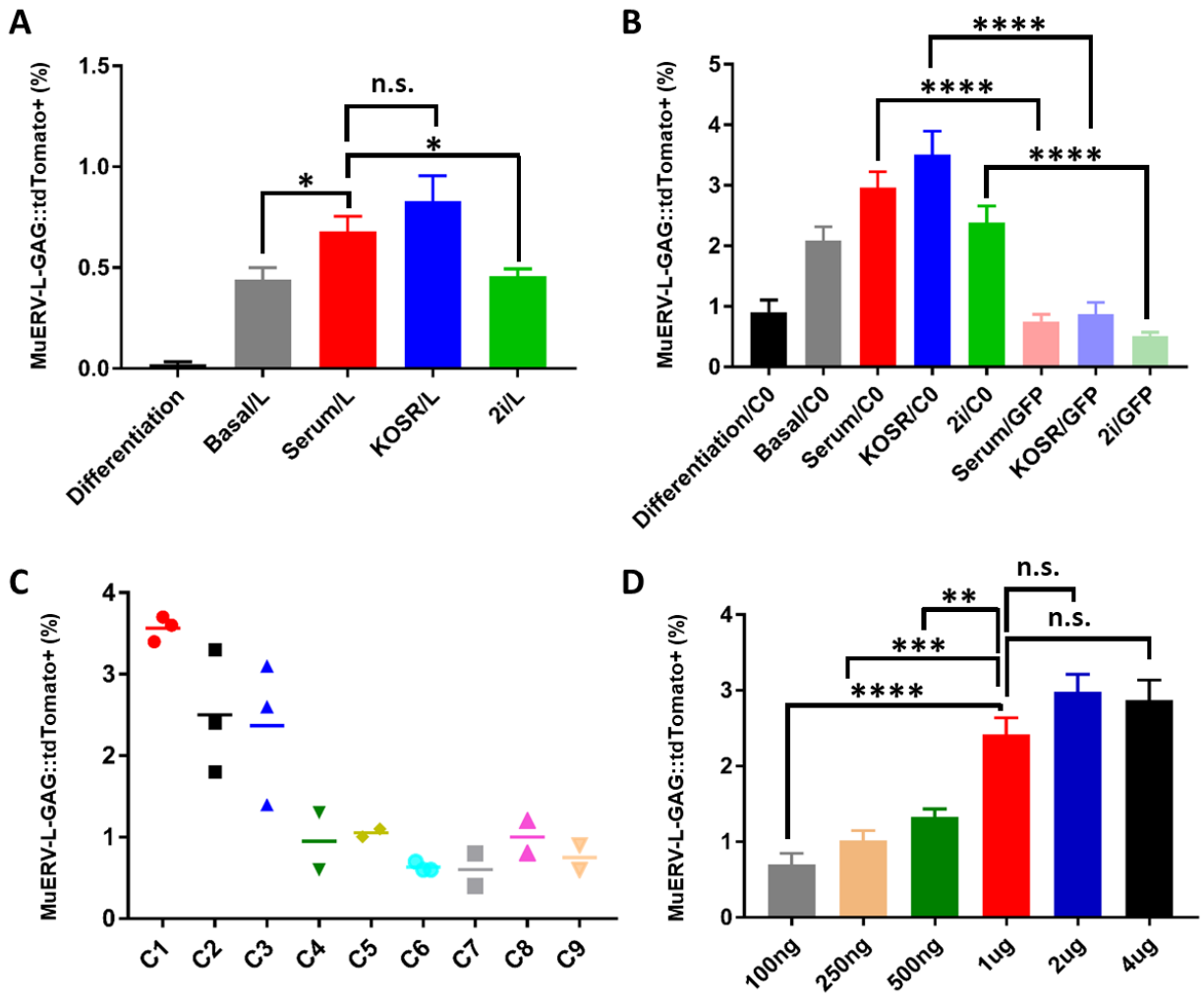


Figure 14: Totipotent-like state can be induced by overexpressing key genes under proper culture conditions. (A) Culture optimization. Common media used in ESCs have been tested. Basal = proprietary differentiation ESC media from Millipore, KOSR = knockout serum replacement, 2i = GSK3β and MEK1/2 inhibitors, L = LIF. Error bars, S.E.M., n = 10. (B) Transfection of mRNA cocktail of candidate genes in

different media conditions. C0 = Hist1h2aa, H3f3b, H1foo, p-Npm2, Zscan4d, Ubtfl1, Zar1, Nobox, Padi6, Khdc3, Sebox, Nlrp5, GFP = only GFP mRNA, LIF was included in all culture conditions except the differentiation media. Error bars, S.E.M., n = 6. (C) Candidate reduction. Working list of candidate genes were further narrowed. C1 = Hist1h2aa, H3f3b, H1foo, p-Npm2, Zscan4d, Ubtfl1, C2 = Zar1, Nobox, Padi6, Khdc3, Sebox, Nlrp5, C3 = Hist1h2aa, H3f3b, H1foo, p-Npm2, C4 = Hist1h2aa, H3f3b, H1foo, C5 = p-Npm2, Zscan4d, Ubtfl1, C6 = Hist1h2aa, p-Npm2, C7 = H3f3b, H1foo, C8 = Zscan4d, Ubtfl1, C9 = Hist1h2aa, H3f3b. (D) Candidate genes have dose response effect in inducing 2C+. Increasing the dose of C1 corresponded to increasing 2C+. There was no statistical differences in 2C+ beyond 1 μg C1. Error bars = S.E.M., n = 6. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$, n.s. = not significant. All P-values were calculated on a basis of a two-tailed unpaired *t*-test.

ESCs were transfected daily with C1 mRNA and quantified for 2C+ every two days. As shown in Figure 15, C1 alone was insufficient to induce the majority of cells towards 2C+. To improve this, exogenous chemicals were also used. Histone deacetylases remove acetyl groups from histones and expose the positive charges that then interact with the net negative charges of the DNA. Thus, histones wrap the DNA more tightly resulting in heterochromatin state. The opposite, more permissive state, occurs during ZGA, therefore, we reasoned that HDAC inhibitors might able to direct cells toward totipotency.[17,75,133,157,164] Trichostatin-A (TSA) is a known HDAC inhibitor that has shown to improve the success rates of SCNT.[148,150,168] When used alone, TSA was able to induce 2C+ albeit at the cost of noticeable cell deaths. Interestingly, when combined with C1, there was an even higher 2C+ increase.

Encouraged by the result, we sought to find another exogenous molecule that can modulate the open chromatin state but through a different mechanism. Previous studies have shown that chromatin assembly factor-1 (CAF1) plays crucial roles in maintaining heterochromatin. As shown in Figure 15, when CAF1 was repressed with siRNA p150, a major subunit of CAF1, increase in 2C+ was observed. When

combined with C1, even higher 2C+ was observed. Combining all three—C1, TSA, p150 siRNA—resulted in the greatest number of 2C+. This, however, accompanied by significantly higher cell deaths. This was especially noticeable in serum-free media. Serum appears to keep the cells alive longer, therefore, serum/L media was used for rest of the experiments unless stated otherwise. Furthermore, the continuous TSA dose was lowered to 10 nM and the p150 siRNA was transfected every third day. Altogether, we were able to continually reprogram for 10 days and increase 2C+ from approximately 0.5% to 50%.

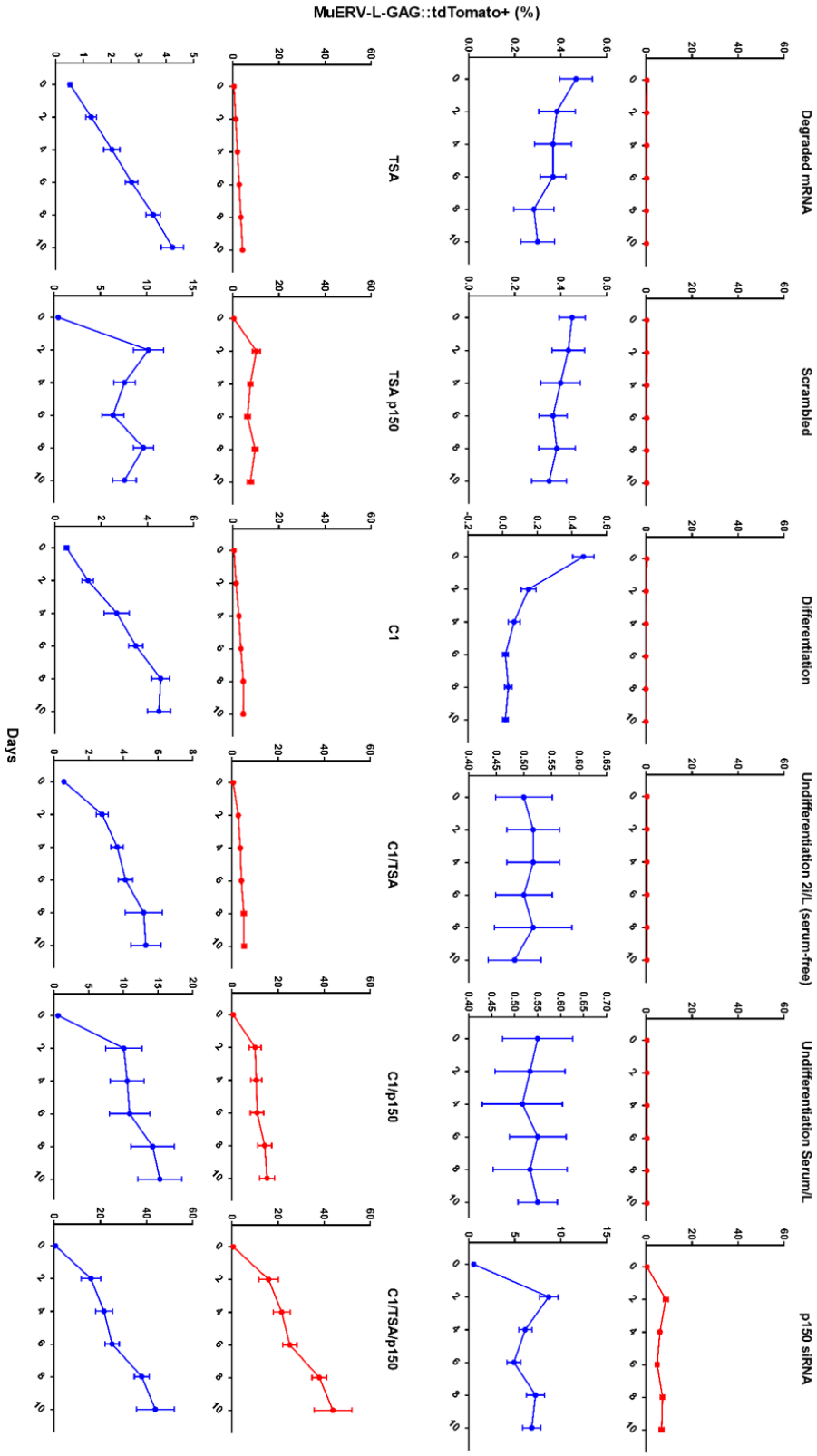

MuERV-L-GAG::tdTomato+ (%)
Degraded mRNA
Scrambled
Differentiation
Undifferentiation 2i/L (serum-free)
Undifferentiation Serum/L
p150 siRNA
TSA
TSA p150
C1
C1/TSA
C1/p150
C1/TSA/p150
Days

Figure 15: Candidate genes in C1 and exogenous chemicals have a synergistic effect in inducing a 2-cell-like state in ESCs. Time course FACS measurements of tdTomato expressions of various conditions are shown. Unless specifically stated, cells were cultured in serum/L. The continuous increase in 2C was observed over ten-day period for C1, TSA, and p150. Combining one with another resulted in higher 2C than when used alone and combining all three resulted in the highest 2C+. The top panels of each (red lines) show the standardized view across all conditions while the bottom panel (blue lines) show much closer view. C1 = Hist1h2aa, H3f3b, H1foo, p-Npm2, Zscan4d, Ubtfl1, TSA = Trichostatin A, p150 = p150 siRNA. Error bars, S.E.M., n = 6.

The typical morphology of ESC colony with defined border and tightly-packed cells began to change over the course of reprogramming as shown in Figure 16. FACS trace diagram shows the conversion of 2C- to 2C+ over time and clear separation between 2C+ and 2C- was seen by day 10, indicating the emergence of the distinct cell population.

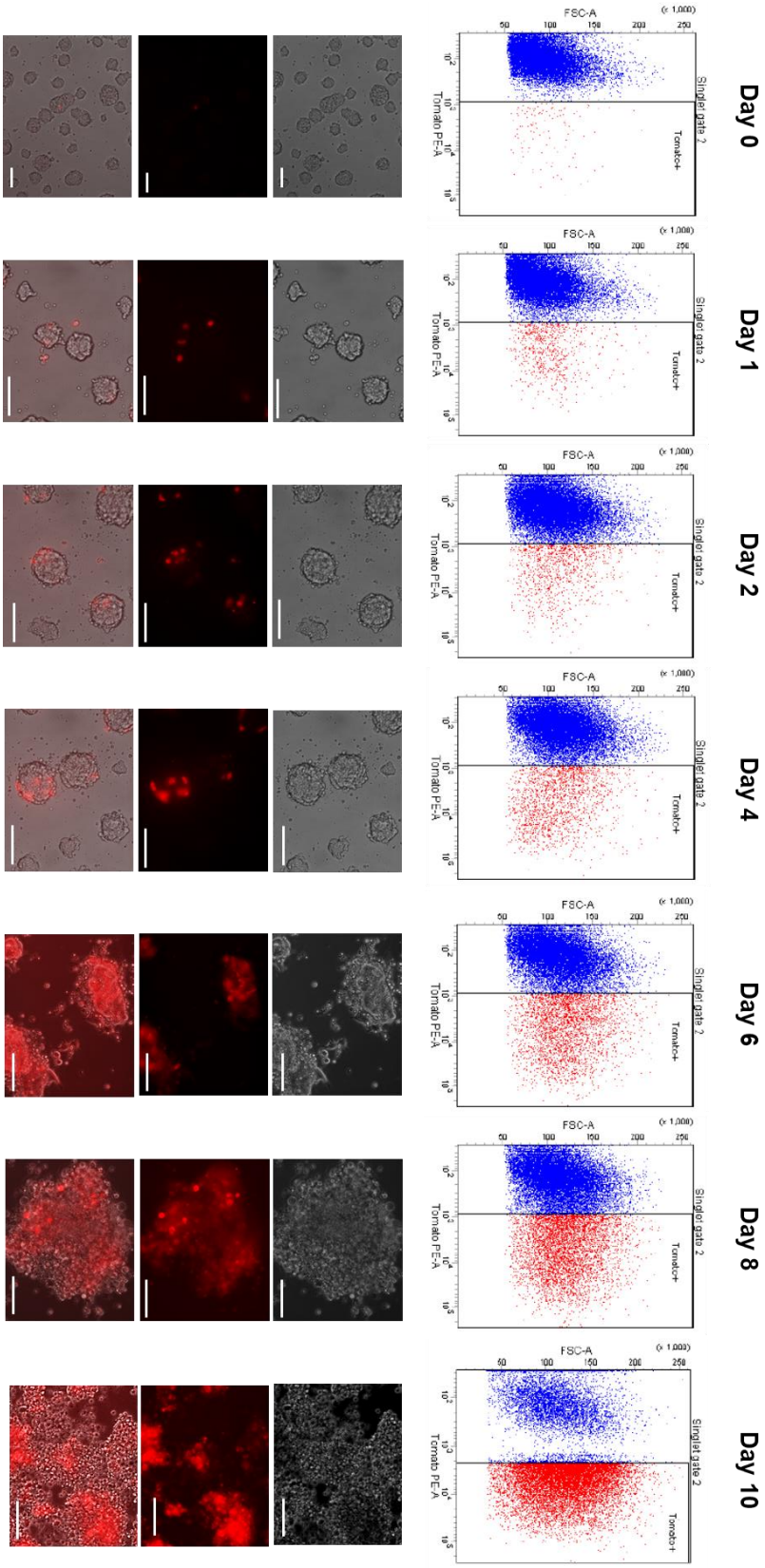

Day 0
Day 1
Day 2
Day 4
Day 6
Day 8
Day 10
FSC-A
(x 1,000)
Singlet gate 2
Tomato+
Tomato PE-A

Figure 16: Changes over the course of reprogramming. Flow trace diagrams show cells acquiring 2C+ resulting in two distinct populations by day ten. The loss of typical ESC morphologies was observed primarily in the 2C+ enriched areas. Scale bars, 100 µm.

Advanced donor age and prolonged cell culture have shown to reduce the efficiency and fidelity of reprogramming.[41,169,170] We wanted to investigate whether 2C induction is also dependent on the age of the cell. We compared our reprogramming results of two different passage numbers. As shown in Figure 17, there were no statistical differences between the two culture groups, indicating the robustness of this reprogramming.

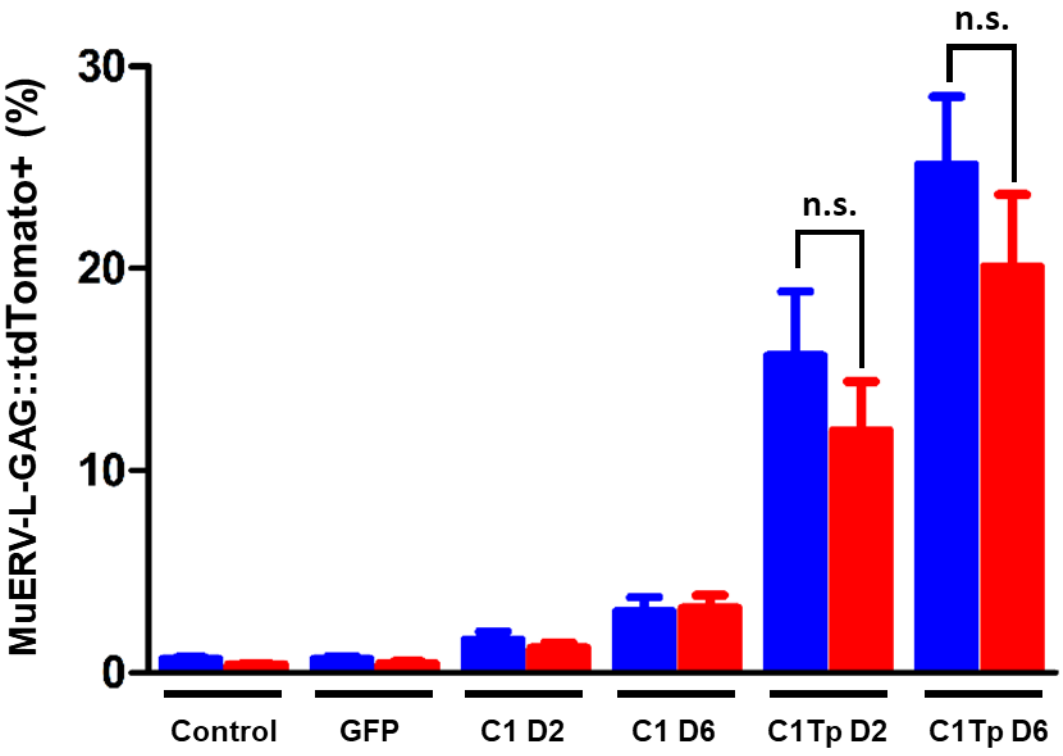


Figure 17: Candidate genes and chemicals can induce 2-cell-like state independent of the age of cells. Earlier passage (p6, blue) and later passage (p34, red) cell cultures showed similar levels of increase for 2C+ in response to reprogramming factors. C1 = Hist1h2aa, H3f3b, H1foo, p-Npm2, Zscan4d, Ubtfl1, T =

TSA, p = p150 siRNA. Error bars = S.E.M., n = 4, n.s. = not significant on the basis of a two-tailed, unpaired *t*-test.

Totipotent-associated genes, maternal factors, in particular, can induce a 2-cell-like state in a dose- and temporal-dependent manner. Cell culture conditions also play a role. Interestingly, we expected more 2C+ on 2i/L medium. However we found serum/L and KOSR/L to be statistically higher. This challenges a previous finding which claims that totipotent cells are more likely to arise from the ground-state culture condition, 2i/L, than from the standard culture condition, serum/L.[171] Despite this differing opinion, in our case, serum/L was the only viable option since serum-free conditions were unable to maintain enough cell viability throughout the reprogramming process. The higher increase in 2C+ from the combined C1, TSA, and p150 siRNA was more than additive. This synergistic effect could indicate that the mechanisms by which the factors of the C1 act to induce 2C state may differ from that of the exogenous chemicals. Aged cells upregulate Ink4/Arf which has been shown to reduce reprogramming efficiency.[172] Our reprogramming did not show age difference, suggesting Ink4/Arf inhibition as a possible mechanism.

VI.

Totipotent Traits

"I love fools' experiments. I am always making them." – Charles Darwin

One hallmark of totipotency is the activation of the MuERV-L element, which showed significant increase during the reprogramming process. We sought out to determine whether these reprogrammed cells possess additional traits of totipotency.

To determine the direction of cell reprogramming—dedifferentiation, no change, or differentiation—genes that are expressed at different stages of cell lineage were quantified using RT-qPCR on reprogrammed 2C+ cells collected by FACS. As shown in Figure 18, MERVL long terminal repeat and major satellites were highly upregulated in reprogrammed cells when compared against ESCs. In contrast, IAP and LINE1 repetitive elements were largely unaffected, and C1, in particular, did not show statistically significant differences to that of control ESCs. This suggests that the reprogramming did not globally affect all repetitive elements indiscriminately. Furthermore, 2-cell-specific totipotent genes all showed a

significant increase in reprogrammed cells. These results indicate that the reprogrammed cells appear to be dedifferentiating towards totipotency. Next, we determined whether this direction is specific towards totipotent state by analyzing other cell lineages.

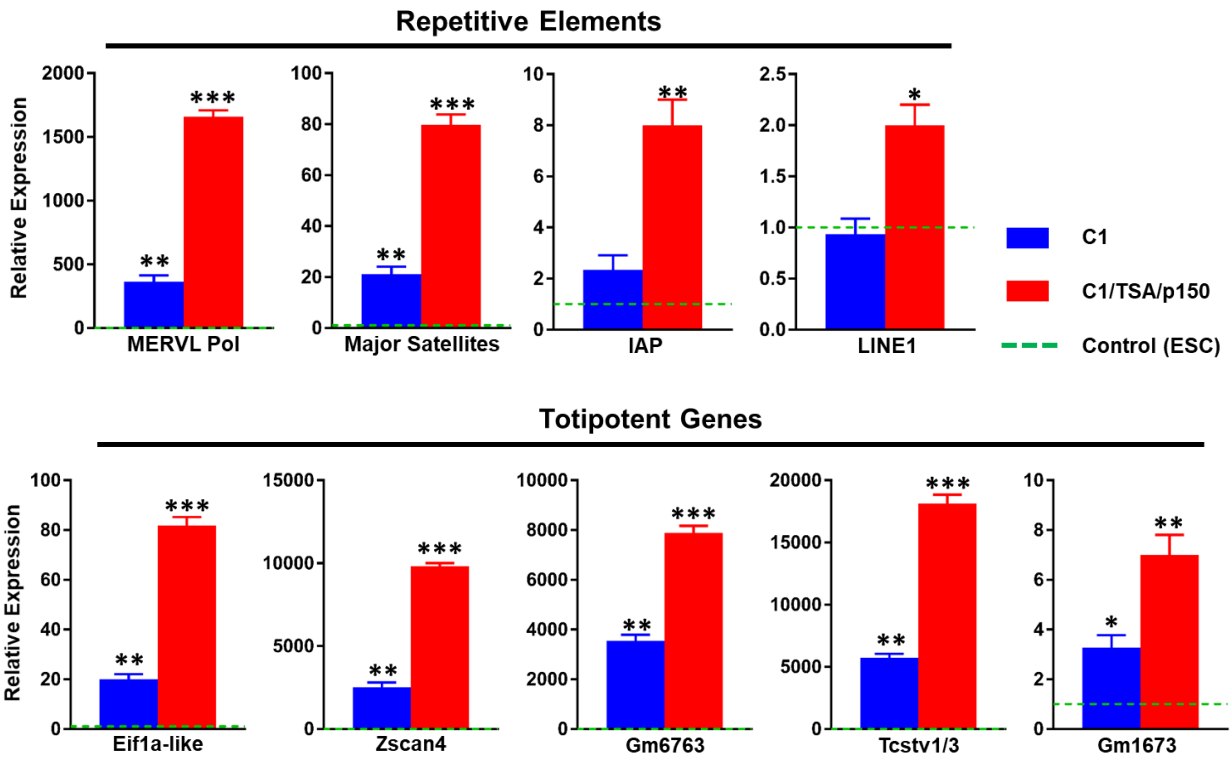


Figure 18: Totipotent genes are upregulated in reprogrammed cells. The relative expressions of the indicated repetitive elements and totipotent genes were determined using RT-qPCR. Normalized to GAPDH. Blue = C1, red = C1, TSA, p150 siRNA, green = ESC. C1 = Hist1h2aa, H3f3b, H1foo, p-Npm2, Zscan4d, Ubtfl1. Error bars, S.D., n = 3. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. All P-values were calculated on the basis of a two-tailed, unpaired t-test, compared to control ESC.

A complete change to a different cell, whether through reprogramming or differentiation, often involves relinquishing the defining traits of the starting cell type. We wanted to determine whether our reprogrammed cells are acquiring totipotent-specific expressions while still maintaining their pluripotency, which would indicate partially reprogrammed state. While some genes do overlap between

totipotency and pluripotency, key pluripotent genes are expressed shortly after totipotency. For example, a key pluripotent-defining gene, Oct4 mRNA and protein are first expressed in 4-cell and 8-cell, respectively, in mouse embryos, indicating that pluripotency follows the immediate aftermath of totipotency.[173,174] Other mammalian species, including human, also show that Oct4 proteins are first detected in 8-cell stage immediately after totipotency, highlighting remarkably precise molecular distinction between these two seemingly connected cell types.[175-177] Interestingly, cdx2 protein expression has also been detected as early as 8-cell stage which indicates the point where two distinct cell populations begin to arise from one.[178] This suggests that totipotency begins at 1-cell, peaks at 2-cell, and falls at 4-cell, with pluripotent pathways being established immediately afterward. As shown in Figure 19, pluripotent pathway genes, including Oct4, are downregulated in reprogrammed cells. The loss of pluripotent genes shows that these reprogrammed cells are undergoing committed change.

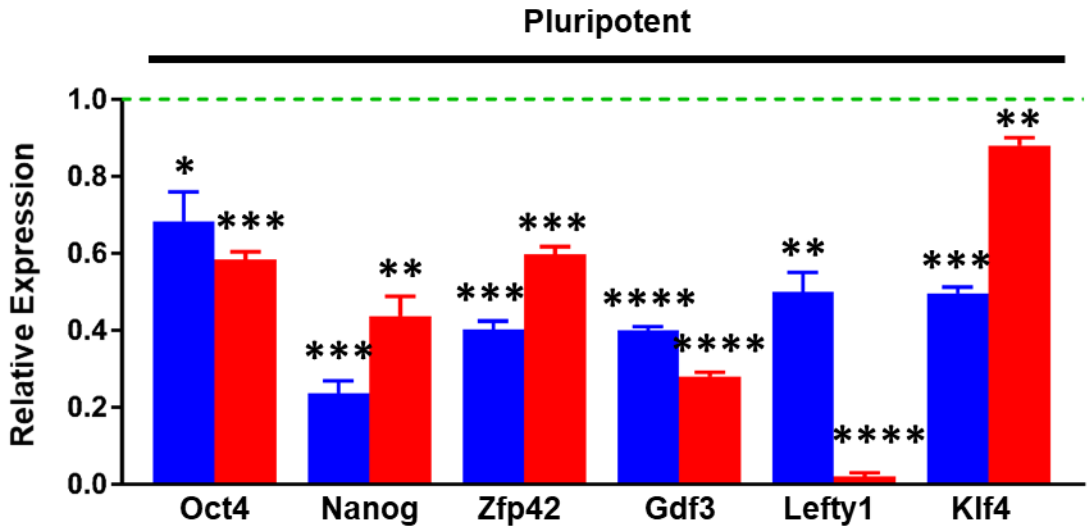
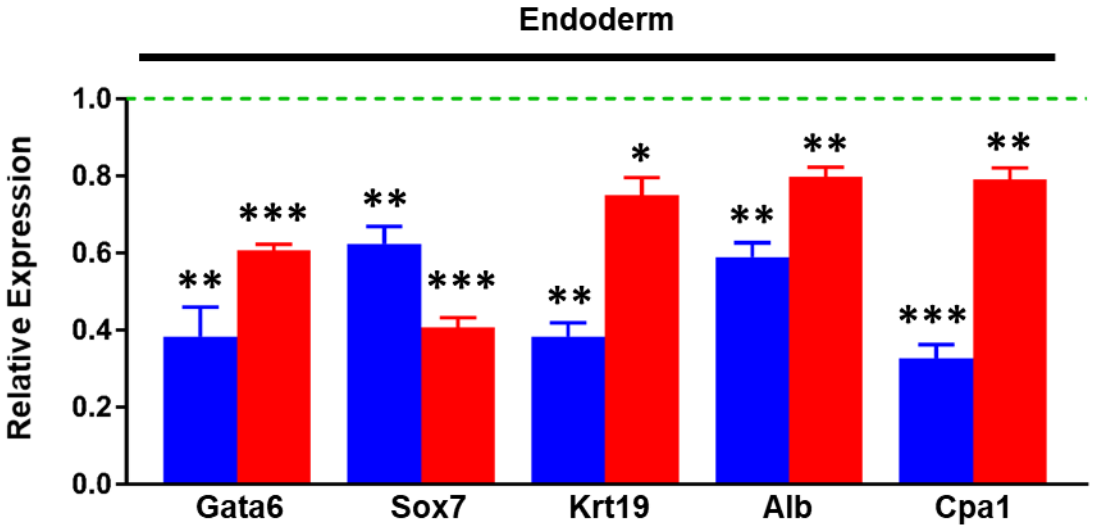


Figure 19: Pluripotent genes are downregulated in reprogrammed cells. The relative expressions of the indicated pluripotent genes were determined using RT-qPCR. Normalized to GAPDH. Blue = C1, red = C1, TSA, p150 siRNA, green = ESC. C1 = Hist1h2aa, H3f3b, H1foo, p-Npm2, Zscan4d, Ubtfl1. Error bars, S.D., n

= 3. * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001. All P-values were calculated on the basis of a two-tailed, unpaired t-test, compared to control ESC.

Lastly, we wanted to determine if this committed change is unidirectional only towards totipotency. To capture all three major differentiation lineages, expression levels of genes that are commonly upregulated at each germ layer have been determined. As shown in Figure 20, genes that define each germ layer are all downregulated in reprogrammed cells. Altogether, the reprogrammed cells are undergoing committed unidirectional change towards totipotency. Hence, we will refer these reprogrammed cells as induced totipotent-like cells (iTLCs).



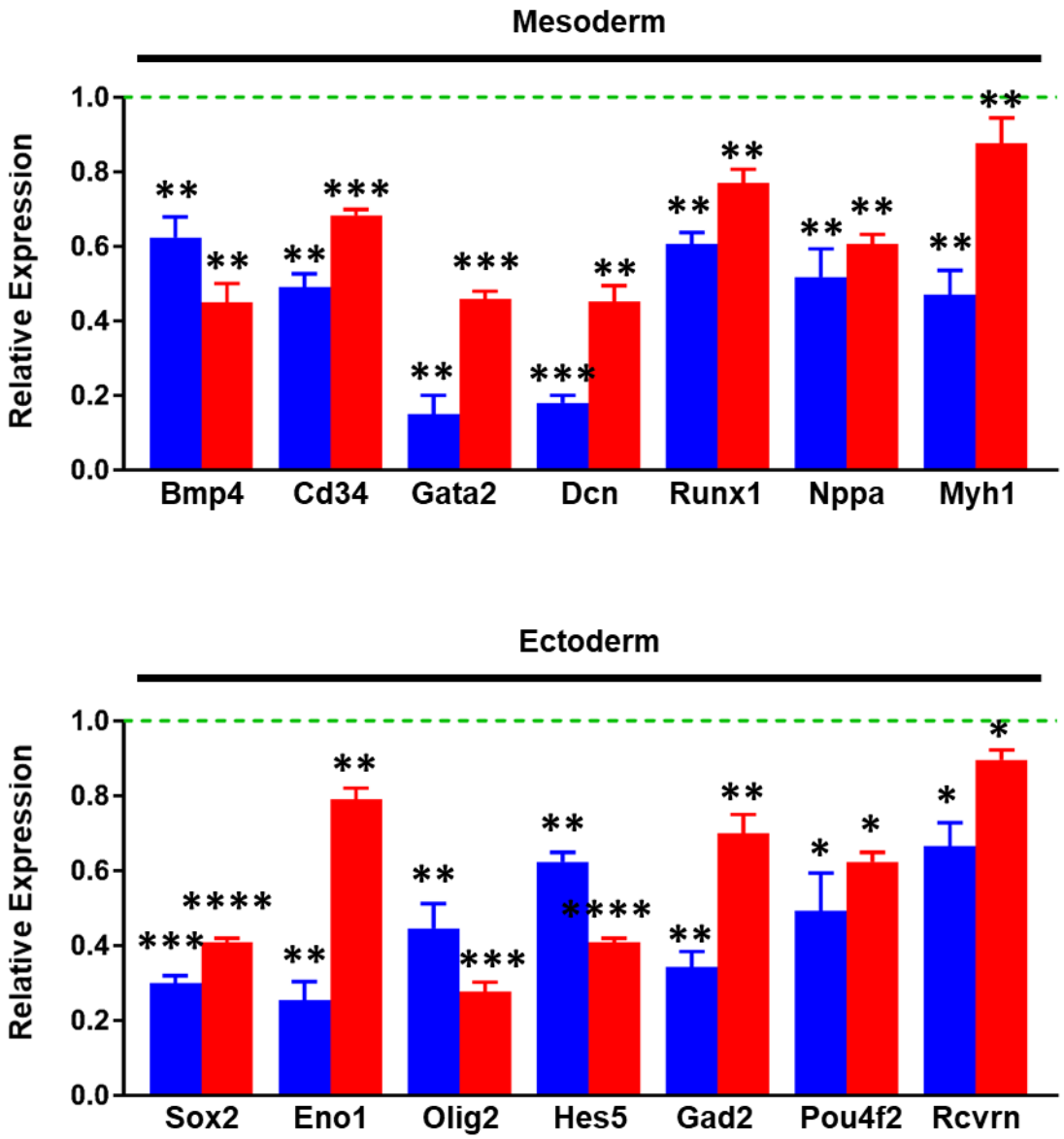


Figure 20: Defining genes from each germ layer are all downregulated in reprogrammed cells. The relative expressions of the indicated germ layer genes were determined using RT-qPCR. Normalized to GAPDH. Blue = C1, red = C1, TSA, p150 siRNA, green = ESC. C1 = Hist1h2aa, H3f3b, H1foo, p-Npm2, Zscan4d, Ubtfl1. Error bars, S.D., n = 3. * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001. All P-values were calculated on a basis of a two-tailed, unpaired t-test, compared to control ESC.

Interestingly, we have also found that the reprogrammed cells exhibit some markers of an oocyte. As shown in Figure 21, genes that are expressed in an oocyte are also upregulated. An oocyte is the only cell that can induce totipotency. One can make the case that based on SCNT, the only difference between an oocyte and a zygote is the number of chromosomes. It is yet unclear if some of the reprogrammed cells are moving more towards oocytes rather than towards zygotes, the transcript distinctions are very subtle due to overlapping developmental period. Karyotype analysis would address whether some of our cells have undergone successful meiosis which would qualify them as oocytes rather than zygotes. Regardless, this highlights the intertwined relationship between the oocytes and totipotency.

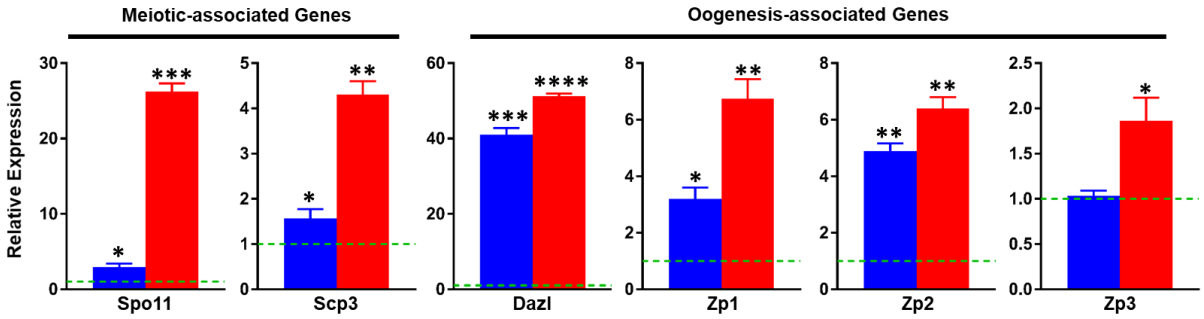


Figure 21: Reprogrammed cells show upregulation of genes that are normally present in oocytes. The relative expressions of the indicated meiotic and oogenesis genes were determined using RT-qPCR. Normalized to GAPDH. Blue = C1, red = C1, TSA, p150 siRNA, green = ESC. C1 = Hist1h2aa, H3f3b, H1foo, p-Npm2, Zscan4d, Ubtfl1. Error bars, S.D., n = 3. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$. All P-values were calculated on a basis of a two-tailed, unpaired t-test, compared to control ESC.

We wanted to assess other features of totipotency from a different angle. Totipotency of a zygote is associated with highly loosened chromatin structure, and the nuclear size progressively shrinks from the zygote to the blastocyst stage.[179,180] Furthermore, increasing the nuclear-to-cytoplasm ratio has shown to activate zygotic gene transcription and the decrease in the ratio delays the zygotic transcription.[181] We

sought out to determine whether reprogrammed iTLCs also show similar nuclear change. As shown in Figure 22, the nuclei of iTLCs are significantly enlarged in size, further displaying totipotent trait.

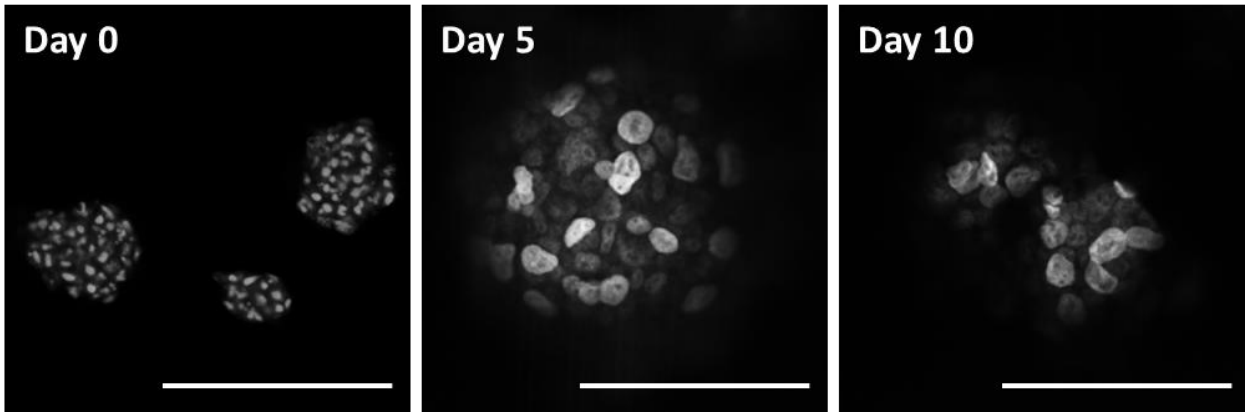
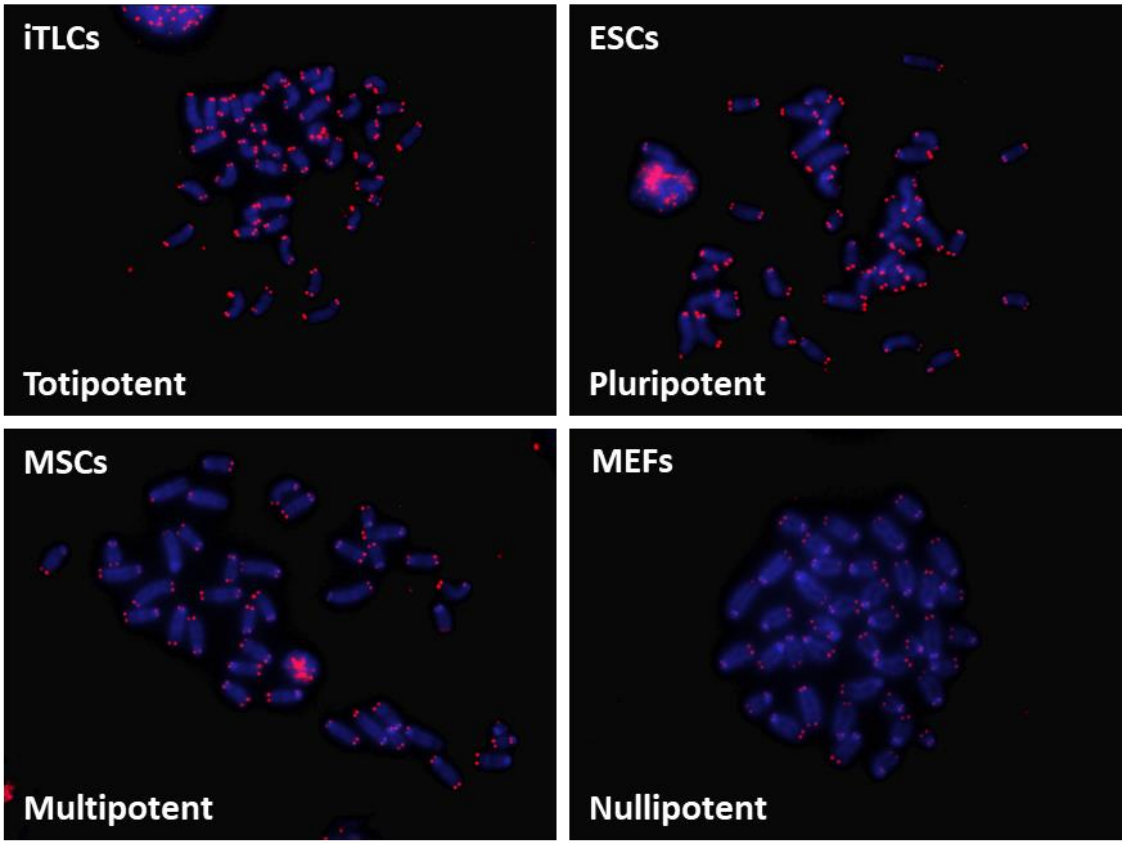
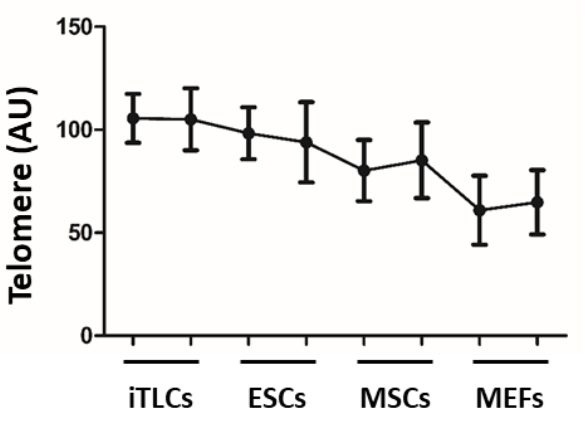


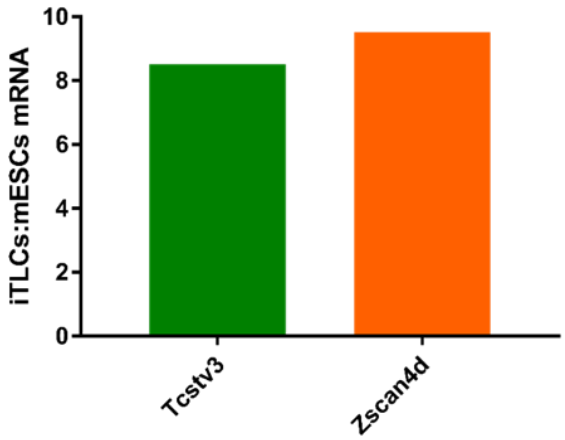
Figure 22: Reprogrammed cells have enlarged nuclei. Nuclear stainings show relative nuclei changes in response to C1/TSA/p150. Control ESCs, left, show nuclei of approximately 1-3 μm in diameters. Day 5, middle, and Day 10, right, of reprogramming show nuclei of approximately 10-15 μm in diameters. Scale bars, 100 μm.

Telomere shortening has been associated with aging.[182] Similarly, cellular aging also follows telomere length. Terminally differentiated somatic cells have diminished telomeres due to no telomerase activity. Adult stem cells have long telomeres due to low levels of telomerase activity.[182] Pluripotent stem cells have even longer telomeres due to higher levels of telomerase activity.[10,183,184] It is unclear whether similar pattern holds towards totipotent cells. Telomere length has been shown to define a cell's differentiation potential.[185] Therefore, totipotent cells, having the highest differentiation potential, would, in theory, have the longest telomere. Not surprisingly, telomere lengths have shown to remarkably increased in one- to two-cell stage embryos in mice.[186] Interested by this, as shown in Figure 23, quantitative fluorescence in situ hybridization (q-FISH) of telomere lengths in reprogrammed cells was performed. After controlling for the duration of cell culture, iTLCs showed greater telomere lengths than other compared cell types.

Upon further investigation, we found that Tcstv3 and Zscan4, predominately expressed in 2-cell embryos, are significantly upregulated and that these have been shown to elongate telomeres.[151,187,188]







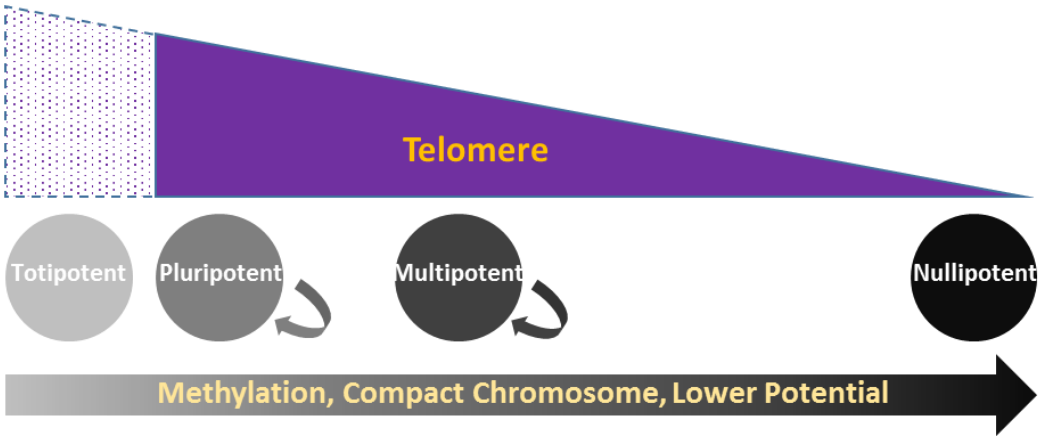


Figure 23: Reprogramming towards totipotency elongates telomere. Top: Telomere FISH images of induced totipotent-like cells, embryonic stem cells, mesenchymal stem cells, and mouse embryonic fibroblasts. Mid-left: Quantification of telomere FISH from two independent samples of each group show corresponding decline in telomere with cell potential. Mid-right: Tcstv3 and Zscan4 are upregulated in iTLCs. Bottom: Proposed model of telomere and cell potential.

Different cell types often require different culture media which can be used to analyze a cell's identity further. The most popular and commonly used medium for retrieving and culturing pre-implantation embryo is called Kalium (K^+) Simplex Optimized Medium (KSOM). Interestingly, ESCs cannot survive in KSOM. This prompted us to culture iTLCs in KSOM. As shown in Figure 24, rapid cell death was observed for ESCs. No cells were alive by day four. In contrast, iTLCs were able to be cultured in KSOM for much longer. This suggests that iTLCs have changed to become more compatible with totipotent culture medium. However, cell deaths were also observed. This may have been due to the eventual loss of totipotency over time.

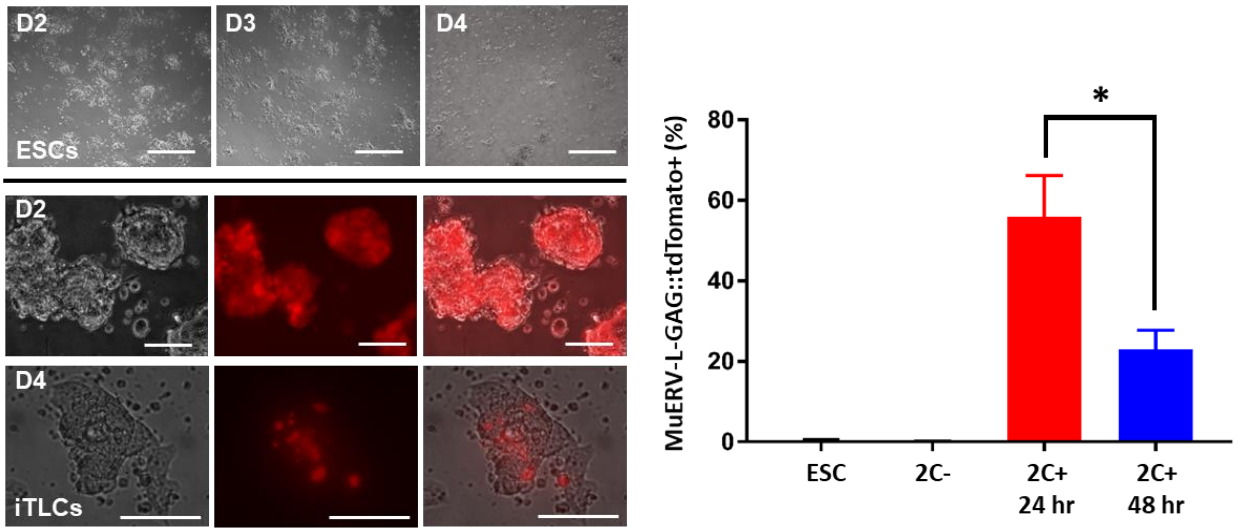


Figure 24: Comparison between ESCs and iTLCs in the totipotent medium. Left: Rapid cell deaths were observed for ESCs while iTLCs were able to be cultured for much longer. Scale bars, 100 μm. Right: The eventual loss of 2C state was observed in iTLCs. Error bars, S.E.M., n = 4. * $P < 0.05$. All P-values were calculated on the basis of a two-tailed, unpaired t-test.

The results show that iTLCs possess tantalizing traits of totipotency. One controversial finding that requires further discussion is the elongated telomere length of iTLCs. Current understanding of telomere length in mammalian embryogenesis indicate that the earliest embryos, zygotes and two-cell stage, have lower telomere length than that of pluripotent stem cells.[189] Surprisingly, telomere length in the oocyte is known to be markedly shorter than somatic cells.[186,190-192] Perhaps then the cause of shorter telomere lengths in zygotes and two-cell embryos, which are derived shortly after oocytes, are not cell intrinsic but rather are results of inheritance. Interestingly, the most significant telomere lengthening takes place following fertilization and coincide with zygotic genome activation.[186,192-194] This is perhaps to counter the relatively short telomere inheritance from the oocytes. Regardless, the telomere lengthening continues and peaks at pluripotent stem cells. Even more interesting is that during the most significant telomere

lengthening, telomerase activity, which is what pluripotent stem cells use, remains undetectable. Instead, the earliest embryos such as totipotent cells elongate by adapting to a favorable chromatin state that allows for high rates of telomere sister chromatid exchange and telomere-specific localization of recombination proteins. This suggests that the chromatin configuration during zygotic genome activation has more purpose than initially realized. Altogether, this means that inducing zygotic genome activation on pluripotent stem cells, will, in theory, result in even higher telomere length. However, this has yet to be shown until now. We believe that this explains why iTLC shows remarkably higher telomere length that differ from an oocyte-derived totipotent cell.

We have explained that the enlarged nucleus of iTLC is the result of ZGA reprogramming. However, the enlarged nucleus is also associated with a tumor cell. Our finding reinforces this concern that iTLCs can survive in KSOM much longer than what is expected from ESCs. RT-qPCR data captured the major cell types based on normal mammalian development. We have yet to address the cell that is outside of normal; the malignant cell.

VII.

Analysis of Malignancy

"An expert is a man who has made all the mistakes which can be made, in a narrow field." – Niels Bohr

ESCs and malignant cells share many similarities. Both express many of the same proto-oncogenes, divide indefinitely, and give rise to many different cell types. Placing enough stress for a prolonged period can transform ESCs into malignant cells. Our reprogramming accompanies cell deaths, indicating that this is a stressful process. It is unclear whether this stress can result in malignant cells.

We wanted to determine whether our reprogramming process affect genomic stability. The routine method to evaluate the genome stability is chromosome karyotype. As shown in Figure 25, a number of different conditions were analyzed. Cytogenetic analysis was performed on twenty G-banded metaphase cells from each condition. First, we established the baseline by analyzing the non-reprogrammed control ESCs. The control analysis showed a majority of cells with normal karyotypes with a few cells demonstrating non-clonal chromosome aberrations that were determined to be artifacts of culture. Then,

we reprogrammed using either C1 or C1/TSA/p150; both resulted in the majority of cells with normal karyotypes. Once we established this, we reprogrammed with C1/TSA/p150 for a prolonged period, which showed similar analysis as the control with a majority of cells demonstrating normal karyotypes. To exclude the possibility that this is unique to our cells, we analyzed a different sex from a different cell line, which also showed normal karyotypes. Altogether, the reprogramming process does not appear to affect genome stability negatively.

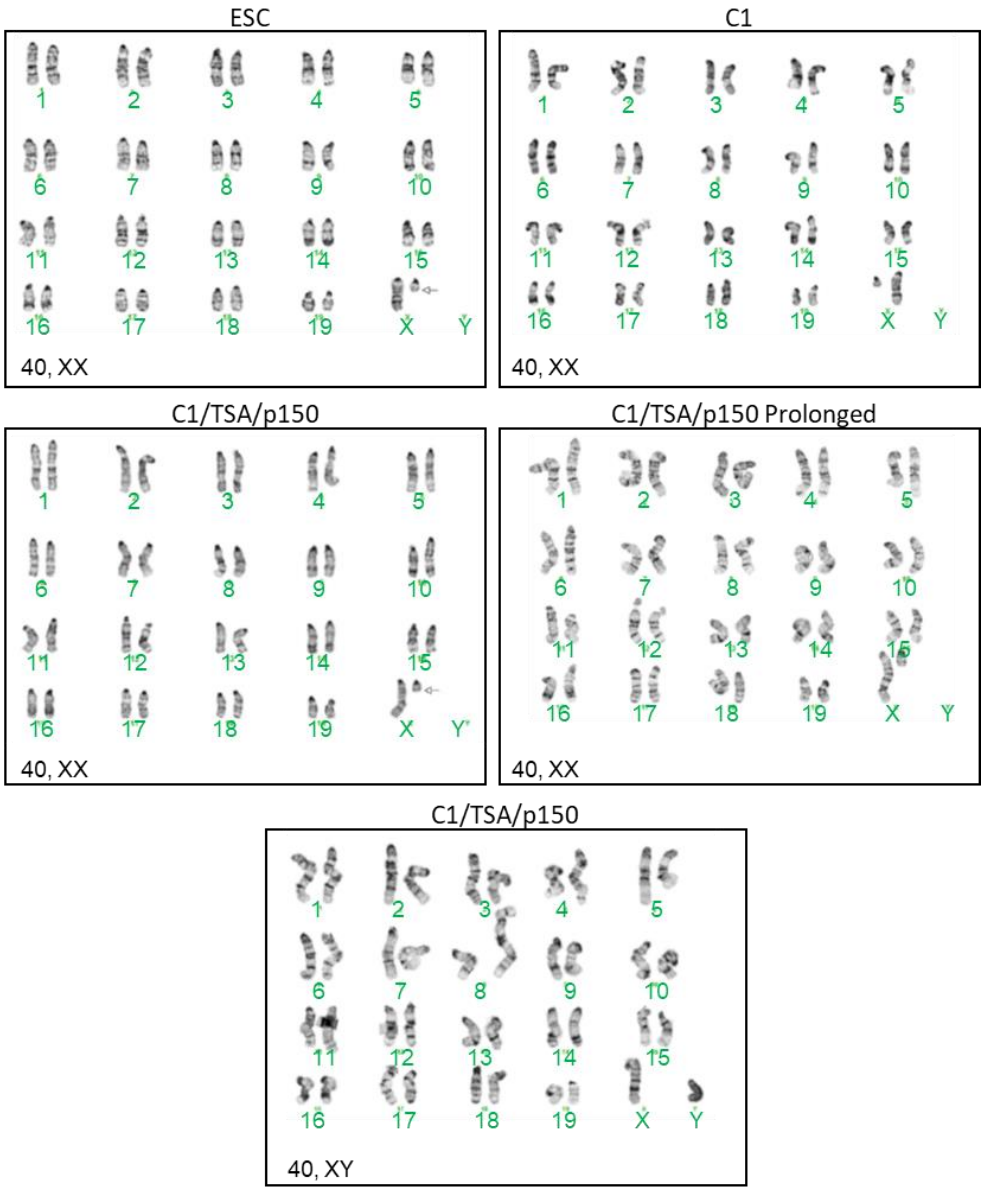
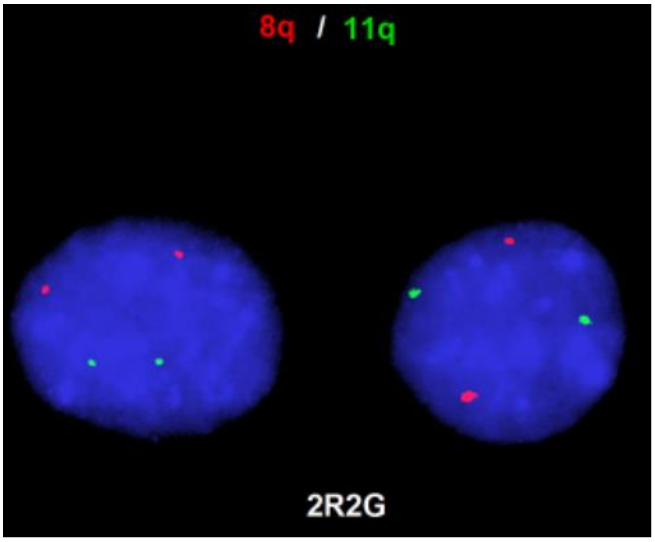
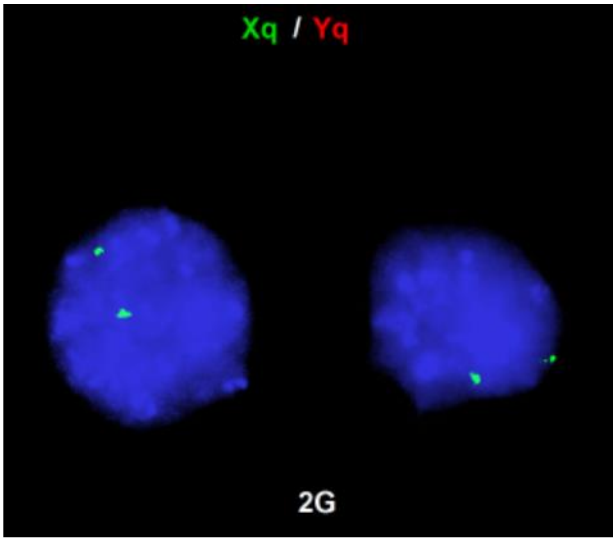


Figure 25: Karyotype analysis. ESC (top left) represents the control. C1 (top right) represents the reprogramming with C1 only. C1/TSA/p150 XX (mid left) represents the earlier time point of the reprogramming. C1/TSA/p150 Prolonged (mid-right) represents the later time point of the reprogramming. C1/TSA/p150 XY (bottom) represents the reprogramming of a different sex from a different cell line.

To further verify the genomic stability in more detail, molecular cytogenetic was performed. Two hundred interphase nuclei per probe set were examined by FISH using the probes that detect changes in chromosome 8, 11, X, and Y copy number, which are the most common genomic aberrations in mouse ESC cultures. We have analyzed C1/TSA/p150 prolonged (Figure 26A) and C1/TSA/p150 of a different cell line (Figure 26B). Both analyses show that majority of cells have diploid copy numbers of chromosomes 8, 11, and correct numbers of sex chromosomes. Altogether, this further validates that the genome stability of iTLCs appears to be unaffected by the reprogramming process.





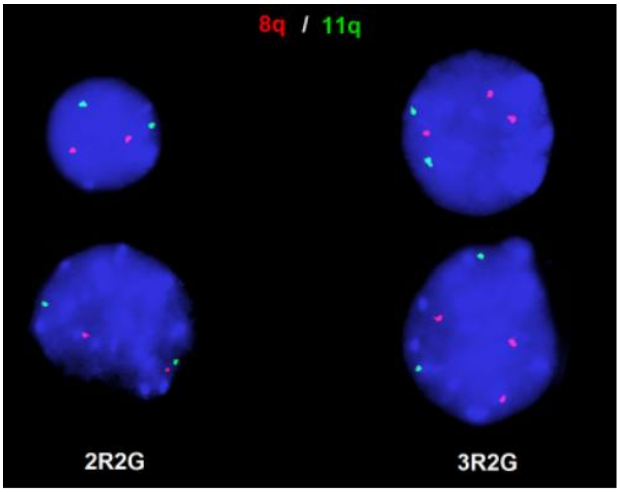
Probe Set: **8E2 (Red) / 11E2 (Green)**

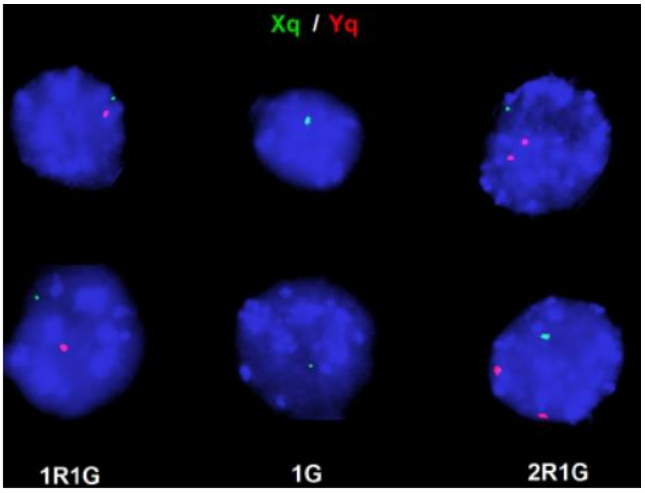
Signal Pattern	Observed % of cells w/signal pattern (Your Results)
2R/2G	95.5%
1R/2G	0.5%
2R/1G	0.5%
3R/2G	3.0%
2R/3G	0.5%

Probe Set: **XF4 (Green) / YA1 (Red)**

Signal Pattern	Observed % of cells w/signal pattern (Your Results)
0R/2G	97.5%
1R/2G	1.0%
0R/3G	1.5%

Figure 26A: FISH analysis of C1/TSA/p150 Prolonged, XX. Probe sets, 8E2 / 11E2 and XF4 / YA1, were used to detect changes in chromosome 8, 11, X, and Y copy number. FISH cut off values have not been determined for these probe sets. Therefore, low signal patterns may represent technical artifacts.





Probe Set: **8E2 (Red) / 11E2 (Green)**

Signal Pattern	Observed % of cells w/signal pattern (Your Results)
2R/2G	90.0%
3R/2G	9.0%
4R/2G	0.5%
2R/3G	0.5%

Probe Set: **XF4 (Green) / YA1 (Red)**

Signal Pattern	Observed % of cells w/signal pattern (Your Results)
1R/1G	92.5%
0R/1G	5.0%
2R/1G	2.0%
1R/0G	0.5%

Figure 26B: FISH analysis of C1/TSA/p150, XY. Probe sets, 8E2 / 11E2 and XF4 / YA1, were used to detect changes in chromosome 8, 11, X, and Y copy number. FISH cut off values have not been determined for these probe sets. Therefore, low signal patterns may represent technical artifacts.

Next, we wanted to assess the functional changes that are commonly present in malignant cells. Malignant cells are often referred to as immortal and have shown to survive under hypoxic and serum depleted conditions.[195,196] We have cultured an equal number of iTLCs in either normal serum (15%) or low serum (1%) and followed cell proliferation over time. As shown in Figure 27, iTLCs proliferated well under normal serum but failed to grow under low serum. By day three, no cells were alive in low serum condition. This suggests that iTLCs do not share the key feature of malignant cells.

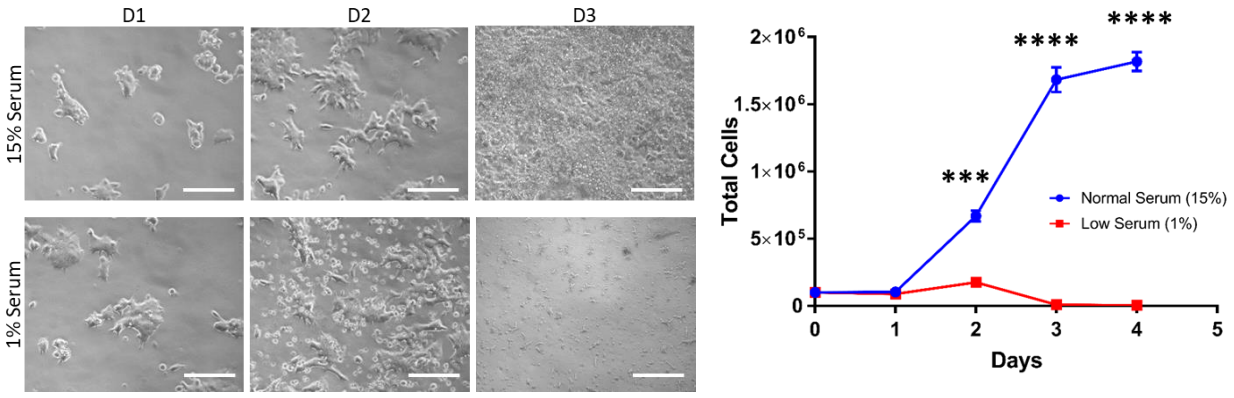
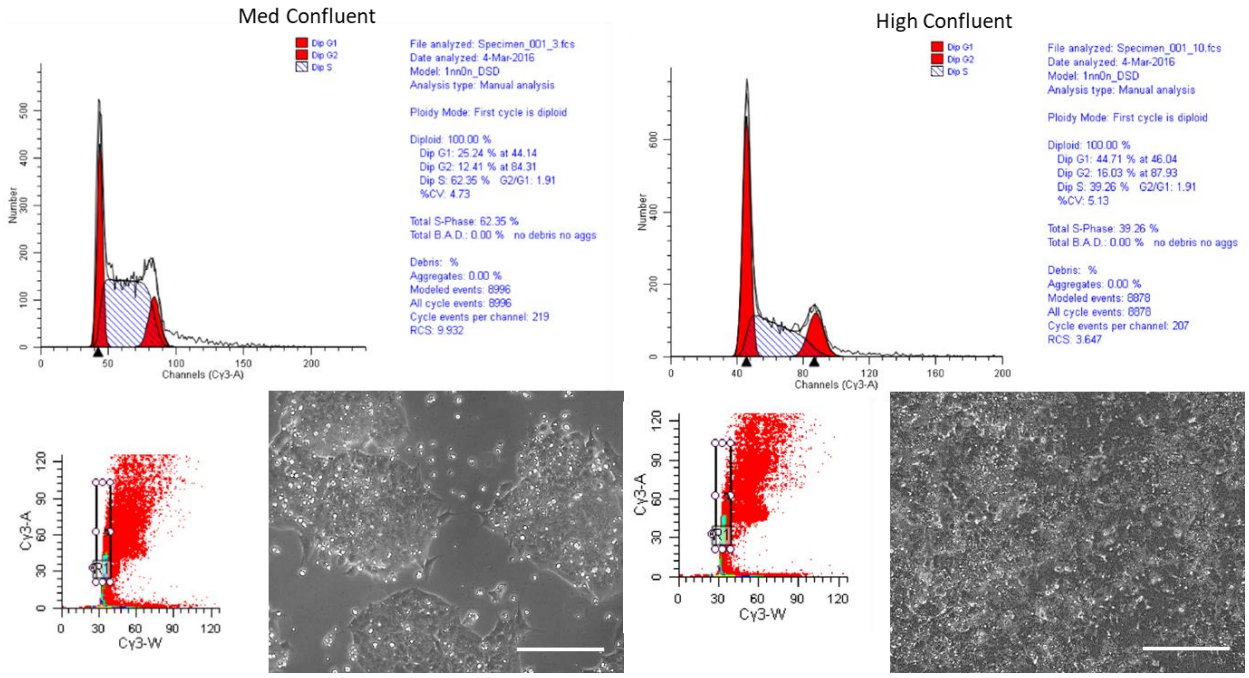


Figure 27: Normal serum and low serum culture conditions. An equal number of reprogrammed cells were cultured in either 15% serum or 1% serum. Representative images are shown on the left. Cells were counted with hemocytometer every day as shown on the right. Scale bars, 100 µm. Error bars, S.E.M., n = 3. *** $P < 0.001$, **** $P < 0.0001$. All P-values were calculated on the basis of a two-tailed, unpaired t-test.

We wanted to assess for other key features of malignant cells. One of the hallmarks of a malignant cell is its loss of contact inhibition. In fact, hypersensitivity to contact inhibition has shown to be cancer resistance.[197] To assess this, iTLCs were cultured at two different confluencies and measured for cell cycle. As shown in Figure 28, iTLCs in medium confluent culture showed a higher number of mitotic cells (approximately 2/3 of the population). In high confluent culture, the number of mitotic cells fell roughly half and percentage of cells in G0 doubled. This suggests that iTLCs exhibit contact inhibition which is something that has shown to be a powerful anticancer mechanism.[197]



	Diploid	G0/1	S	G2	G1/G2	S/G1
Med Confluent	100%	25.24%	62.35%	12.41%	2.03	2.47
High Confluent	100%	44.71%	39.26%	16.03%	2.79	0.88

Figure 28: Cell cycle analysis. iTLCs were cultured at two different confluencies, and their cell cycles were measured. Representative images of medium confluent (left) and of high confluent (right) are shown. The quantified cell cycles are shown in the table (bottom). Scale bars, 100 µm.

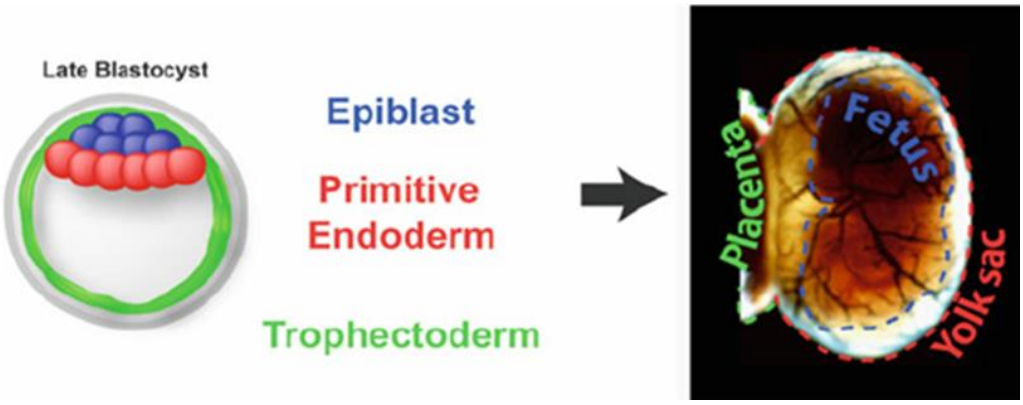
Manipulating and stressing the pluripotent stem cells introduce a possibility of spontaneous transformation towards malignant cells. Our comprehensive genomic analysis showed no aberrations in iTLCs. While we did not sequence the genomic DNA, this was not necessary considering that no DNA was used at any point during the reprogramming process. The synthetic modified mRNA reprogramming method is integration- and virus-free, therefore, no DNA mutations were expected from the reprogramming process. Furthermore, vital functional features of malignant cells were not shown to be present in iTLCs.

VIII.

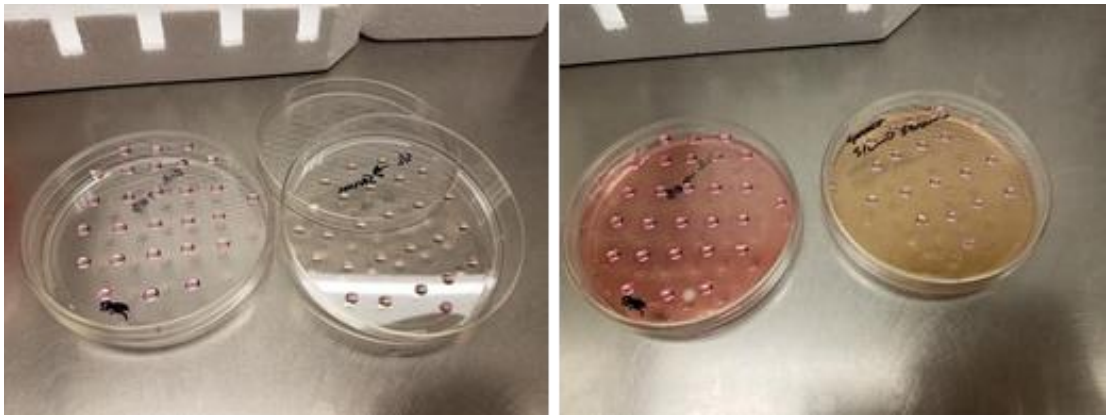
Totipotent Functions

"When I saw the embryo, I suddenly realized there was such a small difference between it and my daughters. I thought, we can't keep destroying embryos for our research. There must be another way." – Shinya Yamanaka

A key defining totipotency is its ability to give rise to the entire organism. Specifically, totipotent cells can differentiate into both embryonic and extraembryonic tissues. To be even more specific, totipotent cells can differentiate into all three distinct lineages of the pre-implantation embryo that are responsible for the entire embryo development process. The following are the three lineages: epiblast which gives rise to the fetus, hypoblast which gives rise to the yolk sac, and trophoblast which gives rise to the placenta. Therefore, we sought out to determine whether iTLCs can differentiate into these three distinct pre-implantation lineages. The following diagram obtained from Google Image summarizes these three lineages.



Embryoid body (EB) differentiation is a three-dimensional aggregate of pluripotent stem cells that presumably resembles some of the early embryogenesis. Instead of using pluripotent stem cells, we used iTLCs to form EB differentiation. There are several different methods to create EBs. However, we found the hanging drop method to be the only one that has yielded positive staining. We have tried both with buffer media and without and found that buffer media was necessary for most time points of analysis. The image of our hanging drop set up is shown below.



The transcription markers for the three pre-implantation lineages are Oct4 for epiblast (fetus), Gata6 for hypoblast (yolk sac), and Cdx2 for trophoblast (placenta). Therefore, we assessed whether iTLC EBs gave rise to these three distinct lineages by analyzing these expressions. As shown in Figure 29, Oct4 and Gata6 were strongly expressed, but Cdx2 was expressed by fewer cells.

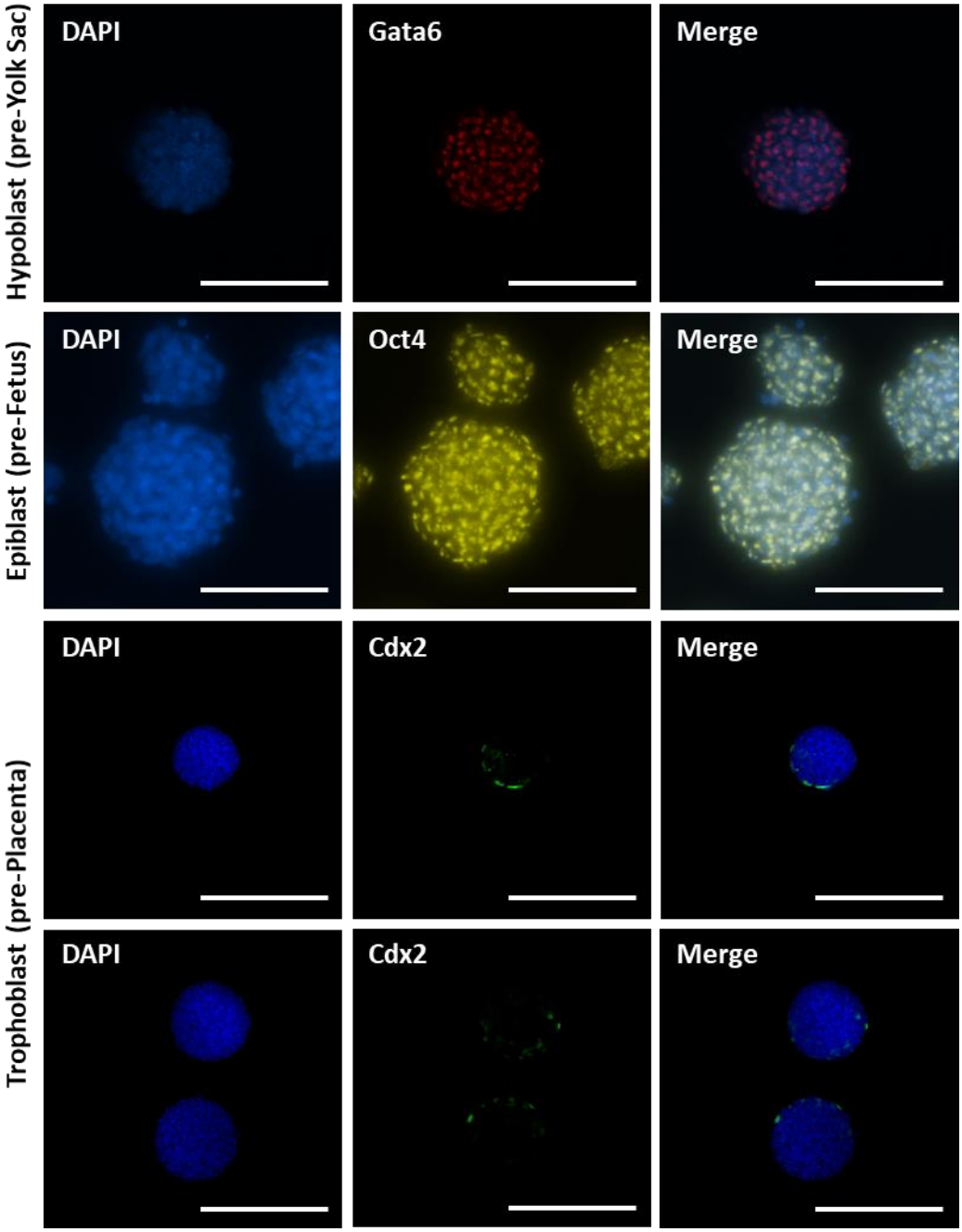


Figure 29: iTLC EBs at day four show expressions of three distinct pre-implantation lineages. Oct4 = epiblast (fetus), Gata6 = hypoblast (yolk sac), Cdx2 = trophoblast (placenta). Scale bars, 100 µm.

We allowed iTLC EBs to differentiate for longer and analyzed the markers of all three germ layers. As shown in Figure 30, iTLCs can give rise to all three germ layers. Interestingly, we found noticeably stronger Cdx2 expression in these EBs.

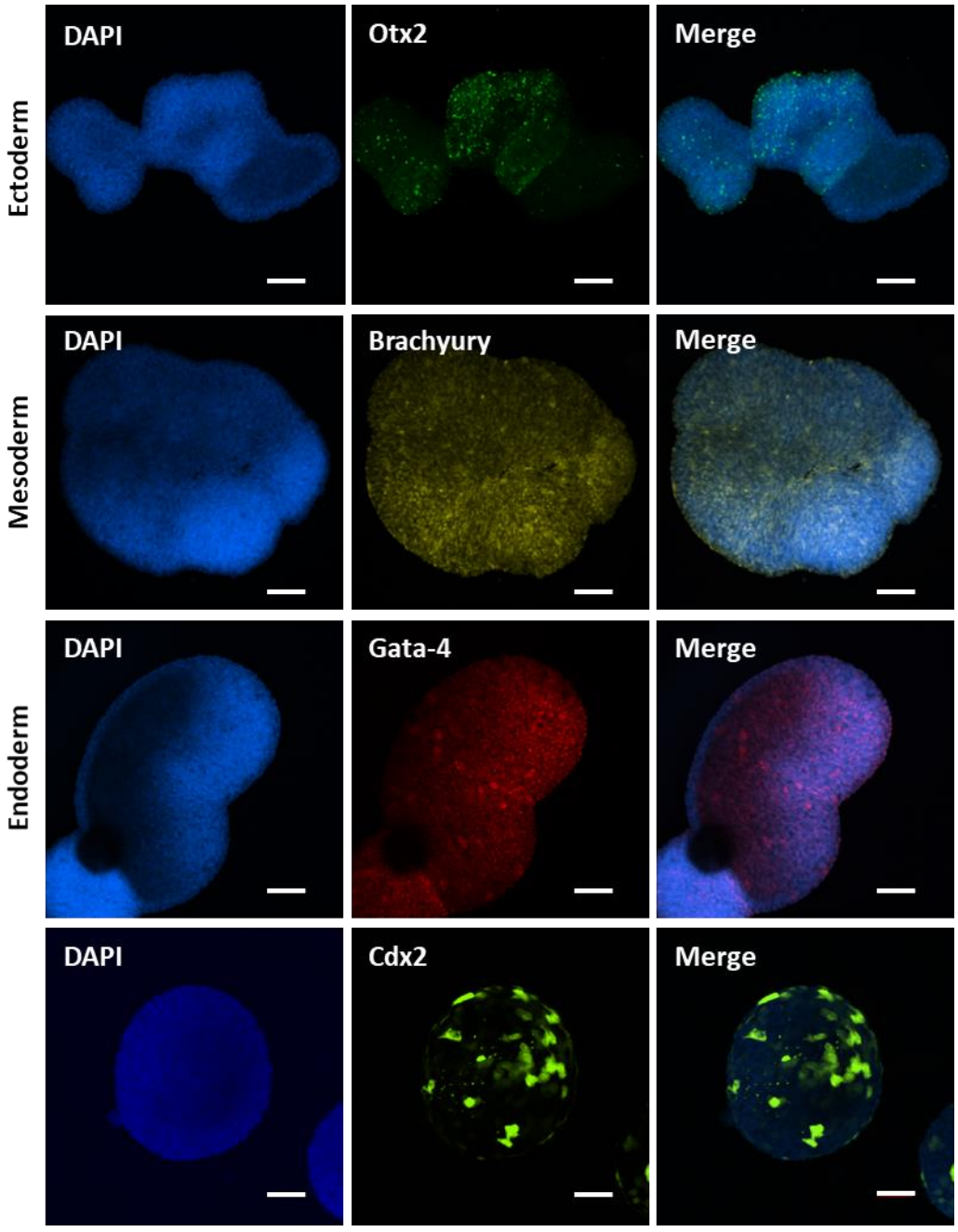


Figure 30: iTLC EBs at day eight show markers of all three germ layers. Cdx2 was also strongly detected. Otx2 = ectoderm, Brachyury = mesoderm, Gata-4 = endoderm. Scale bars, 100 µm.

Observing the stronger Cdx2 expressions at a later time point prompted us to perform a time course quantification analysis of trophectoderm transcripts. As shown in Figure 31, iTLC EBs show increased expression of trophectoderm genes. Eomes and Elf5 expressions were increased significantly compared to ESC EBs. In contrast, iTLC EBs and ESC EBs similarly induced the expressions of Otx2, Brachyury, and Gata-4. Altogether, iTLCs can generate both embryonic and extraembryonic lineages.

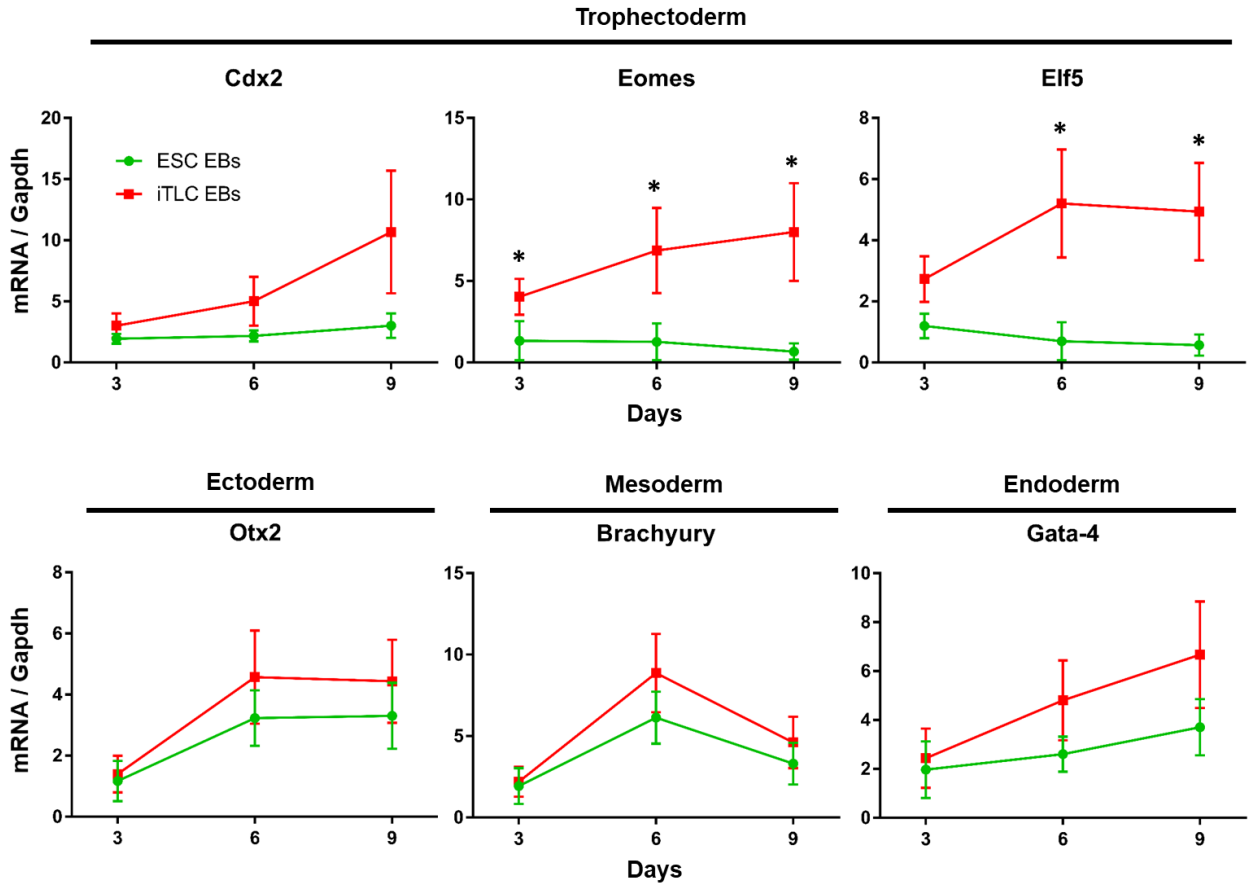


Figure 31: Time course RT-qPCR analysis of the indicated genes from iTLC EBs and ESC EBs. Trophectoderm was noticeably increased in iTLC EBs when compared to ESC EBs. In contrast, three germ layers were

similarly induced in iTLC EBs and ESC EBs. Error bars, S.E.M., n = 3. * $P < 0.05$. All P-values were calculated on the basis of a two-tailed, unpaired t-test.

A notable difference between a totipotent cell and a pluripotent cell is that a totipotent cell can generate both embryonic and extraembryonic tissues while pluripotent stem cell is developmentally restricted to embryonic tissues. iTLCs have shown to express both embryonic and extraembryonic, highlighting an expanded cell fate potential. At last, revealing the true identity of this tantalizing cell required the most advanced technology used in cell profiling; RNA sequencing (RNA-Seq).

IX.

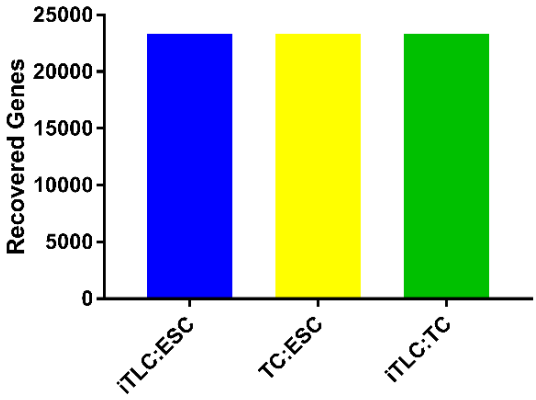
RNA-Seq

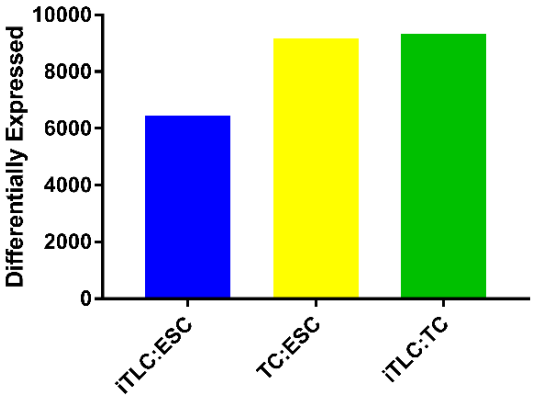
"The development of biology is going to destroy to some extent our traditional grounds for ethical belief and it is not easy to see what to put in their place." – Francis Crick

The final journey towards totipotency required us to suspend our own biases. RNA-Seq generates an unbiased view of the transcriptome and enables the discovery of new cellular states. Three cell types were analyzed; ESC, iTLC, and totipotent cell (TC). The TCs were derived from 2-cell stage mouse embryos.

Triplicate samples of each group were processed, and their average reads were analyzed. As shown in Figure 32, the number of genes that were recovered across all three comparisons were identical. The analysis showed fewer differentially expressed genes between iTLC and ESC (iTLC:ESC) than between the other two comparisons. Since 2-cell stage mouse embryos occur at about 24 hours after fertilization, we reasoned that much of the oocyte's transcripts would still be present at the time when totipotent genes are expressed, resulting in high background noise. A three-way comparison revealed a higher number of iTLC:ESC-only differentially expressed genes when compared to the respective groups. Amongst such

groups, iTLC:TC-only had the lowest differentially expressed genes, indicating that iTLC and TC share more similarities with each other than with ESC.





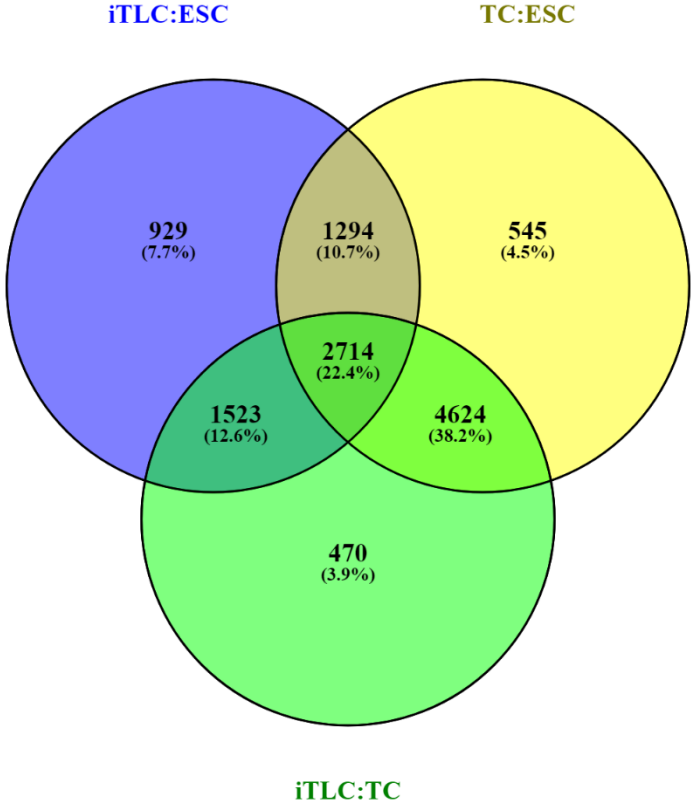


Figure 32: The number of recovered genes and of differentially expressed for the indicated groups are shown (top). Differentially expressed genes of each group were further analyzed in a three-way comparison as shown in the Venn diagram (bottom). Significantly (log2-fold change > 0.60, adjusted p-value < 0.05) up- and down-regulated transcripts were used for differentially expressed analysis.

To verify that the background noise introduced by the persistent oocyte-stored transcripts are impairing the unbiased comparison of totipotency, we analyzed the genes that were most up- and down-regulated as shown in Figure 33. As expected, Chaf1b and Chaf1a were amongst the most downregulated in iTLC when compared against ESC. However, the most upregulated expression in TC when compared against ESC was Omt2a/Omt2b, an oocyte maturation gene. Another expression that was amongst the most upregulated in TC was Oosp1, an oocyte-secreted protein gene. Lastly, the third most upregulated in TC was BTG4 which has known the anti-proliferative function and known to induce G1 arrest in the cell cycle, most likely for the meiotic arrest seen in oogenesis and probably not for the totipotent cells undergoing proliferation. Comparison between iTLC and TC also showed similar finding. The most significant differences came from factors that appear to be important in oogenesis but not in the totipotent state, highlighting the need for much more refined analysis.

iTLC:ESC

Top Analysis-Ready Molecules

Expr Log Ratio up-regulated

Molecules	Expr. Value
Zscan4c (includes others)*	10.196
CHIT1	9.675
Gm5039	9.467
Gm5662 (includes others)*	9.449
Zscan4a	9.392
AU015228	9.225
Gm4971	9.100
Trim75	9.065
Tcstv1	8.964
Gm12794	8.928

Expr Log Ratio down-regulated

Molecules	Expr. Value
FGF17	-6.461
CHAF1B	-4.777
Nkx6-3	-4.285
KIAA0408	-4.080
MYO1F	-4.055
EDN2	-3.869
NTN1	-3.843
INHBB	-3.825
CHAF1A	-3.805
Tdh	-3.685

TC:ESC

Top Analysis-Ready Molecules

Expr Log Ratio up-regulated

Molecules	Expr. Value
Omt2a/Omt2b*	16.446
Gm13088 (includes others)*	16.158
BTG4	15.593
Obox1 (includes others)*	15.421
Oosp1	15.371
Tcl1b4 (includes others)*	15.152
BRDT	15.130
E330034G19Rik	14.748
C86187	14.743
UNC13C	14.607

Expr Log Ratio down-regulated

Molecules	Expr. Value
NR0B1	-12.843
SLC7A3	-12.498
TDGF1	-11.805
EXOC3L1	-11.082
ADAP1	-10.950
DUSP27	-10.799
Krt42	-10.734
CCDC22	-10.661
WFDC2	-10.499
Gm7325	-10.414

iTLC:TC

Top Analysis-Ready Molecules

Expr Log Ratio up-regulated

Molecules	Expr. Value
CA2	13.235
FGFBP1	12.973
TDGF1	11.720
WFDC2	11.650
PSORS1C2	11.579
NR0B1	11.248
FHL1	10.661
Rhox5	10.447
CALCB	10.377
EXOC3L1	10.205

Expr Log Ratio down-regulated

Molecules	Expr. Value
Gm13088 (includes others)*	-15.883
BTG4	-15.330
Oosp1	-15.114
Omt2a/Omt2b*	-15.019
C86187	-14.493
UNC13C	-14.359
KHDC1L*	-14.053
AA545190	-13.762
Tcl1b4 (includes others)*	-13.653
Nlrp9a/Nlrp9c*	-13.210

Figure 33: Ingenuity pathway analysis. Some of the most up- and down-regulated expressions of each comparison as indicated are shown.

Genes that have been shown to express in oocyte and different stages of the embryo were analyzed in each group. As shown in Figure 34, when compared to ESC, iTLC showed little to no difference in oocyte expression while TC, as predicted, showed higher expression. Interestingly, both iTLC and TC showed remarkably high expressions of the embryo at the 2-cell stage but showed notable downregulation of the other stages. When iTLC and TC were compared, iTLC showed downregulation of oocyte expression, as expected. The expressions of 2-cell stage embryo, however, showed the much smaller difference between the two, indicating that they intimately share this particular cell type. Additionally, iTLC showed small upregulation of genes in morula and early blastocyst, suggesting some lingering of pluripotent expressions. Altogether, this shows that iTLC and TC closely resemble each other along the embryo development and are highly expressed for 2-cell stage embryo.

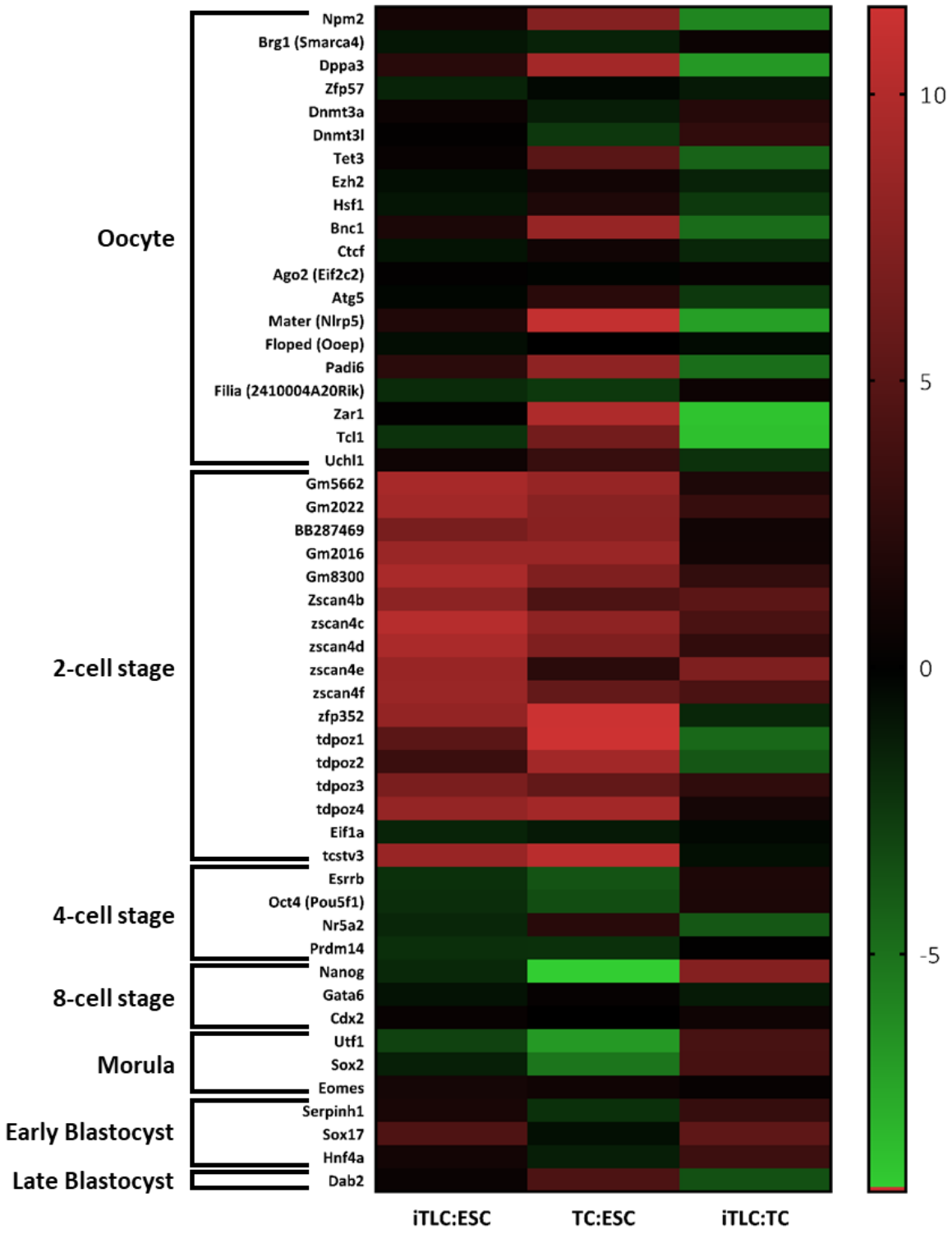


Figure 34A: Heatmap showing the relative expression levels of oocyte and embryo in the indicated groups.

After observing the specific and robust expressions for the 2-cell stage, we wanted to analyze this in further detail. Specifically, ZGA is a known phenomenon that occurs in totipotent state and peaks during the 2-cell stage. We analyzed genes that have been shown to express at different time points during ZGA. When compared to ESC, iTLC showed most upregulation towards the earlier ZGA than later, resembling the pattern found in TC. When iTLC was compared against TC, greater emphasis was put on for earlier genes of ZGA. Med- to- late-ZGA genes were downregulated in comparison to TC, suggesting that TC was at a more mature state of ZGA while iTLC was at the beginning. Recently, Dux has been identified as a key inducer of ZGA in placental mammals.[198] This particular gene was notably upregulated in iTLC and downregulated in TC, further suggesting that iTLC was at the beginning stages of ZGA.

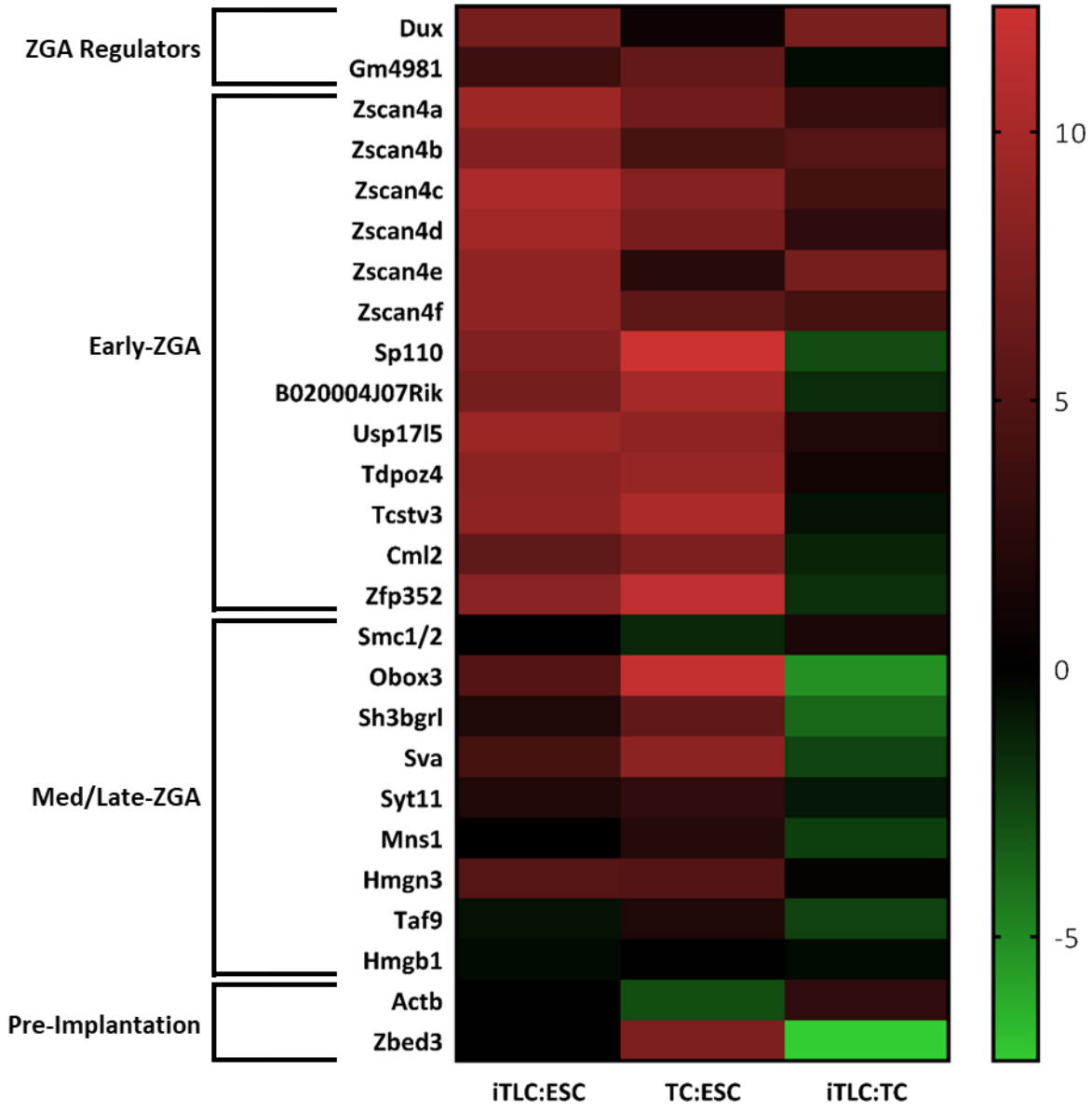


Figure 34B: Heatmap showing the relative expression levels of zygotic genome activation and pre-implantation in the indicated groups.

Encouraged by these results, we questioned whether we could generate zygotes that can recapitulate the embryogenesis. We reasoned that the efficiency, if possible at all, will likely be very low, and our previous analysis showed that iTLC loses totipotency quickly, possibly due to incomplete reprogramming.

Therefore, we tried reprogramming for a much longer duration. Unfortunately, when TSA and p150 siRNA were used, we were unable to maintain enough healthy cells for a very long period. Instead, we reprogrammed with C1 only without passaging to maximize cell viability. About a month later, we noticed a very few exceptionally large cells shown in Figure 35. Some had diameters of approximately 70-100 µm, consistent with a mature oocyte or a zygote. Furthermore, thick transparent membranes resembling zona pellucida surrounded some of these giant cells. This is further supported by our RT-qPCR data that show upregulation of zona pellucid genes in reprogrammed cells. Other giant cells showed what appears to be two pronuclei that are seen in fertilized eggs. Also, other interesting cell populations were observed that resembled blastocysts, embryo hatching, and other stages of embryogenesis.

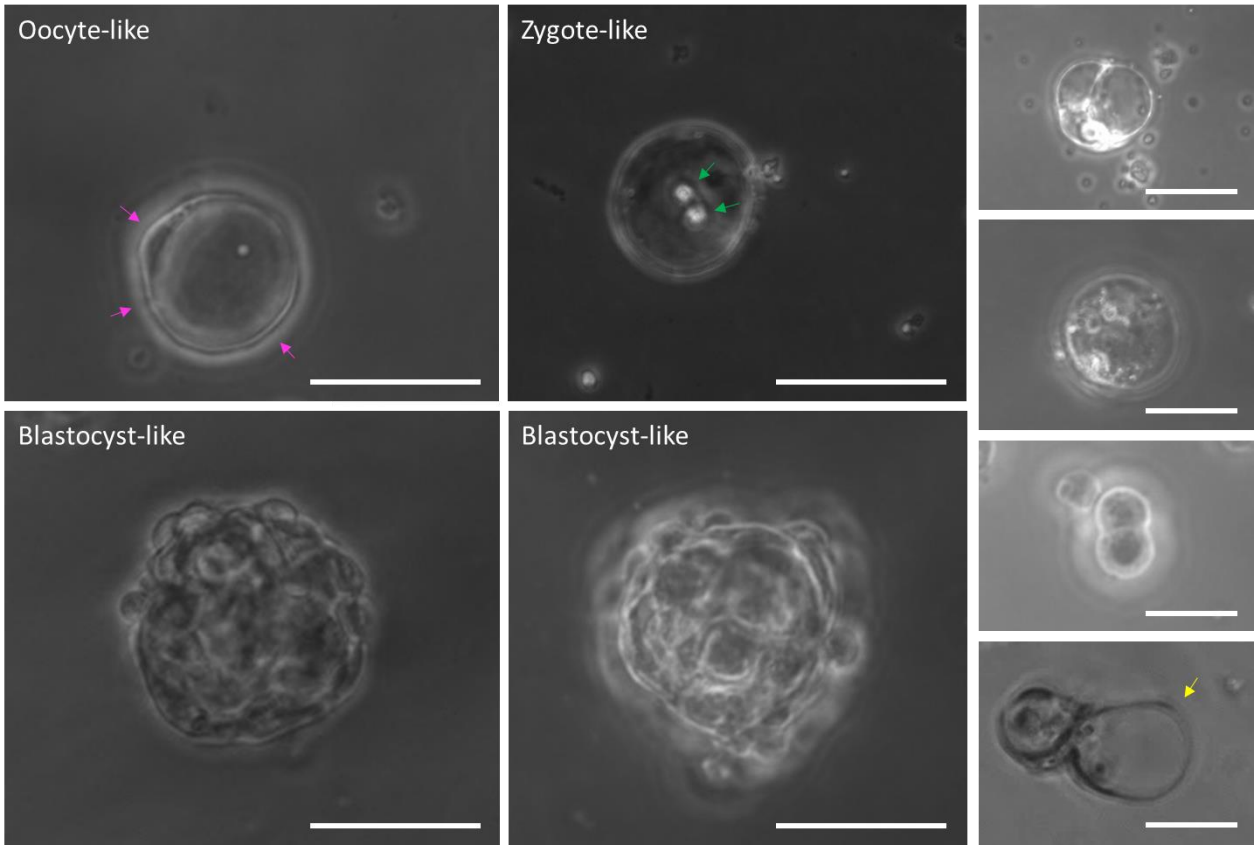


Figure 35: Prolonged reprogramming towards totipotency gave rise to cells that resemble early embryogenesis. Pink = thick transparent membrane resembling zona pellucida, green = distinct nuclei resembling pronucleus, yellow = phenomenon resembling embryo hatching. Scale bars, 100 µm.

Transcriptome analysis of iTLC revealed remarkable similarities to 2-cell stage embryos. When reprogrammed for a long time, cells resembling different stages of embryogenesis were observed. Despite our best effort to isolate and characterize these peculiar cells, especially the oocyte- and zygote-like, we were unable to capture them successfully. Continous reprogramming generated only a few, indicating extremely low efficiency, and most were lost during the isolation process. It is possible that these cells were very fragile and got destroyed in the process. The very few that made the transfer to a new plate were all imaged and shown in Figure 35. While the images show us that these are not cell debris or artifacts, it is possible that these are not totipotent cells but of a different kind. The exact nature of these tantalizing cells remains a mystery. If these are genuinely totipotent, then this development is going to destroy to some extent our traditional belief, and it is not easy to see what the potential implications will be.

X.

Legacy

"Science is like building a building, brick by brick, and you can't build the 10th floor until you build the first." – John "Jack" A. Kessler

Reprogramming pluripotent stem cell towards totipotency was a journey that was inspired by the work that has come before us. SCNT broke new ground by proving that terminally differentiated cell can be reprogrammed back to its beginning. iPSC raised this to a new height by showing that cellular reprogramming can be achieved with defined factors, obviating the need for an oocyte. We wanted to build upon this by showing that defined factors can induce totipotency. This led us to build synthetic modified mRNA for reprogramming, utilize MuERV-L for a reporter, and identify factors and chemicals for totipotency.

We were able to witness some exciting findings. Not all zygotes appear to be totipotent. The stringency of MuERV-L as a totipotent marker may help distinguish the heterogeneous population, possibly aiding in reproductive technology. Maternal factors can induce a 2C-like state that can be further improved upon

with HDAC inhibitor and CAF-1 knockdown, indicating that altering chromatin architecture may be as effective totipotent inducer as direct activation of totipotent-specific genes. The increased totipotent genes and elements, combined with the decrease from pluripotent and three germ layers, showed committed unidirectional change towards totipotency by the reprogrammed cells. Enlarged nuclei, elongated telomere, and responsive to totipotent conditions further highlighted the totipotent traits. The reprogrammed cells maintained genomic stability, nutrient sensitivity, and contact inhibition, reassuring us that transformation towards malignant cell was unlikely. The reprogrammed cells were able to be differentiated towards both embryonic and extraembryonic lineages, suggesting an expanded cell fate potential. This was further supported by RNA-Seq analysis showing remarkable similarities between the reprogrammed cells and the 2-cell stage embryos. Lastly, a tiny fraction of reprogrammed cells appears to have the capacity to recapitulate the early embryogenesis.

Since the beginning of our journey, many others have published totipotent-related findings. One critical shortcoming was our inability to continually update our reprogramming strategies based on the most recent data. Changes to reprogramming method would require repeating the previous experiments which were not an option. This means that our reprogramming strategy is outdated. We will briefly discuss several new findings that may have improved our strategy.

One significant drawback in our reprogramming was that it caused noticeable cell deaths, indicating a very stressful condition. This encouraged us to confirm that no malignant cells were generated, which continues to remain a possibility. More importantly, this limited the duration of the reprogramming process. We identified the exogenous chemicals, TSA and p150 siRNA, to be the most significant contributors of cytotoxicity. Most recent findings showed that microRNA miR-34a restricts cell fate potential in pluripotent stem cell, and when deficient, MuERV-L was strongly induced, and the cells gained the capacity to generate both embryonic and extraembryonic tissues.[199] Also, miR-34a deficiency has also

shown to improve somatic reprogramming efficiency and kinetics significantly.[200] Mechanisms of microRNAs are relatively much more precise and intentional than broad sweeping generic effects of chemical inhibitors. Therefore, this may be an excellent replacement to the exogenous molecules that were used in our reprogramming. Detailed molecular mechanisms revealed that miR-34a directly inhibits Gata2 transcription factor, and Gata2 binds and activates MuERV-L in the absence of miR-34a.[132,199] Furthermore, Gata2 was necessary but insufficient to induce MuERV-L, revealing other possible candidate genes.[132,199]

In addition to replacing exogenous chemicals, the list of candidate genes also needs to be updated based on the latest known mechanisms. Most recently, Dux gene has shown to be a key initiator of ZGA.[198] This particular gene is especially intriguing for our purpose since it recruits other complexes to induce chromatin relaxation but also acts like other known master inducers of ZGA found in other species.[201-203] If Dux is the master initiator of ZGA, its pathway analysis will reveal additional factors that can be served as candidate genes.

Lastly, one notable defining characteristic of a totipotent cell must be addressed. A totipotent cell is particularly unique in that its size is considerable, approximately 70-90 μm for mouse and a bit larger for a human. This is most likely due to its immediate predecessor, oocyte, being of similarly large size. Zygotes that are smaller than average do not undergo successful development and have higher rates of embryo arrests. Not surprisingly, the size of a fertilized embryo is commonly used screening method in IVF clinics. This suggests that a cell may be transcriptionally totipotent but functionally not totipotent. Therefore, proving totipotency requires that the cell is of the correct size. We have shown that our cells are transcriptionally at 2-cell stage, which would indicate that the cell likely had all of the necessary proteins needed for the embryo development. However, we would not expect to observe a cleavage process occurring in a cell that is only a tiny fraction in size. Interestingly, lengthy reprogramming process resulted

in a few cells that were of the comparable size to a zygote. The lengthy duration of reprogramming was carried out without passage in order to maintain maximal cell viabilities, and this resulted in crowding and compaction of cells. Similar compaction plays an essential role in oogenesis.

An oocyte initially starts out as a much smaller cell that grows in size over time with the support of follicular cells surrounding it. This cell-to-cell interaction continues throughout the entire process until the oocyte is fully grown, leaving behind all the others that have contributed to its final development for the sole purpose of starting the next generation. It appears that becoming functionally totipotent requires more than a transcript reprogramming of an individual cell. The correct totipotent reprogramming would involve other types of cells collectively surrounding the 2C, and actively contributing to its growth until it has reached the proper maturity to start the next generation and carry on the legacy.

1 Hooke, R. *Micrographia, or, Some physiological descriptions of minute bodies made by magnifying glasses : with observations and inquiries thereupon*. (London : Printed for James Allestry ... and are to be sold at his shop ..., MDCLXV11 [1667], 1667).

2 Gest, H. The discovery of microorganisms by Robert Hooke and Antoni Van Leeuwenhoek, fellows of the Royal Society. *Notes Rec R Soc Lond.* **58** (2), 187-201 (2004).

3 Martin, J. J., Schatz, V. & Schatz, A. Contribution to the microbiology of the mouth by Antony van Leeuwenhoek (24 October 1632 - 26 August 1732). *Pak Dent Rev.* **21** (2), 43-46 (1971).

4 Shapiro, S. Antony van Leeuwenhoek; a review of his life and work. *J Biol Photogr Assoc.* **23** (2-3), 49-57 (1955).

5 Oliver, F. W. *Makers of British botany; a collection of biographies by living botanists*. (University press, 1913).

6 Gall, J. G., McIntosh, J. R. & American Society for Cell Biology. *Landmark papers in cell biology*. (Cold Spring Harbor Laboratory Press ;, 2001).

7 Mazzarello, P. A unifying concept: the history of cell theory. *Nat Cell Biol.* **1** (1), E13-15, doi:10.1038/8964, (1999).

8 Schultz, M. Rudolf Virchow. *Emerging Infectious Diseases.* **14** (9), 1480-1481, doi:10.3201/eid1409.086672, (2008).

9 Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A.* **78** (12), 7634-7638 (1981).

10 Thomson, J. A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science.* **282** (5391), 1145-1147 (1998).

11 Ramalho-Santos, M. & Willenbring, H. On the origin of the term "stem cell". *Cell Stem Cell.* **1** (1), 35-38, doi:10.1016/j.stem.2007.05.013, (2007).

12 Bleek, W. H. I. & Haeckel, E. *Über den ursprung der sprache*. (H. Boehlau, 1868).

13 Haeckel, E. *Natürliche schöpfungsgeschichte. Gemeinverständliche wissenschaftliche vorträge über die entwickelungslehre im allgemeinen und diejenige von Darwin, Goethe und Lamarck im besonderen*. 5. verb. aufl. Mit dem porträt des verfassers (nach einer photographie) und mit 16 tafeln, 19 holzschnitten, 18 stammbäumen und 19 systematischen tabellen. edn, (G. Reimer, 1874).

14 Morrison, S. J., Shah, N. M. & Anderson, D. J. Regulatory mechanisms in stem cell biology. *Cell.* **88** (3), 287-298 (1997).

15 Weissman, I. L. Stem cells: units of development, units of regeneration, and units in evolution. *Cell.* **100** (1), 157-168 (2000).

16 Lu, F. & Zhang, Y. Cell totipotency: molecular features, induction, and maintenance. *Natl Sci Rev.* **2** (2), 217-225, doi:10.1093/nsr/nwv009, (2015).

17 Mitalipov, S. & Wolf, D. Totipotency, pluripotency and nuclear reprogramming. *Adv Biochem Eng Biotechnol.* **114** 185-199, doi:10.1007/10_2008_45, (2009).

18 Singh, V. K., Saini, A., Kalsan, M., Kumar, N. & Chandra, R. Describing the Stem Cell Potency: The Various Methods of Functional Assessment and In silico Diagnostics. *Front Cell Dev Biol.* **4** 134, doi:10.3389/fcell.2016.00134, (2016).

19 Hanna, J. H., Saha, K. & Jaenisch, R. Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. *Cell.* **143** (4), 508-525, doi:10.1016/j.cell.2010.10.008, (2010).

20 Wuidart, A. *et al.* Quantitative lineage tracing strategies to resolve multipotency in tissue-specific stem cells. *Genes Dev.* **30** (11), 1261-1277, doi:10.1101/gad.280057.116, (2016).

21 Zipori, D. Multipotency and Tissue-Specific Stem Cells. *Biology of Stem Cells and the Molecular Basis of the Stem State.* 39-55, doi:10.1007/978-1-60761-130-1_2, (2009).

22 Majo, F., Rochat, A., Nicolas, M., Jaoude, G. A. & Barrandon, Y. Oligopotent stem cells are distributed throughout the mammalian ocular surface. *Nature.* **456** (7219), 250-254, doi:10.1038/nature07406, (2008).
23 Blanpain, C., Horsley, V. & Fuchs, E. Epithelial stem cells: turning over new leaves. *Cell.* **128** (3), 445-458, doi:10.1016/j.cell.2007.01.014, (2007).
24 Crescenzi, M. *Reactivation of the cell cycle in terminally differentiated cells*. (Landes Bioscience ; Kluwer Academic/Plenum Publishers, 2002).
25 Niwa, H., Miyazaki, J. & Smith, A. G. Quantitative expression of Oct3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet.* **24** 372-376 (2000).
26 Smith, A. G. Embryo-derived stem cells: of mice and men. *Ann Rev Cell Dev Biol.* **17** 455-462 (2001).
27 Niwa, H., Burdon, T., Chambers, I. & Smith, A. G. Self-renewal of pluri-potent embryonic stem cells is mediated via activation of STAT3. *Genes Dev.* **12** 2048-2060 (1998).
28 Smith, A. G., Nichols, J., Robertson, M. & Rathjen, P. D. Differentiation inhibiting activity (DIA/LIF) and mouse development. *Dev Biol.* **151** 339-351 (1992).
29 Ying, Q. L., Nichols, J., Chambers, I. & Smith, A. BMP Induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell.* **115** 281-292 (2003).
30 Donovan, P. J. The germ cell--the mother of all stem cells. *Int J Dev Biol.* **42** (7), 1043-1050 (1998).
31 Donovan, P. J. & de Miguel, M. P. Turning germ cells into stem cells. *Curr Opin Genet Dev.* **13** (5), 463-471 (2003).
32 Kanatsu-Shinohara, M. & Shinohara, T. The germ of pluripotency. *Nat Biotechnol.* **24** (6), 663-664, doi:10.1038/nbt0606-663, (2006).
33 Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell.* **126** (4), 663-676, doi:10.1016/j.cell.2006.07.024, (2006).
34 Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell.* **131** (5), 861-872, doi:10.1016/j.cell.2007.11.019, (2007).
35 Nakagawa, M. *et al.* Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotechnol.* **26** (1), 101-106, doi:10.1038/nbt1374, (2008).
36 Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T. & Yamanaka, S. Generation of mouse induced pluripotent stem cells without viral vectors. *Science.* **322** (5903), 949-953, doi:10.1126/science.1164270, (2008).
37 Warren, L. *et al.* Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell.* **7** (5), 618-630, doi:10.1016/j.stem.2010.08.012, (2010).
38 Roessler, R. *et al.* Detailed analysis of the genetic and epigenetic signatures of iPSC-derived mesodiencephalic dopaminergic neurons. *Stem Cell Reports.* **2** (4), 520-533, doi:10.1016/j.stemcr.2014.03.001, (2014).
39 Marchetto, M. C. *et al.* Transcriptional signature and memory retention of human-induced pluripotent stem cells. *PLoS One.* **4** (9), e7076, doi:10.1371/journal.pone.0007076, (2009).
40 (!!! INVALID CITATION !!! {}).
41 Kim, K. *et al.* Epigenetic memory in induced pluripotent stem cells. *Nature.* **467** (7313), 285-290, doi:10.1038/nature09342, (2010).
42 Bar-Nur, O., Russ, H. A., Efrat, S. & Benvenisty, N. Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet beta cells. *Cell Stem Cell.* **9** (1), 17-23, doi:10.1016/j.stem.2011.06.007, (2011).

43 Osafune, K. *et al.* Marked differences in differentiation propensity among human embryonic stem cell lines. *Nat Biotechnol.* **26** (3), 313-315, doi:10.1038/nbt1383, (2008).

44 Feng, Q. *et al.* Hemangioblastic derivatives from human induced pluripotent stem cells exhibit limited expansion and early senescence. *Stem Cells.* **28** (4), 704-712, doi:10.1002/stem.321, (2010).

45 Hu, B. Y. *et al.* Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proc Natl Acad Sci U S A.* **107** (9), 4335-4340, doi:10.1073/pnas.0910012107, (2010).

46 Narsinh, K. H. *et al.* Single cell transcriptional profiling reveals heterogeneity of human induced pluripotent stem cells. *J Clin Invest.* **121** (3), 1217-1221, doi:10.1172/JCI44635, (2011).

47 Kokkinaki, M., Sahibzada, N. & Golestaneh, N. Human induced pluripotent stem-derived retinal pigment epithelium (RPE) cells exhibit ion transport, membrane potential, polarized vascular endothelial growth factor secretion, and gene expression pattern similar to native RPE. *Stem Cells.* **29** (5), 825-835, doi:10.1002/stem.635, (2011).

48 Beagan, J. A. *et al.* Local Genome Topology Can Exhibit an Incompletely Rewired 3D-Folding State during Somatic Cell Reprogramming. *Cell Stem Cell.* **18** (5), 611-624, doi:10.1016/j.stem.2016.04.004, (2016).

49 Hussein, S. M. *et al.* Copy number variation and selection during reprogramming to pluripotency. *Nature.* **471** (7336), 58-62, doi:10.1038/nature09871, (2011).

50 Laurent, L. C. *et al.* Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. *Cell Stem Cell.* **8** (1), 106-118, doi:10.1016/j.stem.2010.12.003, (2011).

51 Ramos-Mejia, V., Munoz-Lopez, M., Garcia-Perez, J. L. & Menendez, P. iPSC lines that do not silence the expression of the ectopic reprogramming factors may display enhanced propensity to genomic instability. *Cell Res.* **20** (10), 1092-1095, doi:10.1038/cr.2010.125, (2010).

52 Bolton, H. *et al.* Mouse model of chromosome mosaicism reveals lineage-specific depletion of aneuploid cells and normal developmental potential. *Nat Commun.* **7** 11165, doi:10.1038/ncomms11165, (2016).

53 Kimbrel, E. A. & Lanza, R. Current status of pluripotent stem cells: moving the first therapies to the clinic. *Nat Rev Drug Discov.* **14** (10), 681-692, doi:10.1038/nrd4738, (2015).

54 Briggs, R. & King, T. J. Transplantation of Living Nuclei From Blastula Cells into Enucleated Frogs' Eggs. *Proc Natl Acad Sci U S A.* **38** (5), 455-463 (1952).

55 Gurdon, J. B., Elsdale, T. R. & Fischberg, M. Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei. *Nature.* **182** (4627), 64-65 (1958).

56 Bromhall, J. D. Nuclear transplantation in the rabbit egg. *Nature.* **258** (5537), 719-722 (1975).

57 Willadsen, S. M. & Godke, R. A. A simple procedure for the production of identical sheep twins. *Vet Rec.* **114** (10), 240-243 (1984).

58 Prather, R. S. *et al.* Nuclear transplantation in the bovine embryo: assessment of donor nuclei and recipient oocyte. *Biol Reprod.* **37** (4), 859-866 (1987).

59 Campbell, K. H., McWhir, J., Ritchie, W. A. & Wilmut, I. Sheep cloned by nuclear transfer from a cultured cell line. *Nature.* **380** (6569), 64-66, doi:10.1038/380064a0, (1996).

60 Meng, L., Ely, J. J., Stouffer, R. L. & Wolf, D. P. Rhesus monkeys produced by nuclear transfer. *Biol Reprod.* **57** (2), 454-459 (1997).

61 Schnieke, A. E. *et al.* Human factor IX transgenic sheep produced by transfer of nuclei from transfected fetal fibroblasts. *Science.* **278** (5346), 2130-2133 (1997).

62 Byrne, J. A. *et al.* Producing primate embryonic stem cells by somatic cell nuclear transfer. *Nature.*

450 (7169), 497-502, doi:10.1038/nature06357, (2007).
63 Tachibana, M. *et al.* Human embryonic stem cells derived by somatic cell nuclear transfer. *Cell.* **153** (6), 1228-1238, doi:10.1016/j.cell.2013.05.006, (2013).
64 Burns, K. H. *et al.* Roles of NPM2 in chromatin and nucleolar organization in oocytes and embryos. *Science.* **300** (5619), 633-636, doi:10.1126/science.1081813, (2003).
65 Li, L., Baibakov, B. & Dean, J. A subcortical maternal complex essential for preimplantation mouse embryogenesis. *Dev Cell.* **15** (3), 416-425, doi:10.1016/j.devcel.2008.07.010, (2008).
66 Flemr, M., Ma, J., Schultz, R. M. & Svoboda, P. P-body loss is concomitant with formation of a messenger RNA storage domain in mouse oocytes. *Biol Reprod.* **82** (5), 1008-1017, doi:10.1095/biolreprod.109.082057, (2010).
67 Wossidlo, M. *et al.* 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat Commun.* **2** 241, doi:10.1038/ncomms1240, (2011).
68 Iqbal, K., Jin, S. G., Pfeifer, G. P. & Szabo, P. E. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc Natl Acad Sci U S A.* **108** (9), 3642-3647, doi:10.1073/pnas.1014033108, (2011).
69 Gu, T. P. *et al.* The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature.* **477** (7366), 606-610, doi:10.1038/nature10443, (2011).
70 Wang, L. *et al.* Programming and inheritance of parental DNA methylomes in mammals. *Cell.* **157** (4), 979-991, doi:10.1016/j.cell.2014.04.017, (2014).
71 Guo, F. *et al.* Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. *Cell Stem Cell.* **15** (4), 447-459, doi:10.1016/j.stem.2014.08.003, (2014).
72 Wyatt, G. R. & Cohen, S. S. The bases of the nucleic acids of some bacterial and animal viruses: the occurrence of 5-hydroxymethylcytosine. *Biochem J.* **55** (5), 774-782 (1953).
73 Adenot, P. G., Mercier, Y., Renard, J. P. & Thompson, E. M. Differential H4 acetylation of paternal and maternal chromatin precedes DNA replication and differential transcriptional activity in pronuclei of 1-cell mouse embryos. *Development.* **124** (22), 4615-4625 (1997).
74 Bultman, S. J. *et al.* Maternal BRG1 regulates zygotic genome activation in the mouse. *Genes Dev.* **20** (13), 1744-1754, doi:10.1101/gad.1435106, (2006).
75 Schultz, R. M. Regulation of zygotic gene activation in the mouse. *Bioessays.* **15** (8), 531-538, doi:10.1002/bies.950150806, (1993).
76 Davis, W., Jr. & Schultz, R. M. Developmental change in TATA-box utilization during preimplantation mouse development. *Dev Biol.* **218** (2), 275-283, doi:10.1006/dbio.1999.9486, (2000).
77 Peaston, A. E. *et al.* Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev Cell.* **7** (4), 597-606, doi:10.1016/j.devcel.2004.09.004, (2004).
78 Macfarlan, T. S. *et al.* Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature.* **487** (7405), 57-63, doi:10.1038/nature11244, (2012).
79 Ishiuchi, T. *et al.* Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat Struct Mol Biol.* **22** (9), 662-671, doi:10.1038/nsmb.3066, (2015).
80 Biswas, D. & Jiang, P. Chemically Induced Reprogramming of Somatic Cells to Pluripotent Stem Cells and Neural Cells. *Int J Mol Sci.* **17** (2), 226, doi:10.3390/ijms17020226, (2016).
81 Long, Y., Wang, M., Gu, H. & Xie, X. Bromodeoxyuridine promotes full-chemical induction of mouse pluripotent stem cells. *Cell Res.* **25** (10), 1171-1174, doi:10.1038/cr.2015.96, (2015).
82 Zhang, Y., Li, W., Laurent, T. & Ding, S. Small molecules, big roles -- the chemical manipulation of stem cell fate and somatic cell reprogramming. *J Cell Sci.* **125** (Pt 23), 5609-5620, doi:10.1242/jcs.096032, (2012).

83 Van Craenenbroeck, K., Vanhoenacker, P. & Haegeman, G. Episomal vectors for gene expression in mammalian cells. *Eur J Biochem.* **267** (18), 5665-5678 (2000).

84 Cheng, L. *et al.* Low incidence of DNA sequence variation in human induced pluripotent stem cells generated by nonintegrating plasmid expression. *Cell Stem Cell.* **10** (3), 337-344, doi:10.1016/j.stem.2012.01.005, (2012).

85 Hu, K. *et al.* Efficient generation of transgene-free induced pluripotent stem cells from normal and neoplastic bone marrow and cord blood mononuclear cells. *Blood.* **117** (14), e109-119, doi:10.1182/blood-2010-07-298331, (2011).

86 Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science.* **318** (5858), 1917-1920, doi:10.1126/science.1151526, (2007).

87 Schlaeger, T. M. *et al.* A comparison of non-integrating reprogramming methods. *Nat Biotechnol.* **33** (1), 58-63, doi:10.1038/nbt.3070, (2015).

88 Nakanishi, M. & Otsu, M. Development of Sendai virus vectors and their potential applications in gene therapy and regenerative medicine. *Curr Gene Ther.* **12** (5), 410-416 (2012).

89 Yamamoto, A., Kormann, M., Rosenecker, J. & Rudolph, C. Current prospects for mRNA gene delivery. *Eur J Pharm Biopharm.* **71** (3), 484-489, doi:10.1016/j.ejpb.2008.09.016, (2009).

90 Mandal, P. K. & Rossi, D. J. Reprogramming human fibroblasts to pluripotency using modified mRNA. *Nat Protoc.* **8** (3), 568-582, doi:10.1038/nprot.2013.019 10.1038/nprot.nprot.2013.019, (2013).

91 Barreau, C., Dutertre, S., Paillard, L. & Osborne, H. B. Liposome-mediated RNA transfection should be used with caution. *RNA.* **12** (10), 1790-1793, doi:10.1261/rna.191706, (2006).

92 Stepinski, J., Waddell, C., Stolarski, R., Darzynkiewicz, E. & Rhoads, R. E. Synthesis and properties of mRNAs containing the novel "anti-reverse" cap analogs 7-methyl(3'-O-methyl)GpppG and 7-methyl (3'-deoxy)GpppG. *RNA.* **7** (10), 1486-1495 (2001).

93 Strenkowska, M. *et al.* Towards mRNA with superior translational activity: synthesis and properties of ARCA tetraphosphates with single phosphorothioate modifications. *New J Chem.* **34** (5), 993-1007, doi:10.1039/b9nj00644c, (2010).

94 Zohra, F. T., Chowdhury, E. H., Tada, S., Hoshiba, T. & Akaike, T. Effective delivery with enhanced translational activity synergistically accelerates mRNA-based transfection. *Biochem Biophys Res Commun.* **358** (1), 373-378, doi:10.1016/j.bbrc.2007.04.059, (2007).

95 Eberhardt, W., Doller, A., Akool el, S. & Pfeilschifter, J. Modulation of mRNA stability as a novel therapeutic approach. *Pharmacol Ther.* **114** (1), 56-73, doi:10.1016/j.pharmthera.2007.01.002, (2007).

96 Coutts, M., Krowczynska, A. & Brawerman, G. Protection of mRNA against nucleases in cytoplasmic extracts of mouse sarcoma ascites cells. *Biochim Biophys Acta.* **1173** (1), 49-56 (1993).

97 Schoenberg, D. R. The end defines the means in bacterial mRNA decay. *Nat Chem Biol.* **3** (9), 535-536, doi:10.1038/nchembio0907-535, (2007).

98 Gallie, D. R. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes Dev.* **5** (11), 2108-2116 (1991).

99 Gingras, A. C., Raught, B. & Sonenberg, N. eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu Rev Biochem.* **68** 913-963, doi:10.1146/annurev.biochem.68.1.913, (1999).

100 Izaurralde, E. *et al.* A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell.* **78** (4), 657-668 (1994).

101 Izaurralde, E., Stepinski, J., Darzynkiewicz, E. & Mattaj, I. W. A cap binding protein that may mediate nuclear export of RNA polymerase II-transcribed RNAs. *J Cell Biol.* **118** (6), 1287-1295

(1992).
102 Mathonnet, G. *et al.* MicroRNA inhibition of translation initiation in vitro by targeting the cap-binding complex eIF4F. *Science.* **317** (5845), 1764-1767, doi:10.1126/science.1146067, (2007).
103 Caput, D. *et al.* Identification of a common nucleotide sequence in the 3'-untranslated region of mRNA molecules specifying inflammatory mediators. *Proc Natl Acad Sci U S A.* **83** (6), 1670-1674 (1986).
104 Klausner, R. D., Rouault, T. A. & Harford, J. B. Regulating the fate of mRNA: the control of cellular iron metabolism. *Cell.* **72** (1), 19-28 (1993).
105 Holtkamp, S. *et al.* Modification of antigen-encoding RNA increases stability, translational efficacy, and T-cell stimulatory capacity of dendritic cells. *Blood.* **108** (13), 4009-4017, doi:10.1182/blood-2006-04-015024, (2006).
106 Chang, H., Lim, J., Ha, M. & Kim, V. N. TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications. *Mol Cell.* **53** (6), 1044-1052, doi:10.1016/j.molcel.2014.02.007, (2014).
107 Brawerman, G. The Role of the poly(A) sequence in mammalian messenger RNA. *CRC Crit Rev Biochem.* **10** (1), 1-38 (1981).
108 Elango, N., Elango, S., Shivshankar, P. & Katz, M. S. Optimized transfection of mRNA transcribed from a d(A/T)100 tail-containing vector. *Biochem Biophys Res Commun.* **330** (3), 958-966, doi:10.1016/j.bbrc.2005.03.067, (2005).
109 Peng, J., Murray, E. L. & Schoenberg, D. R. In vivo and in vitro analysis of poly(A) length effects on mRNA translation. *Methods Mol Biol.* **419** 215-230, doi:10.1007/978-1-59745-033-1_15, (2008).
110 Gallie, D. R. A tale of two termini: a functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation. *Gene.* **216** (1), 1-11 (1998).
111 Michel, Y. M., Poncet, D., Piron, M., Kean, K. M. & Borman, A. M. Cap-Poly(A) synergy in mammalian cell-free extracts. Investigation of the requirements for poly(A)-mediated stimulation of translation initiation. *J Biol Chem.* **275** (41), 32268-32276, doi:10.1074/jbc.M004304200, (2000).
112 Newbury, S. F. Control of mRNA stability in eukaryotes. *Biochem Soc Trans.* **34** (Pt 1), 30-34, doi:10.1042/BST20060030, (2006).
113 Kariko, K., Buckstein, M., Ni, H. & Weissman, D. Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. *Immunity.* **23** (2), 165-175, doi:10.1016/j.immuni.2005.06.008, (2005).
114 Akira, S., Uematsu, S. & Takeuchi, O. Pathogen recognition and innate immunity. *Cell.* **124** (4), 783-801, doi:10.1016/j.cell.2006.02.015, (2006).
115 Kariko, K. *et al.* Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. *Mol Ther.* **16** (11), 1833-1840, doi:10.1038/mt.2008.200, (2008).
116 Nallagatla, S. R. & Bevilacqua, P. C. Nucleoside modifications modulate activation of the protein kinase PKR in an RNA structure-specific manner. *RNA.* **14** (6), 1201-1213, doi:10.1261/rna.1007408, (2008).
117 Uzri, D. & Gehrke, L. Nucleotide sequences and modifications that determine RIG-I/RNA binding and signaling activities. *J Virol.* **83** (9), 4174-4184, doi:10.1128/JVI.02449-08, (2009).
118 Kariko, K., Ni, H., Capodici, J., Lamphier, M. & Weissman, D. mRNA is an endogenous ligand for Toll-like receptor 3. *J Biol Chem.* **279** (13), 12542-12550, doi:10.1074/jbc.M310175200, (2004).
119 Malik, N. & Rao, M. S. A review of the methods for human iPSC derivation. *Methods Mol Biol.* **997** 23-33, doi:10.1007/978-1-62703-348-0_3, (2013).
120 Rao, M. S. & Malik, N. Assessing iPSC reprogramming methods for their suitability in translational

medicine. *J Cell Biochem.* **113** (10), 3061-3068, doi:10.1002/jcb.24183, (2012).

121 Weinberg, D. E. & Gross, J. D. Molecular biology: Messenger RNAs marked for longer life. *Nature.* **541** (7637), 293-294, doi:10.1038/nature21109, (2017).

122 Mauer, J. *et al.* Reversible methylation of m6Am in the 5' cap controls mRNA stability. *Nature.* **541** (7637), 371-375, doi:10.1038/nature21022, (2017).

123 Pardi, N. *et al.* Expression kinetics of nucleoside-modified mRNA delivered in lipid nanoparticles to mice by various routes. *J Control Release.* **217** 345-351, doi:10.1016/j.jconrel.2015.08.007, (2015).

124 Ferizi, M. *et al.* Stability analysis of chemically modified mRNA using micropattern-based single-cell arrays. *Lab Chip.* **15** (17), 3561-3571, doi:10.1039/c5lc00749f, (2015).

125 Leonhardt, C. *et al.* Single-cell mRNA transfection studies: delivery, kinetics and statistics by numbers. *Nanomedicine.* **10** (4), 679-688, doi:10.1016/j.nano.2013.11.008, (2014).

126 Phua, K. K., Leong, K. W. & Nair, S. K. Transfection efficiency and transgene expression kinetics of mRNA delivered in naked and nanoparticle format. *J Control Release.* **166** (3), 227-233, doi:10.1016/j.jconrel.2012.12.029, (2013).

127 Kariko, K., Muramatsu, H., Ludwig, J. & Weissman, D. Generating the optimal mRNA for therapy: HPLC purification eliminates immune activation and improves translation of nucleoside-modified, protein-encoding mRNA. *Nucleic Acids Res.* **39** (21), e142, doi:10.1093/nar/gkr695, (2011).

128 Zhou, L. Q. & Dean, J. Reprogramming the genome to totipotency in mouse embryos. *Trends Cell Biol.* **25** (2), 82-91, doi:10.1016/j.tcb.2014.09.006, (2015).

129 Schoorlemmer, J., Perez-Palacios, R., Climent, M., Guallar, D. & Muniesa, P. Regulation of Mouse Retroelement MuERV-L/MERVL Expression by REX1 and Epigenetic Control of Stem Cell Potency. *Front Oncol.* **4** 14, doi:10.3389/fonc.2014.00014, (2014).

130 Kigami, D., Minami, N., Takayama, H. & Imai, H. MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. *Biol Reprod.* **68** (2), 651-654 (2003).

131 Yoder, J. A., Walsh, C. P. & Bestor, T. H. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13** (8), 335-340 (1997).

132 Maoz, N. & Buganim, Y. Moving towards totipotency without a single miR-acle. *Cell Res.* **27** (5), 600-601, doi:10.1038/cr.2017.30, (2017).

133 Hamatani, T., Carter, M. G., Sharov, A. A. & Ko, M. S. Dynamics of global gene expression changes during mouse preimplantation development. *Dev Cell.* **6** (1), 117-131 (2004).

134 Rother, F. *et al.* Importin alpha7 is essential for zygotic genome activation and early mouse development. *PLoS One.* **6** (3), e18310, doi:10.1371/journal.pone.0018310, (2011).

135 Kim, J. *et al.* Maternal Setdb1 Is Required for Meiotic Progression and Preimplantation Development in Mouse. *PLoS Genet.* **12** (4), e1005970, doi:10.1371/journal.pgen.1005970, (2016).

136 Yu, C. *et al.* BTG4 is a meiotic cell cycle-coupled maternal-zygotic-transition licensing factor in oocytes. *Nat Struct Mol Biol.* **23** (5), 387-394, doi:10.1038/nsmb.3204, (2016).

137 Castillo, A. & Justice, M. J. The kinesin related motor protein, Eg5, is essential for maintenance of pre-implantation embryogenesis. *Biochem Biophys Res Commun.* **357** (3), 694-699, doi:10.1016/j.bbrc.2007.04.021, (2007).

138 Zheng, W. *et al.* Maternal phosphatidylinositol 3-kinase signalling is crucial for embryonic genome activation and preimplantation embryogenesis. *EMBO Rep.* **11** (11), 890-895, doi:10.1038/embor.2010.144, (2010).

139 Li, X. Y., Cui, X. S. & Kim, N. H. Transcription profile during maternal to zygotic transition in the mouse embryo. *Reprod Fertil Dev.* **18** (6), 635-645 (2006).

140 Guo, F. *et al.* Single-cell multi-omics sequencing of mouse early embryos and embryonic stem cells. *Cell Res.* **27** (8), 967-988, doi:10.1038/cr.2017.82, (2017).
141 Fan, X. *et al.* Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. *Genome Biol.* **16** 148, doi:10.1186/s13059-015-0706-1, (2015).
142 Shinagawa, T. *et al.* Histone variants enriched in oocytes enhance reprogramming to induced pluripotent stem cells. *Cell Stem Cell.* **14** (2), 217-227, doi:10.1016/j.stem.2013.12.015, (2014).
143 Jullien, J. *et al.* HIRA dependent H3.3 deposition is required for transcriptional reprogramming following nuclear transfer to Xenopus oocytes. *Epigenetics Chromatin.* **5** (1), 17, doi:10.1186/1756-8935-5-17, (2012).
144 Gaspar-Maia, A. *et al.* MacroH2A histone variants act as a barrier upon reprogramming towards pluripotency. *Nat Commun.* **4** 1565, doi:10.1038/ncomms2582, (2013).
145 Nashun, B., Akiyama, T., Suzuki, M. G. & Aoki, F. Dramatic replacement of histone variants during genome remodeling in nuclear-transferred embryos. *Epigenetics.* **6** (12), 1489-1497, doi:10.4161/epi.6.12.18206, (2011).
146 Kunitomi, A. *et al.* H1foo Has a Pivotal Role in Qualifying Induced Pluripotent Stem Cells. *Stem Cell Reports.* **6** (6), 825-833, doi:10.1016/j.stemcr.2016.04.015, (2016).
147 Yang, P., Wu, W. & Macfarlan, T. S. Maternal histone variants and their chaperones promote paternal genome activation and boost somatic cell reprogramming. *Bioessays.* **37** (1), 52-59, doi:10.1002/bies.201400072, (2015).
148 Lee, M. J. *et al.* Trichostatin A promotes the development of bovine somatic cell nuclear transfer embryos. *J Reprod Dev.* **57** (1), 34-42 (2011).
149 Maalouf, W. E. *et al.* Trichostatin A treatment of cloned mouse embryos improves constitutive heterochromatin remodeling as well as developmental potential to term. *BMC Dev Biol.* **9** 11, doi:10.1186/1471-213X-9-11, (2009).
150 Jeong, Y. I. *et al.* Effects of Trichostatin A on In vitro Development of Porcine Embryos Derived from Somatic Cell Nuclear Transfer. *Asian-Australas J Anim Sci.* **26** (12), 1680-1688, doi:10.5713/ajas.2013.13029, (2013).
151 Falco, G. *et al.* Zscan4: a novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells. *Dev Biol.* **307** (2), 539-550, doi:10.1016/j.ydbio.2007.05.003, (2007).
152 Wu, X. *et al.* Zygote arrest 1 (Zar1) is a novel maternal-effect gene critical for the oocyte-to-embryo transition. *Nat Genet.* **33** (2), 187-191, doi:10.1038/ng1079, (2003).
153 Wu, X., Wang, P., Brown, C. A., Zilinski, C. A. & Matzuk, M. M. Zygote arrest 1 (Zar1) is an evolutionarily conserved gene expressed in vertebrate ovaries. *Biol Reprod.* **69** (3), 861-867, doi:10.1095/biolreprod.103.016022, (2003).
154 Shinojima, Y. *et al.* Identification and analysis of an early diagnostic marker for malignant melanoma: ZAR1 intra-genic differential methylation. *J Dermatol Sci.* **59** (2), 98-106, doi:10.1016/j.jdermsci.2010.04.016, (2010).
155 Akiyama, T. *et al.* Transient bursts of Zscan4 expression are accompanied by the rapid derepression of heterochromatin in mouse embryonic stem cells. *DNA Res.* **22** (5), 307-318, doi:10.1093/dnares/dsv013, (2015).
156 Huynh, L. M., Shinagawa, T. & Ishii, S. Two Histone Variants TH2A and TH2B Enhance Human Induced Pluripotent Stem Cell Generation. *Stem Cells Dev.* **25** (3), 251-258, doi:10.1089/scd.2015.0264, (2016).
157 Marcho, C., Cui, W. & Mager, J. Epigenetic dynamics during preimplantation development. *Reproduction.* **150** (3), R109-120, doi:10.1530/REP-15-0180, (2015).
158 Messerschmidt, D. M., Knowles, B. B. & Solter, D. DNA methylation dynamics during epigenetic

reprogramming in the germline and preimplantation embryos. *Genes Dev.* **28** (8), 812-828, doi:10.1101/gad.234294.113, (2014).

159 Hayakawa, K., Ohgane, J., Tanaka, S., Yagi, S. & Shiota, K. Oocyte-specific linker histone H1foo is an epigenomic modulator that decondenses chromatin and impairs pluripotency. *Epigenetics.* **7** (9), 1029-1036, doi:10.4161/epi.21492, (2012).

160 Teranishi, T. *et al.* Rapid replacement of somatic linker histones with the oocyte-specific linker histone H1foo in nuclear transfer. *Dev Biol.* **266** (1), 76-86 (2004).

161 Tamada, H. *et al.* Chromatin decondensation and nuclear reprogramming by nucleoplasmin. *Mol Cell Biol.* **26** (4), 1259-1271, doi:10.1128/MCB.26.4.1259-1271.2006, (2006).

162 Amano, T. *et al.* Zscan4 restores the developmental potency of embryonic stem cells. *Nat Commun.* **4** 1966, doi:10.1038/ncomms2966, (2013).

163 Lin, C. J., Koh, F. M., Wong, P., Conti, M. & Ramalho-Santos, M. Hira-mediated H3.3 incorporation is required for DNA replication and ribosomal RNA transcription in the mouse zygote. *Dev Cell.* **30** (3), 268-279, doi:10.1016/j.devcel.2014.06.022, (2014).

164 Ma, J. Y., Zhang, T., Shen, W., Schatten, H. & Sun, Q. Y. Molecules and mechanisms controlling the active DNA demethylation of the mammalian zygotic genome. *Protein Cell.* **5** (11), 827-836, doi:10.1007/s13238-014-0095-3, (2014).

165 Gaume, X. & Torres-Padilla, M. E. Regulation of Reprogramming and Cellular Plasticity through Histone Exchange and Histone Variant Incorporation. *Cold Spring Harb Symp Quant Biol.* **80** 165-175, doi:10.1101/sqb.2015.80.027458, (2015).

166 Sonehara, H., Nagata, M. & Aoki, F. Roles of the first and second round of DNA replication in the regulation of zygotic gene activation in mice. *J Reprod Dev.* **54** (5), 381-384 (2008).

167 Svoboda, P., Franke, V. & Schultz, R. M. Sculpting the Transcriptome During the Oocyte-to-Embryo Transition in Mouse. *Curr Top Dev Biol.* **113** 305-349, doi:10.1016/bs.ctdb.2015.06.004, (2015).

168 Inoue, K. *et al.* Trichostatin A specifically improves the aberrant expression of transcription factor genes in embryos produced by somatic cell nuclear transfer. *Sci Rep.* **5** 10127, doi:10.1038/srep10127, (2015).

169 Wang, B. *et al.* Reprogramming efficiency and quality of induced Pluripotent Stem Cells (iPSCs) generated from muscle-derived fibroblasts of mdx mice at different ages. *PLoS Curr.* **3** RRN1274, doi:10.1371/currents.RRN1274, (2011).

170 Trokovic, R., Weltner, J., Noisa, P., Raivio, T. & Otonkoski, T. Combined negative effect of donor age and time in culture on the reprogramming efficiency into induced pluripotent stem cells. *Stem Cell Res.* **15** (1), 254-262, doi:10.1016/j.scr.2015.06.001, (2015).

171 Morgani, S. M. *et al.* Totipotent embryonic stem cells arise in ground-state culture conditions. *Cell Rep.* **3** (6), 1945-1957, doi:10.1016/j.celrep.2013.04.034, (2013).

172 Li, H. *et al.* The Ink4/Arf locus is a barrier for iPS cell reprogramming. *Nature.* **460** (7259), 1136-1139, doi:10.1038/nature08290, (2009).

173 Nichols, J. *et al.* Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell.* **95** (3), 379-391 (1998).

174 Liu, L., Czerwiec, E. & Keefe, D. L. Effect of ploidy and parental genome composition on expression of Oct-4 protein in mouse embryos. *Gene Expr Patterns.* **4** (4), 433-441, doi:10.1016/j.modgep.2004.01.004, (2004).

175 Khan, D. R. *et al.* Expression of pluripotency master regulators during two key developmental transitions: EGA and early lineage specification in the bovine embryo. *PLoS One.* **7** (3), e34110, doi:10.1371/journal.pone.0034110, (2012).

176 Niakan, K. K. & Eggan, K. Analysis of human embryos from zygote to blastocyst reveals distinct

gene expression patterns relative to the mouse. *Dev Biol.* **375** (1), 54-64, doi:10.1016/j.ydbio.2012.12.008, (2013).

177 Blakeley, P. *et al.* Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development.* **142** (20), 3613, doi:10.1242/dev.131235, (2015).

178 Ralston, A. & Rossant, J. Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo. *Dev Biol.* **313** (2), 614-629, doi:10.1016/j.ydbio.2007.10.054, (2008).

179 Borsos, M. & Torres-Padilla, M. E. Building up the nucleus: nuclear organization in the establishment of totipotency and pluripotency during mammalian development. *Genes Dev.* **30** (6), 611-621, doi:10.1101/gad.273805.115, (2016).

180 Ooga, M., Fulka, H., Hashimoto, S., Suzuki, M. G. & Aoki, F. Analysis of chromatin structure in mouse preimplantation embryos by fluorescent recovery after photobleaching. *Epigenetics.* **11** (1), 85-94, doi:10.1080/15592294.2015.1136774, (2016).

181 Jevtic, P. & Levy, D. L. Nuclear size scaling during Xenopus early development contributes to midblastula transition timing. *Curr Biol.* **25** (1), 45-52, doi:10.1016/j.cub.2014.10.051, (2015).

182 Blasco, M. A. Telomere length, stem cells and aging. *Nat Chem Biol.* **3** (10), 640-649, doi:10.1038/nchembio.2007.38, (2007).

183 Niida, H. *et al.* Severe growth defect in mouse cells lacking the telomerase RNA component. *Nat Genet.* **19** (2), 203-206, doi:10.1038/580, (1998).

184 Allsopp, R. Telomere length and iPSC re-programming: survival of the longest. *Cell Res.* **22** (4), 614-615, doi:10.1038/cr.2012.6, (2012).

185 Aguado, T. *et al.* Telomere Length Defines the Cardiomyocyte Differentiation Potency of Mouse Induced Pluripotent Stem Cells. *Stem Cells.* **35** (2), 362-373, doi:10.1002/stem.2497, (2017).

186 Liu, L. *et al.* Telomere lengthening early in development. *Nat Cell Biol.* **9** (12), 1436-1441, doi:10.1038/ncb1664, (2007).

187 Zalzman, M. *et al.* Zscan4 regulates telomere elongation and genomic stability in ES cells. *Nature.* **464** (7290), 858-863, doi:10.1038/nature08882, (2010).

188 Zhang, Q. *et al.* Tcstv1 and Tcstv3 elongate telomeres of mouse ES cells. *Sci Rep.* **6** 19852, doi:10.1038/srep19852, (2016).

189 Kalmbach, K., Robinson, L. G., Jr., Wang, F., Liu, L. & Keefe, D. Telomere length reprogramming in embryos and stem cells. *Biomed Res Int.* **2014** 925121, doi:10.1155/2014/925121, (2014).

190 Keefe, D. L., Liu, L. & Marquard, K. Telomeres and aging-related meiotic dysfunction in women. *Cell Mol Life Sci.* **64** (2), 139-143, doi:10.1007/s00018-006-6466-z, (2007).

191 Kalmbach, K. H. *et al.* Telomeres and human reproduction. *Fertil Steril.* **99** (1), 23-29, doi:10.1016/j.fertnstert.2012.11.039, (2013).

192 Wang, F. *et al.* Robust measurement of telomere length in single cells. *Proc Natl Acad Sci U S A.* **110** (21), E1906-1912, doi:10.1073/pnas.1306639110, (2013).

193 Schaetzlein, S. *et al.* Telomere length is reset during early mammalian embryogenesis. *Proc Natl Acad Sci U S A.* **101** (21), 8034-8038, doi:10.1073/pnas.0402400101, (2004).

194 Treff, N. R., Su, J., Taylor, D. & Scott, R. T., Jr. Telomere DNA deficiency is associated with development of human embryonic aneuploidy. *PLoS Genet.* **7** (6), e1002161, doi:10.1371/journal.pgen.1002161, (2011).

195 Lin, S. P. *et al.* Survival of cancer stem cells under hypoxia and serum depletion via decrease in PP2A activity and activation of p38-MAPKAPK2-Hsp27. *PLoS One.* **7** (11), e49605, doi:10.1371/journal.pone.0049605, (2012).

196 Michalopoulou, E., Bulusu, V. & Kamphorst, J. J. Metabolic scavenging by cancer cells: when the

going gets tough, the tough keep eating. *Br J Cancer.* **115** (6), 635-640, doi:10.1038/bjc.2016.256, (2016).

197 Seluanov, A. *et al.* Hypersensitivity to contact inhibition provides a clue to cancer resistance of naked mole-rat. *Proc Natl Acad Sci U S A.* **106** (46), 19352-19357, doi:10.1073/pnas.0905252106, (2009).

198 De Iaco, A. *et al.* DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat Genet.* **49** (6), 941-945, doi:10.1038/ng.3858, (2017).

199 Choi, Y. J. *et al.* Deficiency of microRNA miR-34a expands cell fate potential in pluripotent stem cells. *Science.* **355** (6325), doi:10.1126/science.aag1927, (2017).

200 Choi, Y. J. *et al.* miR-34 miRNAs provide a barrier for somatic cell reprogramming. *Nat Cell Biol.* **13** (11), 1353-1360, doi:10.1038/ncb2366, (2011).

201 Choi, S. H. *et al.* DUX4 recruits p300/CBP through its C-terminus and induces global H3K27 acetylation changes. *Nucleic Acids Res.* **44** (11), 5161-5173, doi:10.1093/nar/gkw141, (2016).

202 Sun, Y. *et al.* Zelda overcomes the high intrinsic nucleosome barrier at enhancers during Drosophila zygotic genome activation. *Genome Res.* **25** (11), 1703-1714, doi:10.1101/gr.192542.115, (2015).

203 Schulz, K. N. *et al.* Zelda is differentially required for chromatin accessibility, transcription factor binding, and gene expression in the early Drosophila embryo. *Genome Res.* **25** (11), 1715-1726, doi:10.1101/gr.192682.115, (2015).