

Spring 5-6-2012

# Generating Induced Pluripotent Stem Cells from Human Fibroblast Utilizing Controlled Expression of Transcription Factors

Ju-Sung Song

*University of Connecticut - Storrs, [songjs712@gmail.com](mailto:songjs712@gmail.com)*

Follow this and additional works at: [https://opencommons.uconn.edu/srhonors\\_theses](https://opencommons.uconn.edu/srhonors_theses)



Part of the [Medical Cell Biology Commons](#), and the [Pharmacy and Pharmaceutical Sciences Commons](#)

---

## Recommended Citation

Song, Ju-Sung, "Generating Induced Pluripotent Stem Cells from Human Fibroblast Utilizing Controlled Expression of Transcription Factors" (2012). *Honors Scholar Theses*. 230.

[https://opencommons.uconn.edu/srhonors\\_theses/230](https://opencommons.uconn.edu/srhonors_theses/230)

**Generating Induced Pluripotent Stem Cells from Human  
Fibroblast Utilizing Controlled Expression of Transcription Factors**

**Ju-Sung Song**

Principal Investigator: Theodore P. Rasmussen

Department of Pharmaceutical Science

School of Pharmacy Honors Program

University of Connecticut

## **Table of Contents**

Abstract.....	3
Background and Introduction.....	3~8
Materials and Methods.....	8~12
Results.....	12~17
Discussion.....	17~21
Figures and Tables.....	22~36
References.....	37

**Abstract:**

Human Embryonic stem cells (hESCs) represent a pluripotent cell population derived from the inner cell mass (ICM) of the blastocyst stage of the developing embryo. Ethical concerns have been raised regarding the derivation process, as the procedure ultimately results in the destruction of the embryos. Recently, alternative approach has been devised that encompasses reprogramming of terminally differentiated cells into an hESC-like state. This process relies on the over-expression of four pluripotency-associated factors, and successfully reprogrammed cells are termed induced pluripotent stem cells (iPSCs). These cells hold great promise for the use in future cell-based therapeutics such as evaluating drug induced toxicity and developing personalized medicines. However, the intricacies of the reprogramming mechanism are not fully understood, mostly due to the random nature of the expression of the four factors. In this study, I propose a series of experiments involving a novel molecular switch, known as Rheostat system, along with the doxycycline inducible system, that will regulate the expression of the four transcription factors in a precise temporal manner. Controlled expression of the transcription factors will help to elucidate underlying molecular mechanisms of the reprogramming process.

**Background and Introduction:**

Stem cells can be categorized into two major categories according to their potency of differentiation: adult stem cells and pluripotent stem cells.<sup>1</sup> Adult stem cells, such as hematopoietic and mesenchymal stem cells, generally have limited potential for growth and differentiation. In contrast, pluripotent stem cells such as embryonic stem cells (ESCs) have the capacity to give rise to any type of cell in the body, and to grow indefinitely in the laboratory setting.<sup>1</sup> Therefore, pluripotent stem cells offer greater potential to be used therapeutically than

adult stem cells.<sup>1,2</sup> For example, ESCs provide indefinite resources for generating somatic cells that can be employed in cell-based regenerative therapies. This is of particular relevance for diseases such as Parkinson's disease, leukemia and muscular dystrophy in which cell transplantation can functionally surrogate the tissue damages that are caused by environmental agents, mutations or injury.<sup>1,2</sup> Unfortunately, ESC research is currently obstructed by two major issues. One is the ethical issue associated with the derivation of human ESCs from human fertilized eggs, and the other is the immunological incompatibility between ESC-derived donor organs or cells and the recipients because of differences in the histocompatibility factor.<sup>2</sup>

The development of Induced Pluripotent Stem cells (iPSCs), pioneered by Takahashi and Yamanaka, have advanced stem cell research by circumventing these issues.<sup>3</sup> Unlike the derivation of ESCs from ICM, iPSCs can be obtained by reprogramming readily accessible patient's somatic cells with either sets of four transcription factors, Oct4, Sox2, cMyc, and Klf4 or Oct4, Nanog, Sox2 and Lin28.<sup>3</sup> These reprogrammed cells are in an ESC-like state that convey most of the therapeutic values that ESCs offer.<sup>3</sup> Because iPSCs are derived from each individual, they eliminated histocompatibility issues that were previously associated with ESCs.<sup>4</sup> Therefore, iPSCs can be differentiated into certain tissues and be transplanted into patients with no potential risk of immune rejection.<sup>5</sup> These tissues can also be used in *in-vitro* studies to elucidate patient-specific disease pathogenesis and discover methods yielding the best course of treatment.<sup>5</sup>

Use of iPSC technology is also useful in predicting the therapeutic benefit of certain medications. Not all medications have an equal effect in patients.<sup>6</sup> Some patients receive a beneficial effect from medication while others do not and small subsets of patients exhibit undesirable side effects that pose serious consequences.<sup>5</sup> Those side effects often occur in

cardiac, hepatic, and neural cells and they can be fatal.<sup>5</sup> Unfortunately, those side effects are often idiosyncratic and are discovered after patients have taken the medication.<sup>7</sup> By using patient-specific iPSCs production followed by differentiation into cardiac, hepatic and neural cells, it is now possible to determine individualized patient responses to a certain medication prior to administration.<sup>7</sup> By conducting several drug efficacy tests in those major tissues *in-vitro*, it is practicable to predict which medication would have the best therapeutic effect and have the least risk of adverse reaction on an individual basis.<sup>6,7</sup> Screening of iPSCs may also lead to the discovery of genetic defects in the patients and can help find the most effective interaction of a drug with those defective genes to eliminate certain physiological effect on patients.<sup>8</sup> Therefore, patients can receive personalized medication that will be most effective through the usage of iPSC technology. iPSC technology is certainly a key application that has many beneficial uses and is a key to amend the technical and ethical problems associated with ESC research.

Despite their great therapeutic promise, the value of iPSCs is hampered by the prolonged duration of more than 21 days for reprogramming and by a low success rate of less than 0.1 percent.<sup>6</sup> Also, increased mutational loads that have accumulated in somatic cells over the patient's life span are found to be responsible for varying results of reprogramming process.<sup>7</sup> There have been intense studies to understand the underlying transcriptional and chromatin-based mechanisms during reprogramming process.<sup>8</sup> However, there are still many questions that are left unanswered. The answers to these questions will be crucial for improving the reprogramming efficiency and further advancing the therapeutic application of iPSCs. Therefore, I focused to study the temporal requirement of expression of each gene during the course of the reprogramming process by utilizing novel inducible molecular switch, Rheostat system, along with doxycycline system to expound the way to increase the efficiency in generating iPSCs. The

temporal tuning of factor expression permits to analyze at which time span a particular gene is most active. This, in turn, will provide information about a specific time period when cells need to be induced to express specific genes.

### **Preliminary Data:**

Currently, our lab utilizes a doxycycline-dependent system to regulate the expression of four transcription factors (Figure 1). This system ensures simultaneous expression of the four genes. All transcription factors are induced three days after the transfection by the addition of doxycycline and the induction is ceased fifteen days later by the removal of doxycycline. This system may be more efficient in inducing gene expressions, but it does not reveal any specific information about how each transcription factor affects the reprogramming process and may be less efficient in generating iPSCs. Controlling gene expression by more than one gene switch is advantageous because each gene expression can be separately induced. Therefore, if each transcription factor is regulated by an independent molecular switch, the effects of each gene during the reprogramming process can be observed by the several induction and discharge of reprogramming genes during the of reprogramming process which would ultimately elucidate the temporal requirement of gene expression in various time periods.

### **Rheostat System**

The Rheostat system is a novel ecdysone receptor (EcR)-based gene switch, which can be implicated in various molecular applications such as gene therapy, recombinant protein production, functional genomics, and regulation of transgenic gene expression.<sup>11,12</sup> EcR is a member of the nuclear receptor superfamily that plays an important role in regulating development and reproduction in insects.<sup>11</sup> In the case of insects and some plants, the application of an EcR-based gene switch could be limited by the presence of endogenous ecdysteroids that

could potentially activate the system.<sup>12</sup> However, an EcR-based gene switch is attractive for mammalian application because both the receptor and ligand are foreign to mammalian cells.<sup>12</sup> Absence of endogenous ecdysteroid in mammalian cells lowers the potential to cross-activate the system, however, it was reported that exogenous steroid such as growth hormone present in media or endogenous steroid produced within the mammalian cells can also have some affinity to EcR and can activate transcriptional response.<sup>12</sup> Therefore, the Rheostat system was developed in order to decrease receptor affinity to steroidal ligands but retain high affinity to a synthetic nonsteroidal ligand so that the expression can solely be controlled by the presence of synthetic ligand.<sup>12</sup> The mutation of a critical alanine residue, A110, to proline in the ligand binding domain of EcR was introduced, leading to a selective disruption of both binding and transactivation with steroids but not with nonsteroidal ligand.<sup>12,13</sup> A synthetic non-steroidal ligand, known as RSG-1, was found to have high affinity to this single point-mutant receptor and was shown to activate ecdysone responsive genes effectively.<sup>13</sup> Furthermore, the target gene expression was revealed to be increased in ligand dose dependent manner and can repeatedly turned on and off by the addition and removal of ligand.<sup>12</sup> Therefore, Rheostat system seems to be suitable gene switch machinery, which can be utilized in conjunction with the doxycycline-dependent system in studying temporal tuning of reprogramming genes.

The Rheostat system is comprised of three components: Rheo Receptor (Rheo R), Rheo Activator (Rheo A), and Rheo Response Element (RRE). Rheo R, in concert with RSG-1 heterodimerizes with Rheo A and moves across the nuclear membrane. The binding of the two proteins transactivates and binds to RRE to generate gene expression (Figure 2).<sup>13</sup> In order to utilize this system in our fibroblasts, three components of Rheo system were cloned into two



lentiviral vector, FU-ΔGWMCS-T2A and FU-ΔeGFP-T2A-LUC. Concise overview of construct preparation is explained below.

## **Materials and Methods:**

### **Cell Culture**

The DNA plasmids in pNERB-R1 backbone containing gene segments of Rheo R and Rheo A were purchased from Milipore stably expressed in 293T cells (Figure 3). The 293T cells were cultured on 10cm-diameter culture plate in 1X DMEM medium, supplemented with 10% fetal bovine serum, 0.6% Geneticin 418 and 0.5% Doxycycline. Within 24-48 h, the cells achieved greater than 80% confluence, and were dissociated by 0.05% trypsin-EDTA and passaged onto three 10cm-diameter-culture dishes to cultivate high yield. Cells were harvested at 80% confluency by dissociation with 0.05% trypsin-EDTA and were pelleted by five minutes of centrifugation at 1,000rpm.

### **Extraction of genomic DNA**

The genomic DNA was extracted by lysing cell pellet with 3 ml of lysis buffer (50mM Tris-HCl in pH 8, 50 mM EDTA, 1% SDS and 10mM NaCl). Lysed cells were centrifuged at 13,100 rpm for 10 minutes to remove cell debris. Fifteen µl of Proteinase K (NEB, Ipswich, MA) was added to the supernatant and was incubated for 18 hours at 37°C. Three ml of phenol was then added to the supernatant and was placed in the shaker for 12 hours. The supernatant was centrifuged at 13,100 rpm for 15 minutes at 4°C and the bottom phase of the supernatant was removed with a P1000 pipette. Three ml of phenol were added to the residual supernatant and the procedure was repeated once again. Three ml of chloroform were added to the upper phase of the supernatant from the phenol extraction, and was placed on the shaker for 12 hours. The upper phase of the supernatant was discarded by P1000 pipette and 2X volume of ethanol and 0.1X

volume of sodium acetate were added to precipitate genomic DNA. Precipitated genomic DNA was diluted in 1M of Tris-EDTA buffer (pH 8) and was stored in 4°C.

## **PCR**

Rheo R and Rheo A were amplified by polymerase chain reaction (PCR) using Biorad's Icyler. Primers were designed to recognize each segments of Rheo R and Rheo A in the pNERB-R1 vector and the sequences of the primers are shown in Table 1 with their respective melting points. Two separate PCR were performed for Rheo R and Rheo A. For both reactions, PCR was prepared by adding 10 µM of its forward and reverse primers, 5 µl of 10X Buffer, 10 mM of dNTPs and 1 µg of Taq polymerase to 1 µg of genomic DNA. PCR was carried out in 30 cycles of 60, 60, and 90 seconds at 95°C, 58°C, and 68°C and an additional 10 minutes of extension period was provided at the end of the reaction at 72°C. For Rheo A, PCR was carried out in 30 cycles of 60, 60, and 90 seconds at 95°C, 60°C, and 68°C, and an additional 10 minutes of extension period was provided at the end of the reaction at 72°C as well.

## **Electrophoresis**

The agarose gel was prepared by adding 400mg of SeaKem LE Agarose (Lonza, Rockland, ME) in 50ml of 1X TAE (Bio-Rad, Hercules, CA) and the addition of 5 µg of fluorescent dye, Ethidium Bromide (Bio-rad, Hercules, CA). The agarose was dissolved in TAE by heating up to boiling point, and was cooled down in 50°C water bath. The agarose was then poured into 8cm x 10cm gel box with the comb properly inserted. The 2X volume of bromophenol blue was added to the PCR products, and was loaded into the wells after the gel was solidified. The gel was transferred to electrophoresis machine containing a sufficient volume of 1X TAE running buffer to submerge the gel. Electrophoresis was performed at 100V until

bromophenol blue has run  $\frac{3}{4}$  the length of the gel. The gel was then observed under Benchtop 3UV<sup>TM</sup> Transilluminator.

### **Gel Extraction**

All gel extraction was performed by using QIAquick gel extraction Kit (Qiagen, Valencia, CA), following the manufacturer's instruction.

### **TOPO Cloning and Transformation**

The PCR products were introduced into pCR2.1-TOPO vector using Invitrogen's TOPO-TA cloning kit (Figure 5, Invitrogen, [www.invitrogen.com](http://www.invitrogen.com)). The ligated TOPO vector were then transformed into TOP10 chemically competent E.coli cells using standard protocol and plated onto 10cm diameter LB plates supplemented with, 100µg of ampicillin, 50 µg of X-GAL and IPTG. Ten colonies were picked including one light blue colony for negative control after 12 hours of incubation at 37°C. Each colony was transferred into the glass tube in 5ml of LB broth, supplemented with 100µg of ampicillin and placed in the shaker for 12-14 hours at 37°C. All ten samples were centrifuged for 3 minutes at 13,100 rpm and were pelleted for genomic DNA extraction.

### **DNA Extraction**

All DNA of the samples were extracted by using QIAprep Spin Miniprep kit (Qiagen, Valencia, CA). The procedure was executed following manufacturer's instruction.

### **Restriction Enzyme Digestion**

All restriction enzymes used in this experiment were purchased from BioLab, Inc. and was performed following manufacturer's instruction. One µg of the samples was incubated for 90~120 minutes at 37°C and were separated by electrophoresis.

### **Ligation**

All ligation of plasmids and vectors were carried out by using T4 DNA ligase (BioLab, Inc., 400,000U/ml). The ligation reactions were comprised of 1 µg of lentiviral vector, 1 µg of Rheo R or Rheo A that were diluted into 1:10, 1:100, and 1:500 ratios, 1 µl of 10X Buffer for T4 DNA Ligase with 10 mM ATP, and 1 µl of T4 DNA ligase. The ligation reactions were incubated at 16°C for 12-16 hours by using Programmable Thermal Controller (MJ Research, Inc. [www.mj-research.com](http://www.mj-research.com)) and were transformed into HB101 competent cells. The RRE was also ligated into the vector following the same procedure.

### **Transfection**

293T cells were passaged 12 hours prior transfection to a 6-wells plate at a density of  $3 \times 10^5$  cells per well. Cells were transfected with each of the construct by using TransIT-2020 transfection protocol (Mirus, Madison, WI). Two and half µg of each construct were separately diluted in 250 µl of 1X Opti-MEM I Reduced-Serum Media (Gibco, [www.invitrogen.com](http://www.invitrogen.com)) in a sterile tube and 7.5 µl of TransIT-2020 Reagent was added. The mixtures were then vortexed to mix completely and were incubated at room temperature for 30 minutes. After 30 minutes, mixtures were added drop-wise to different areas of the wells to complete transfection. Cells were induced with different concentration of RSG-1 (5nM, 50nM, 500nM and 1000nM) 12 hours after the transfection and were harvested 36 hours after for luciferase assay.

### **RSG-1**

RSG-1 was purchased from Milipore<sup>TM</sup> (Catalog number: 20-382) and was supplied in 5mM in 100% DMSO. RSG-1 was diluted in cell media to concentrations of 5nM, 50nM, 500nM and 1000nM and was stored at 4°C. Manufacturer reported that the ligand can be stored in cell media at 4°C for several weeks with no loss of activity and that media containing the

ligand can be added to cells in culture and incubated at 37°C for many days while still retaining full functionality.

### **Luciferase Assay**

Luciferase activity was measured by L-Max Molecular Devices 48 hours after the induction. The cells were scraped from the plates and their cell counts were calculated by using hemocytometer. Cells were pelleted by 5 minutes of centrifugation at 13,100 rpm and were lysed with 200 µl of 1X passive lysis buffer (Promega, Madison, WI). Cell debris was removed by 10 minutes of centrifugation at 13,100 rpm and 8 µl of supernatant was mixed with 50 µl of Luciferase Assay Reagent right before the reading. All experiments were carried out in triplicate and the average values were computed in graph. All luciferase activity was normalized by their cell counts that were calculated prior harvesting.

### **eGFP Activity**

The transfected cells were observed under the fluorescent microscope, Olympus IX51, to detect eGFP activity prior to harvesting for luciferase assay.

## **Results:**

### **Subcloning of Rheo R and Rheo A**

The DNA plasmids in pNERB-R1 backbone containing gene segments of Rheo R and Rheo A were purchased from Milipore stably expressed in 293T cells (Figure 3). Cells were cultured and harvested and their genomic DNA was extracted by several phenol and chloroform extractions. Rheo R and Rheo A were amplified by two separate PCR and the PCR products were separate by electrophoresis (Figure 4). The PCR products were excised from the gel and extracted by using Qiagen's gel extraction kit. Purified PCR products were then introduced to PCR2.1-TOPO vector using standard protocol (Figure 5). The ligated TOPO vectors were

transformed into TOP10 chemically competent cells and cultured in LB plates as described above. The 10 colonies were picked including one light blue colony for negative control and were cultured in LB broth for 16 hours. The cells were centrifuged and their genomic DNA were extracted by using Qiagen's miniprep kit. Three  $\mu\text{g}$  of each genomic DNA from 10 samples were digested with 20 units of EcoRI restriction enzyme (10,000U/ml) for 90 minutes at 37 °C and were separated by electrophoresis in 0.8% gel. The proper insertions of Rheo R and Rheo A into pCR2.1-TOPO vector are confirmed by observing 1000bp and 1500bp fragments excised from the vectors (Figure 6). All genomic DNA samples that showed proper insertions were sequenced in order to detect any mutation that may have occurred during PCR and transformation. Their sequences were compared to the parent Rheo R and Rheo A sequences obtained from Genbank and analyzed by using ApE software. For Rheo R, various mismatches and 17 bp insertion near the 3' end were discovered in all 9 samples. The common mismatches were at A678, T838, and A1222 residues, which caused the mutation to threonine, glycine, and glycine respectively. Also, a 17 bp insertion near the 3' end generated a protein truncation by 10 amino acids. Since, all nine samples contained similar mutation patterns, it was deduced that the published sequence from Genbank was not reflecting the Rheo R sequence used by the manufacturer. The Rheo R sample that contained lowest number of mismatches was selected to clone into the lentiviral construct, FU- $\Delta$ GWMCs-T2A (Figure 7). For Rheo A, all samples carried at least one mismatch at 171bp region that caused the mutation of glycine to threonine. Since, this mutation was observed in all other samples, it was deduced that it was likely due to an error in the published sequence. The Rheo A sample that only contained one mismatch at 171bp region was selected to be cloned into FU- $\Delta$ GWRR-T2A (Figure 8).

### **Lentivector Construction**

The lentiviral construct, FU-ΔGWMCS-T2A plasmid, was provided by Dr. Krueger. This plasmid contains human ubiquitous promoter (hUBC), multiple cloning site (MCS), and T2A peptide (Figure 9). T2A is one of the “self-cleaving” 2A peptides that enable dual expression of the gene from single promoter through a ribosomal skip mechanism. Therefore, Rheo R and Rheo A were cloned into our lentiviral construct separated by T2A peptide for bicistronic expression of Rheo R and Rheo A in regulation of single promoter, hUBC.

### **Rheo R into FU-ΔGWMCS-T2A**

Rheo R was excised by 20 units of NheI (10,000U/ml) and StuI (10,000U/ml) restriction enzymes from the 5 µg of pCR2.1-TOPO vector and was separated and purified by electrophoresis and Qiagen’s gel extraction kit (Figure 10B). Five µg of the lentiviral construct, FU-ΔGWMCS-T2A, was cut by 20 units of NheI (10,000U/ml) and HpaI (10,000U/ml) restriction enzymes and purified by gel extraction using Qiagen’s gel extraction kit (Figure 10A). These two constructs were then ligated by the T4 DNA ligase and were transformed into HB101 competent cells. Cells were cultured and picked in the same condition as described before. Their genomic DNA were extracted using Qiagen’s miniprep kit. The resulting plasmid was designated as FU-ΔGWRR-T2A (Figure 11C) and proper insertion of Rheo R was confirmed by SmaI digestion (Figure 11D).

### **Rheo A into FU-ΔGWRR-T2A**

Both Rheo A and FU-ΔGWRR-T2A contain Sfi I restriction enzyme recognition site and were both excised and extracted by using Qiagen’s gel extraction kit (Figure 12). Two extracted plasmids were ligated by T4 DNA ligase and transformed into HB101 competent cells. Cells were grown and harvested and their genomic DNA was extracted by using Qiagen’s miniprep

kit. The resulting plasmid was designated as FU- $\Delta$ GWRR-T2A-RA (Figure 13C) and the proper insertion of Rheo A was confirmed by ScaI digestion (Figure 13D).

### **Rheo Response Element**

RRE was not provided from the manufacturer and it was constructed by hybridizing four oligonucleotides that were synthesized based on the sequence data obtained from Genbank. The schematic drawing of RRE oligonucleotides are shown in Figure 14A and the sequences of the four oligonucleotides are available in Table 2. The 5  $\mu$ l of each of the four oligonucleotides (100 $\mu$ M) were combined along with 5  $\mu$ l of 10X NEB 2 buffer and 25  $\mu$ l of water to bring the final volume to 50  $\mu$ l. These nucleotides were annealed by nucleic acid hybridization which was placed into the beaker containing 95°C water that was cooled down to room temperature over night. Annealed RRE was then cloned into FU- $\Delta$ GW-eGFP-T2A-LUC vector, which was provided from Dr. Gurevich in our lab. FU- $\Delta$ GW-eGFP-T2A-LUC vector was excised by PacI and BamHI digestion to remove hUBC from the vector and was replaced with RRE (Figure 14). RRE and digested FU- $\Delta$ GW-eGFP-T2A-LUC vector were transformed into HB101 competent cells. The cells were cultured and picked in the same condition as discussed earlier and their genomic DNA was extracted by using Qiagen's miniprep. The confirmation of RRE insertion into FU- $\Delta$ GW-eGFP-T2A-LUC was verified by ScaI digestion and the resulting plasmid was designated as FU- $\Delta$ RRE-eGFP-T2A-LUC (Figure 15). The samples that were confirmed the proper insertion, were processed to be sequenced in order to verify accurate nucleic acid hybridization and transformation. No mismatch was found in all samples (Figure 16).

### **Confirmation of Rheostat gene switch mechanism**

293 T cells were transfected with both FU- $\Delta$ RRE-eGFP-T2A-LUC and FU- $\Delta$ GWRR-T2A-RA vectors and were induced with different concentration of RSG-1 to confirm proper



transfection and integration of the two lentiviral constructs within the genome. The cells were observed under the fluorescent microscope to detect eGFP activity (Figure 17). The eGFP activity was shown to increase in a ligand-dependent manner; the cells induced with higher concentration of ligand showed higher activity of eGFP activity. There was not any eGFP activity observed in the cells that were not transfected (NT). However, some eGFP activity was also observed from the transfected cells that were not induced with RSG-1 (0nM). In order to quantify the gene expression activity of the Rheostat system, all cells were harvested to measure their luciferase activity. Luciferase data of all cells in different experiment conditions in triplicate are shown in Table 3. Luciferase assay showed the dose-dependent increase in luciferase activity as well (Figure 18). Highest luciferase activity was observed from the cells induced with 1000nM and the lowest from the cells induced with 5nM. Luciferase activity was also detected in the negative control sample as expected (0 nM). To further test this abnormal expression, another luciferase assay was performed to see how each of the constructs has an effect on abnormal expression. Cells were only transfected with either construct and were induced with 500nM of RSG-1 prior to harvesting for luciferase activity (Figure 19). Luciferase activity was only detected in the cells that had RRE transfected. No luciferase activity was detected in the cells that were only transfected with FU- $\Delta$ GWRR-T2A-RA. This conferred that FU- $\Delta$ RRE-eGFP-T2A-LUC itself can induce gene expression when integrated into 293T cells without FU- $\Delta$ GWRR-T2A-RA. It also showed that Rheo R and Rheo A are functioning accurately, since the luciferase activity was increased in a ligand-dependent manner as it was observed in previous experiment and the expression was halted when RRE was not present. It is not clear whether exogenous steroids in cell media activate certain nuclear receptors that share similar binding

domain as RRE to bind and initiate gene expression or RRE itself initiate expression without any regulatory control.

In order to elucidate whether other exogenous chemicals such as growth hormones present in FBS serum effects expression, cells were cultured in four different concentration of serum (10%, 7.5%, 5%, 2.5%) prior to and after the transfection with FU- $\Delta$ RRE-eGFP-T2A-LUC . Cells were not induced with RSG-1 and their luciferase activity were measured 36 hours after the transfection (Figure 20). There was a slight reduction in luciferase activity from the cells grown in lower concentration of media but that reduction did not correlate with decrease in serum concentration. The lowest luciferase activity was detected in cells grown in 2.5% serum concentration, and the highest in 10%. However, there was slightly higher expression of luciferase activity in cells grown in 5% than the ones from 7.5%. There was some variability in cell counts especially with the cells grown in lower serum concentration. Cell counts were about 30% less in the cells in grown in 2.5% serum and more cells were detached from the plates when they were harvested for luciferase activity. Also, there was an inconsistency of luciferase data between the triplicated samples, which hinders the validity and accuracy of the data produced in this experiment. Therefore, this experiment can be performed again to ensure similar cell counts between different experimental conditions to produce more reliable luciferase activity data.

## **Discussion:**

The iPSC technology holds a great therapeutic advantage and provides a vast potential for developing personalized treatments for diseases and for drug screening. However, the use of iPSC is hampered by the unknown mechanism of prolonged duration of reprogramming and a low success rate of less than 0.1%. The intricacies of the reprogramming mechanism are not fully understood, mostly due to the random nature of the expression of the reprogramming genes.

Also, most of the studies have used single inducible system to express all genes simultaneously, which has made it more difficult to observe the effect of each reprogramming genes during the reprogramming process. Therefore, the EcR based molecular switch, Rheostat system was utilized in conjunction with doxycycline-dependent molecular switch to explicate the effects of each reprogramming gene during the reprogramming process. Components of Rheostat system were cloned into two lentiviral vectors through subsequent steps of restriction enzyme digestions, ligations, and cloning, and were delivered into 293T cells. Rheostat system's inducibility was tested with varying concentration of its ligand, RSG-1. Induced cells were observed under the fluorescent microscope to detect eGFP activity and were harvested to compare luciferase activity. The result showed that the Rheostat system can express the gene expression in ligand dose-dependent manner. However, gene expression was also detected in non-induced sample, which conferred that Rheostat system can be activated by other mechanisms. There are two possibilities in which Rheostat system can be activated without its ligand. One is that there are other steroidal ligands that have some affinity to Rheo R that bind and activate the system. The other is that RRE may have similar binding sites with other activated nuclear receptors, which can activate the system without its ligand RSG-1 and its receptors. Following experiment revealed that the abnormal activation of the system was depended on RRE as RRE itself generated gene expression in absence of its receptor, Rheo R and Rheo A. Rheo R and Rheo A were found to have no role in expressing random gene expression since there was no activation of the system when RRE was absence. Rheo R was binding RSG-1 and heterodimerized with Rheo A correctly and moved across the nuclear membrane to bound RRE. However, the RRE was absence and the gene expression was not

initiated. Therefore, it was deduced that RRE was sensitive to other activated nuclear receptors and has initiated the gene response.

The third experiment was performed to confirm RRE's sensitivity to other nuclear receptor that can be activated by exogenous chemicals such as growth hormones in FBS serum. The result showed that RRE has some responsiveness to serum concentrations, as its expression was lower in the cells grown in lower serum concentration and higher in cells grown in higher serum concentration. However, this result was not reliable since the cells grown in 5% serum showed higher expression than the cells grown in 7.5%. This may be due to the serum concentration directly affecting cell growth. When the cells were harvested for luciferase assay, cell numbers were tended to be lower in the cells that were cultured in lower concentration of serum. This indicated that the low serum concentration has affected the growth of the cells, which may have further interrupted the transfection and integration efficiency of the construct into the DNA. Therefore, luciferase activity may not represent true RRE sensitivity to the serum concentration. In order to illustrate whether decreased in cellular growth or decreased in serum concentration led to lowering of the luciferase activity, experiment can be performed to demonstrate transfection efficiency of the vector, FU- $\Delta$ RRE-eGFP-T2A-LUC of all cells while maintaining good cellular environment for growth. The transfection efficiency then can be used to normalize luciferase data so that the different cell numbers among different experimental conditions would not skew the luciferase activity. This experiment would exemplify that luciferase activity can be equally compared without the concern of the cell growth among other samples and would show true responsiveness of RRE to serum concentration. The data obtained from this experiment can be crucial in determining RRE sensitivity to steroid and may assist in finding the cause of the abnormal RRE expression without an induction.

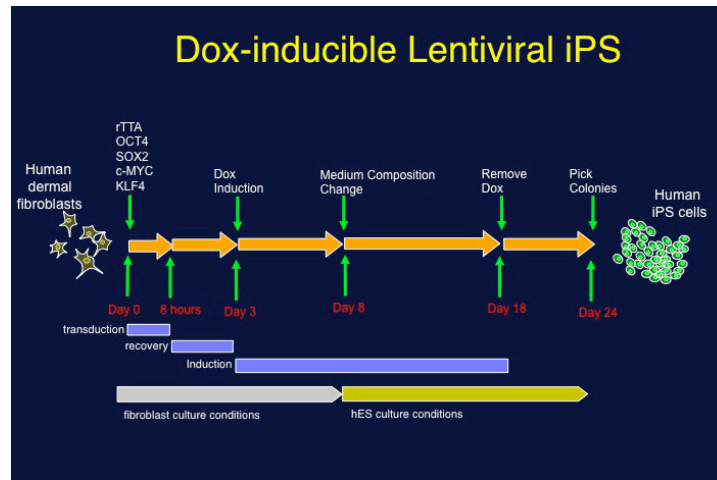
It is unclear whether the manufacturer of Rheostat system utilizes different RRE sequences that were obtained from the Genebank or have incorporated co-regulatory genes to regulate the sensitivity of RRE to other nuclear receptor. Therefore, if there is a co-regulatory genes that assist in decreasing the sensitivity of RRE, new oligonucleotides can be synthesized to have such co-regulatory genes incorporated into the upstream of the binding sites of RRE. This new RRE then can be incorporated into FU-ΔGW-eGFP-T2A-LUC after excising hUBC by PacI and BamHI restriction enzyme digestion, and be transfected to 293T cells to confirm tight control of the gene expression. When the sensitivity of RRE to other nuclear receptor is completely eliminated, the Rheostat system can be incorporated into our fibroblast cell lines along with the doxycycline system to study temporal requirements of certain gene expression during the reprogramming process.

#### **Future Research Plan: Controlled Expression of Reprogramming Factors for Investigation of iPSC Mechanism.**

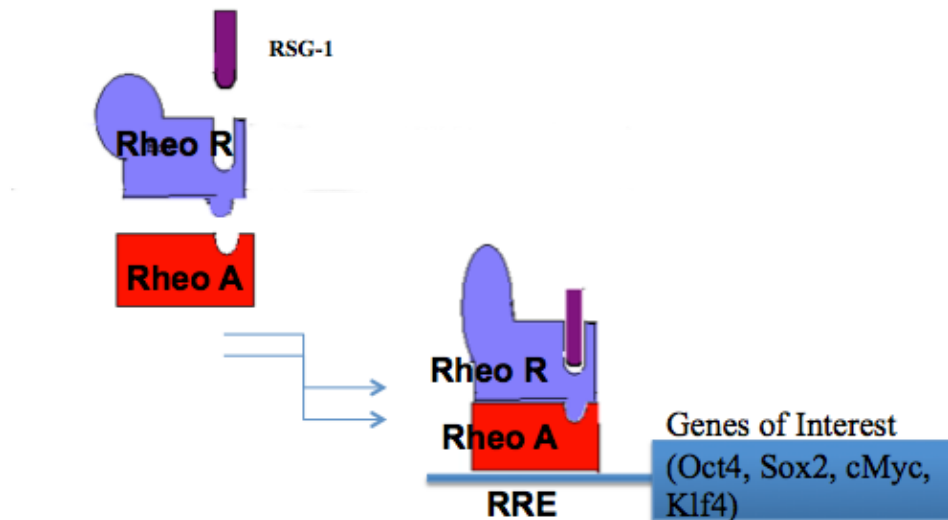
The combination of Rheostat and doxycycline inducible system offers an advantage in regulation of expression of transcription factors by two independent mechanisms. This provides many possible ways to investigate how each gene expression can affect reprogramming process over time. First, the affect of two factors that are found to be most essential during reprogramming process, Oct4 and Sox2 can be investigated.<sup>6</sup> Oct4 plasmid will be cloned into FU-ΔRRE-T2A-Luc and be regulated by RSG-1. Sox2 plasmid will be cloned into our docycycline-dependent vector and will be regulated by doxycycline. The other two factors, cMyc and Klf4, will be cloned into the vector that is designed to produce expression constitutively. The experiments will be designed in a specific way to observe the reprogramming efficiency by expressing each gene in varying time period. For example, an equal number of fibroblast cells

will be cultured in each well of 6-well plate and have all genes expressed. The first well, designated as Day 1, will have the expression of Sox2 turned off by stopping the addition of doxycycline into the well. Five days later, the expression of Sox2 in the second well will be halted. Following these procedures, the expression of Sox2 in each well will be halted in 5 day intervals. The cells will be observed under the microscope during this 25 days time-controlled experiment to detect the appearance of *bona fide*, the iPSC colonies. The number of iPSC colonies will be counted in each well and the results will be compared with our lab's previous reprogramming data to estimate the reprogramming efficiency. The same experiment can be utilized by halting the expression of Oct4 in 5 day intervals. This experiment data can be compared with the one from Sox2 dependent experiment and our lab's previous reprogramming data to assess the efficiency. If halting the expression of either gene does not affect the efficiency, the experiment can be employed in an opposite way; instead of halting the expression, expression can be induced in 5 day intervals and the results will be compared with previous data. In a parallel iteration, Oct4 and Sox2 will be expressed constitutively and Klf4 and cMyc will be placed under inducible control of expression and the experiment repeated as described above. Therefore, by having two molecular systems regulating the expression of the four genes, they permit many possible ways to study temporal requirement of reprogramming process by regulating each gene expression in varying time period. These experiments will allow observation as to which gene expression is most necessary during the course of the reprogramming process. Thus, it will ultimately facilitate the finding of optimal reprogramming condition, which can increase the efficiency and success rates of reprogramming fibroblast into iPSCs.

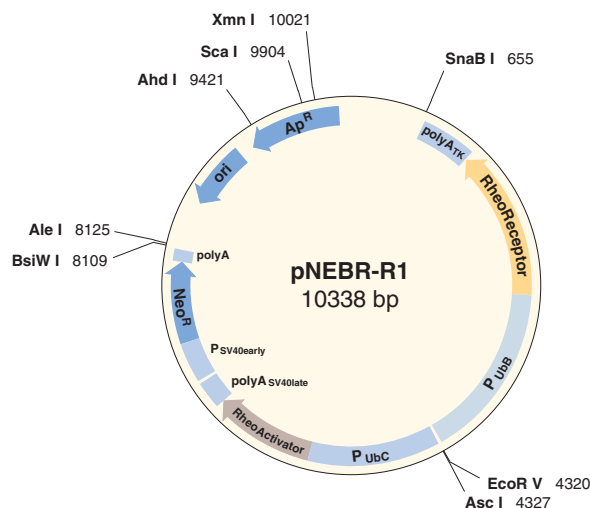
## Figures and Tables:



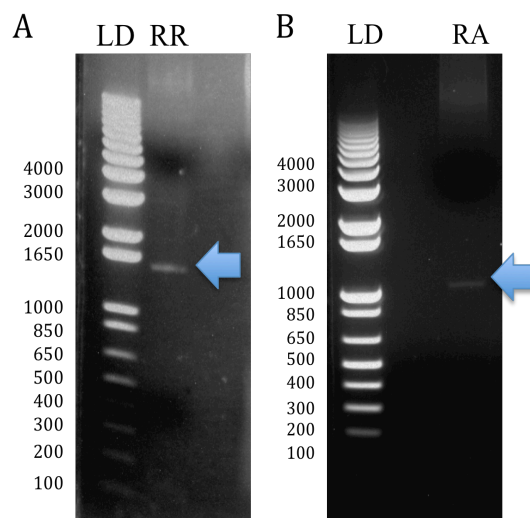
**Figure 1.** Diagram of doxycycline dependent induction of transcription factors used in generating iPSCs. Fibroblasts are transduced with four reprogramming genes and the expression is induced by the addition of doxycycline at day 3. Induction ceases at day 18 by removing doxycycline.



**Figure 2.** Diagram of inducible mechanism of Rheostat system. When RSG-1 binds Rheo R, Rheo R heterodimerizes with Rheo A and moves across the nuclear membrane. The Rheo R and Rheo A complex then binds to RRE to generate gene expression.

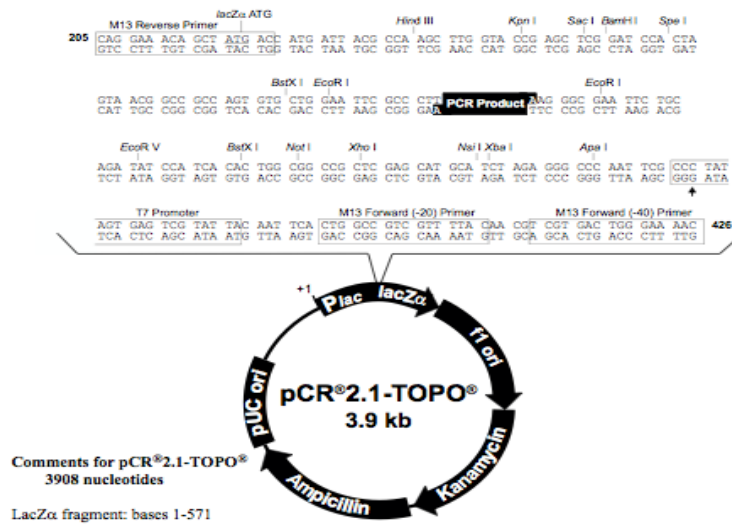


**Figure 3.** Schematic drawing of the DNA plasmid in pNERB-R1 backbone, containing gene segments of Rheo R and Rheo A. The position of Rheo R and Rheo A is indicated in yellow and grey respectively. Gene segments of Rheo R and Rheo A are about 1,500 bp and 1,000bp long respectively.

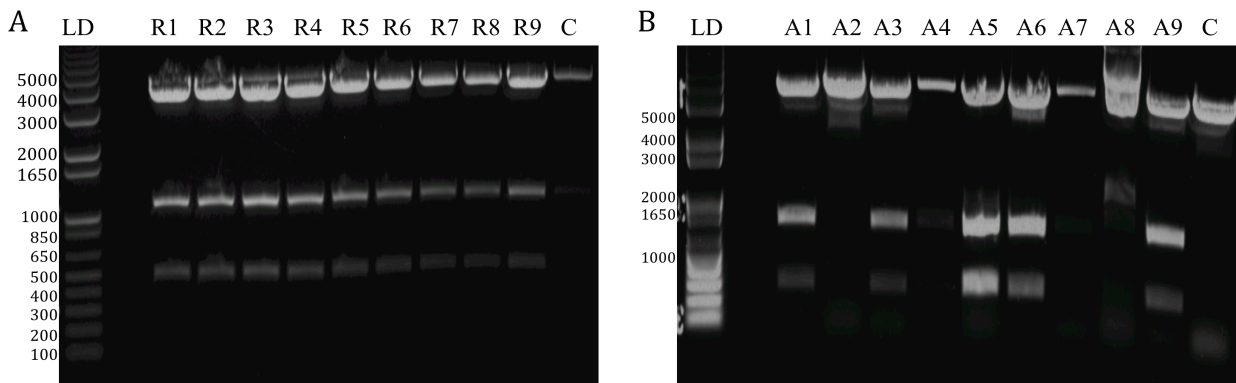


**Figure 4. (A).** PCR of 1 µg of pNERB-R1 vector using Rheo R specific forward and reverse primers. Sequences of Rheo R specific primers are shown in Table 1 with their respective melting points. Arrow indicates successful amplification of Rheo R. PCR Condition: 30 cycles of 60, 60 and 90 seconds at 95°C, 58°C and 68°C respectively. 10 mins of extension period at the end of the reaction at 72°C. **(B)** PCR of 1 µg of pNERB-R1 vector using Rheo A specific forward and reverse primers. Sequences of Rheo A specific primers are shown in Table 1 with their respective melting points. Arrow indicates successful amplification of Rheo A. PCR condition: 30 cycles of 60, 60 and 90 seconds at 95°C, 58°C and 68°C respectively. 10 mins of extension period at the end of the reaction at 72°C.





**Figure 5.** Sequences and positions of the restriction enzyme recognition sites in *LacZα* fragment of pCR2.1-TOPO vector. Insertion site for PCR product is highlighted in black.



**Figure 6. (A)** EcoRI digestion of PCR2.1-TOPO vector containing Rheo R. Digested vectors from 10 colonies were separated by electrophoresis in 0.8% agarose gel. Proper insertion of Rheo R is confirmed in all nine samples (R1-R9) by observing 1000 bp fragments excised from the vector. C indicates the genomic DNA attained from light blue colony for negative control. As expected, the digestion of the vector was not observed. **(B)** EcoRI digestion of PCR2.1-TOPO vector containing Rheo A. Digested vectors from 10 colonies were separated by electrophoresis in 0.8% agarose gel. Confirmation of proper insertion of Rheo A is confirmed in sample A1, A3, A5, A6 and A9 by observing 1500bp fragments excised from the vector. C indicates negative control and the digestion was not observed.

RheoReceptor\_CDS1209-2675.ape from 1 to 1510

Alignment to

RheoR9.ape from 53 to 1580

Matches(|):1507

Mismatches(#):3

Gaps( ):18

Unattempted(.):0

```

      *      *      *      *      *      *      *      *      *      *
1  ggctagcgcgatgccaccATGAAGCTACTGTCTTCTATCGAACAAGCATGCGATATTTGCCGACTTAAAAAGCTCAAGTGCTCCAAAGAAAAACCGAAGTGC 100
      |
53 GGCTAGCGATGCCACCATGAAGCTACTGTCTTCTATCGAACAAGCATGCGATATTTGCCGACTTAAAAAGCTCAAGTGCTCCAAAGAAAAACCGAAGTGC 152
      *      *      *      *      *      *      *      *      *      *
101 GCCAAGTGTCTGAAGAACAACTGGGAGTGTGCTACTCTCCCAAACCAAAGGTCTCCGCTGACTAGGGCACATCTGACAGAAGTGGAATCAAGGCTAG 200
      |
153 GCCAAGTGTCTGAAGAACAACTGGGAGTGTGCTACTCTCCCAAACCAAAGGTCTCCGCTGACTAGGGCACATCTGACAGAAGTGGAATCAAGGCTAG 252
      *      *      *      *      *      *      *      *      *      *
201 AAAGACTGGAACAGCTATTTCTACTGATTTTCTCCTCGAGAAGACCTTGACATGATTTTGAAATGGATTCTTTACAGGATATAAAAGCATTGTTAACAGG 300
      |
253 AAAGACTGGAACAGCTATTTCTACTGATTTTCTCCTCGAGAAGACCTTGACATGATTTTGAAATGGATTCTTTACAGGATATAAAAGCATTGTTAACAGG 352
      *      *      *      *      *      *      *      *      *      *
301 ATTATTTGTACAAGATAAATGTGAATAAAGATGCCGTCACAGATAGATTGGCTTCAGTGGAGACTGATATGCCTCTAACATTGAGACAGCATAGAATAAGT 400
      |
353 ATTATTTGTACAAGATAAATGTGAATAAAGATGCCGTCACAGATAGATTGGCTTCAGTGGAGACTGATATGCCTCTAACATTGAGACAGCATAGAATAAGT 452
      *      *      *      *      *      *      *      *      *      *
401 GCGACATCATCATCGGAAGAGAGTAGTAACAAAGGTCAAAGACAGTTGACTGTATCGCCGGAATTCCTCGGGGATCCGGCTGAGTGCCTAGTACCCGAGA 500
      |
453 GCGACATCATCATCGGAAGAGAGTAGTAACAAAGGTCAAAGACAGTTGACTGTATCGCCGGAATTCCTCGGGGATCCGGCTGAGTGCCTAGTACCCGAGA 552
      *      *      *      *      *      *      *      *      *      *
501 CTCAGTGCGCCATGAAGCGGAAAGAGAAGAAAGACACAGAAGGAGAAGGACAACTGCCTGTGACGACGACGCGTGGACGACCACATGCCGCCATTAT 600
      |
553 CTCAGTGCGCCATGAAGCGGAAAGAGAAGAAAGACACAGAAGGAGAAGGACAACTGCCTGTGACGACGACGCGTGGACGACCACATGCCGCCATTAT 652
      *      *      *      *      *      *      *      *      *      *
601 GCAGTGTGAACCTCCACCTCCTGAAGCAGCAAGGATTCACGAAGTGGTCCCAAGGTTTCTCTCCGACAAGCTGTTGGAGACAAACCGGCAGAAAAACATC 700
      |
653 GCAGTGTGAACCTCCACCTCCTGAAGCAGCAAGGATTCACGAAGTGGTCCCAAGGTTTCTCTCCGACAAGCTGTTGG#GACAAACCGGCAGAAAAACATC 752
      *      *      *      *      *      *      *      *      *      *
701 CCCAGTTGACAGCCAACAGCAGTTCCCTTATCGCCAGGCTCATCTGGTACCAGGACGGGTACGAGCAGCCTTCTGATGAAGATTGAAGAGGATTACGC 800
      |
753 CCCAGTTGACAGCCAACAGCAGTTCCCTTATCGCCAGGCTCATCTGGTACCAGGACGGGTACGAGCAGCCTTCTGATGAAGATTGAAGAGGATTACGC 852
      *      *      *      *      *      *      *      *      *      *
801 AGACGTGGCAGCAAGCGGACGATGAAAACGAAGAGTCTGACACTCCCTTCCGCCAGATCACAGAGATGACTATCCTCAGGGTCCAATTATCGTGGAGTT 900
      |
853 AGACGTGGCAGCAAGCGGACGATGAAAACGAAGAGTCTGACACTCCCTTCCGCCAGATCACAGAGATGACTATCCTCAGGGTCCAATTATCGTGGAGTT 952
      *      *      *      *      *      *      *      *      *      *
901 CGCGAAGGGATTGCCAGGGTTCGCCAAGATCTCGCAGCCTGATCAAATTACGCTGCTTAAGGCTTGCTCAAGTGAGGTAATGATGCTCCGAGTCGCGCGA 1000
      |
953 CGCGAAGGGATTGCCAGGGTTCGCCAAGATCTCGCAGCCTGATCAAATTACGCTGCTTAAGGCTTGCTCAAGTGAGGTAATGATGCTCCGAGTCGCGCGA 1052
      *      *      *      *      *      *      *      *      *      *
1001 CGATACGATGCGGCCCTCAGACAGTATTCTGTTTCGCGAACAACCAAGCGTACACTCGCGACAACACCGCAAGGCTGGCATGGCCGAGGTCATCGAGGATC 1100
      |
1053 CGATACGATGCGGCCCTCAGACAGTATTCTGTTTCGCGAACAACCAAGCGTACACTCGCGACAACACCGCAAGGCTGGCATGGCCGAGGTCATCGAGGATC 1152
      *      *      *      *      *      *      *      *      *      *
1101 TACTGCACCTTCTGCCGGTGCATGTACTCTATGGCGTTGGACAACATCCATTACGCGCTGCTCAGGCTGTCGTCATCTTTTCTGACCGGCCAGGGTTGGA 1200
      |
1153 TACTGCACCTTCTGCCGGTGCATGTACTCTATGGCGTTGGACAACATCCATTACGCGCTGCTCAGGCTGTCGTCATCTTTTCTGACCGGCCAGGGTTGGA 1252
      *      *      *      *      *      *      *      *      *      *
1201 GCAGCGCAACTGGTGGAGAAATCCAGCGGTACTACCTGAATACGCTCCGCATCTATATCCTGAACAGCTGAGCGGGTCGCGCGGTCGTCCTCGTCATA 1300
      |
1253 GCAGCGCAACTGGTGGAGAAATCCAGCGGTACTACCTGAATACGCTCCGCATCTATATCCTGAACAGCTGAGCGGGTCGCGCGGTCGTCCTCGTCATA 1352
      *      *      *      *      *      *      *      *      *      *
1301 TACGGCAAGATCCTCTCAATCCTCTCTGAGCTACGCACGCTCGGCATGCAAACTCCAACATGTGCATCTCCCTCAAGCTCAAGAACAGAAAGCTGCCGC 1400
      |
1353 TACGGCAAGATCCTCTCAATCCTCTCTGAGCTACGCACGCTCGGCATGCAAACTCCAACATGTGCATCTCCCTCAAGCTCAAGAACAGAAAGCTGCCGC 1452
      *      *      *      *      *      *      *      *      *      *
1401 CTTTCCTCGAGGAGATCTGGGATGTGGCGGACATGTCGCACACCCAACGCCCGCCTATCCTCGAGTCCCCCAGCAATCTCTA----- 1482
      |
1453 CTTTCCTCGAGGAGATCTGGGATGTGGCGGACATGTCGCACACCCAACGCCCGCCTATCCTCGAGTCCCCCAGCAATCTCTA#GGCGGCCGCTTTTCGATC 1552
      *      *      *      *      *      *      *      *      *      *

```

```

      *      *      *
1483 CCCC GGCTGCAGGGAGATGGGGAGGC 1510
      |||||
1553 CCCC GGCTGCAGGGAGATGGGGAGGC 1580
      *      *      *

```

**Figure 7.** The sequence alignment of Rheo R that was cloned into lentiviral construct, FU-ΔGWMCS-T2A with the published sequence data obtained from Genebank. The mismatches and a 17bp insertion are highlighted in red.

```

RheoActivator_CD85565-6539.ape from 19 to 983
Alignment to
RheoA2.ape from 38 to 1002
Matches(1):964
Mismatches(#):1
Gaps( ):0
Unattempted(.):0

      *      *      *      *      *      *      *      *      *      *
19  ATGGGCCCTAAAAAGAAGCGTAAAGTCGCCCCCGACCGATGTCAGCCTGGGGACGAGCTCCACTTAGACGGCGAGGACGTGGCGATGGCGCATGCCG 118
      |||||
38  ATGGGCCCTAAAAAGAAGCGTAAAGTCGCCCCCGACCGATGTCAGCCTGGGGACGAGCTCCACTTAGACGGCGAGGACGTGGCGATGGCGCATGCCG 137
      *      *      *      *      *      *      *      *      *      *

      *      *      *      *      *      *      *      *      *      *
119 ACGGCGTAGACGATTTTCGATCTGGACATGTTGGGGGACGGGGATTCCCGGGGCGGGATTACCCCCACGACTCCGCCCCCTACGGCGCTCTGGATAT 218
      |||||
138 ACGGCGTAGACGATTTTCGATCTGGACATGTTGGGGGACGGGGATTCCCGGGGCGGGATTACCCCCACGACTCCGCCCCCTACGGCGCTCTGGATAT 237
      *      *      *      *      *      *      *      *      *      *
      *      *      *      *      *      *      *      *      *      *
219 GGCCGACTTCGAGTTTTCGACAGATGTTTACCGATGCCCTTGAATTGACGAGTACGGTGGGGAAATTCGAGATGCCTGTGGACAGGATCCTGGAGGCAGAG 318
      |||||
238 GGCCGACTTCGAGTTTTCGACAGATGTTTACCGATGCCCTTGAATTGACGAGTACGGTGGGGAAATTCGAGATGCCTGTGGACAGGATCCTGGAGGCAGAG 337
      *      *      *      *      *      *      *      *      *      *

      *      *      *      *      *      *      *      *      *      *
319 CTTGCTGTGGAACAGAAGAGTGAACAGGGCGTTGAGGGTCTGGGGGAACCGGGGTAGCGGCAGCAGCCAAATGACCTGTGACTAACATCTGTCAGG 418
      |||||
338 CTTGCTGTGGAACAGAAGAGTGAACAGGGCGTTGAGGGTCTGGGGGAACCGGGGTAGCGGCAGCAGCCAAATGACCTGTGACTAACATCTGTCAGG 437
      *      *      *      *      *      *      *      *      *      *

      *      *      *      *      *      *      *      *      *      *
419 CAGCTGACAAACAGCTATTACGCTTGTGAGTGGGCGAAGAGGATCCACACTTTTCTCCTTGCTCTGGATGATCAGGTCATATTGCTGCGGGCAGG 518
      |||||
438 CAGCTGACAAACAGCTATTACGCTTGTGAGTGGGCGAAGAGGATCCACACTTTTCTCCTTGCTCTGGATGATCAGGTCATATTGCTGCGGGCAGG 537
      *      *      *      *      *      *      *      *      *      *

      *      *      *      *      *      *      *      *      *      *
519 CTGGAATGAACCTCCTCATTGCCTCCTTTTACACCGATCCATTGATGTTTCGAGATGGCATCCTCCTTGCCACAGGCTTTCACGTGCACCGCAACTCAGCC 618
      |||||
538 CTGGAATGAACCTCCTCATTGCCTCCTTTTACACCGATCCATTGATGTTTCGAGATGGCATCCTCCTTGCCACAGGCTTTCACGTGCACCGCAACTCAGCC 637
      *      *      *      *      *      *      *      *      *      *

      *      *      *      *      *      *      *      *      *      *
619 CATTGACGAGGATAGGAGCCATCTTTGATCGGGTGCTGACAGAGCTAGTGTCAAAAATGCGTGACATGAGGATGGACAAGACAGAGCTTGGCTGCCTGA 718
      |||||
638 CATTGACGAGGATAGGAGCCATCTTTGATCGGGTGCTGACAGAGCTAGTGTCAAAAATGCGTGACATGAGGATGGACAAGACAGAGCTTGGCTGCCTGA 737
      *      *      *      *      *      *      *      *      *      *

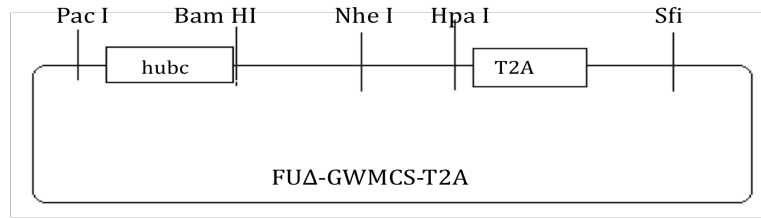
      *      *      *      *      *      *      *      *      *      *
719 GGGCAATCATTCTGTTTAAATCCAGAGGTGAGGGGTTTGAATCCGCCAGGAAGTTGAACCTTCTACGTGAAAAAGTATATGCCGCTTTGGAAGAATATAC 818
      |||||
738 GGGCAATCATTCTGTTTAAATCCAGAGGTGAGGGGTTTGAATCCGCCAGGAAGTTGAACCTTCTACGTGAAAAAGTATATGCCGCTTTGGAAGAATATAC 837
      *      *      *      *      *      *      *      *      *      *

      *      *      *      *      *      *      *      *      *      *
819 TAGAACAACACATCCCGATGAACAGGAAGATTGCAAAACTTTTGCTTCGCTGCTTCTTTACGTTCCATAGGCCTTAAGTGTGAGCATTGTTT 918
      |||||
838 TAGAACAACACATCCCGATGAACAGGAAGATTGCAAAACTTTTGCTTCGCTGCTTCTTTACGTTCCATAGGCCTTAAGTGTGAGCATTGTTT 937
      *      *      *      *      *      *      *      *      *      *

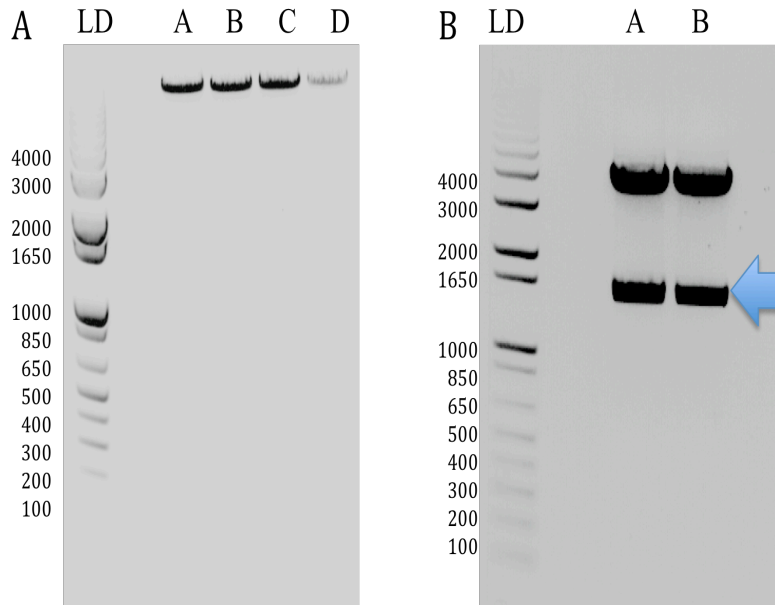
      *      *      *      *      *      *      *      *      *      *
919 TTCTTTCGCCTTATTGGAGATGTTCCAATTGATACGTTCTGATGGAGATGCTTGAATCACCTTC 983
      |||||
938 TTCTTTCGCCTTATTGGAGATGTTCCAATTGATACGTTCTGATGGAGATGCTTGAATCACCTTC 1002
      *      *      *      *      *      *      *      *      *      *

```

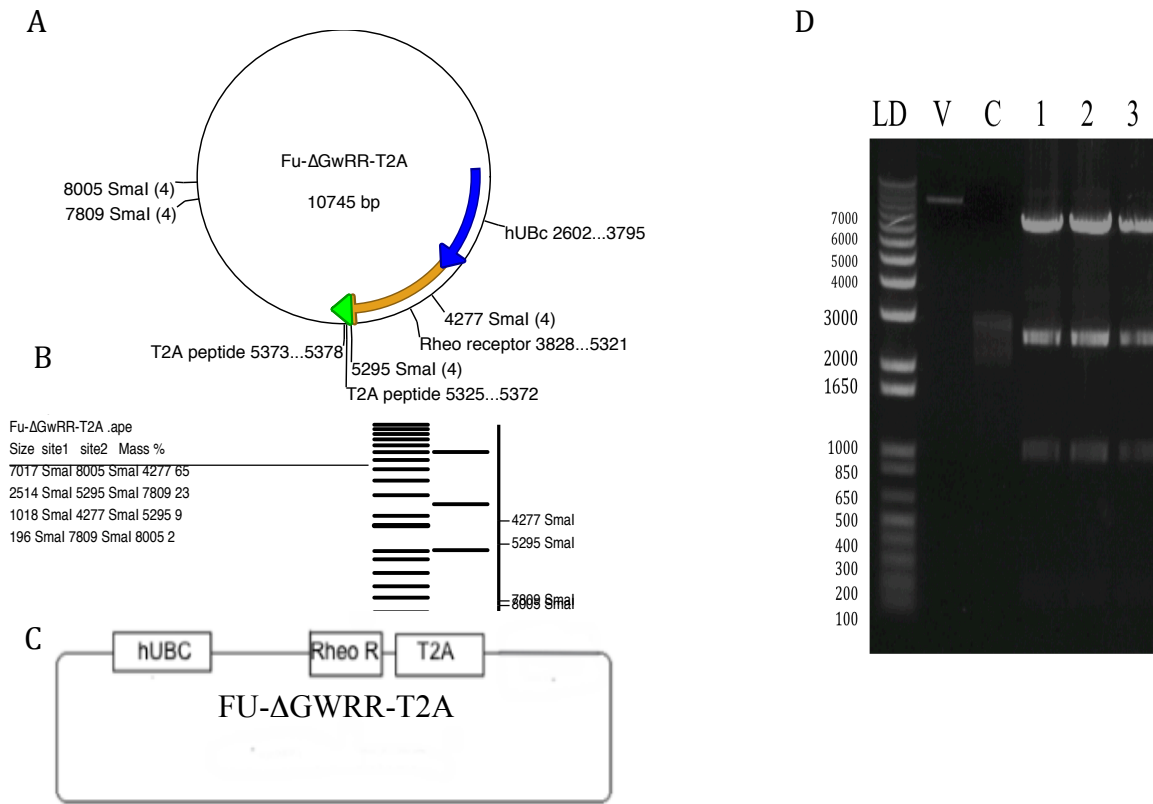
**Figure 8.** The sequence alignment of Rheo A with the published sequence data obtained from Genebank. The Rheo A that was cloned into the lentiviral construct, FU-ΔGWRR-T2A, contained one mutation at 171bp region, which caused the single point mutation substituting threonine for glycine.



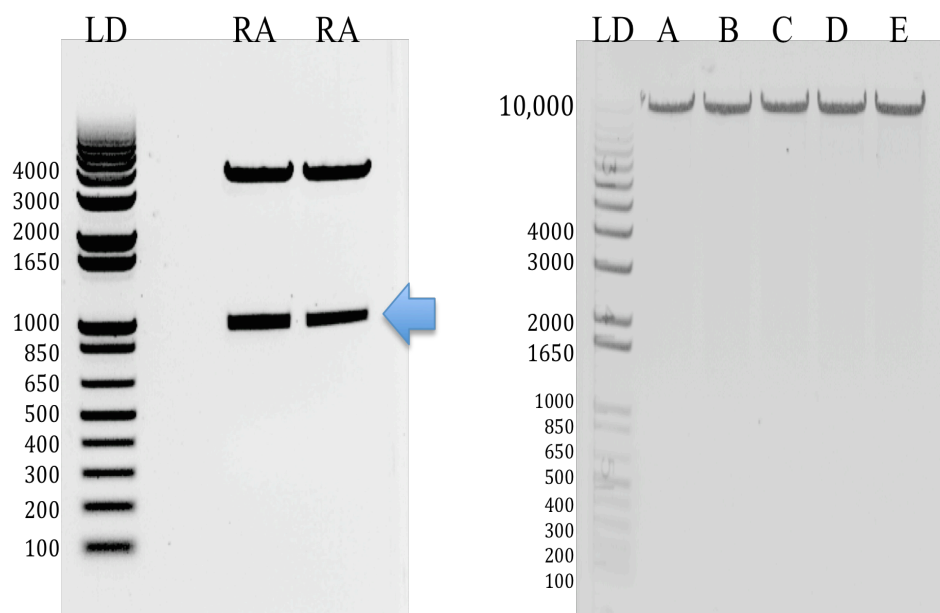
**Figure 9.** Schematic drawing of the lentivirus transfer vector, FU-ΔGWMCS-T2A, indicating the position of the hUBC, T2A peptides, and respective restriction enzyme recognition sites that were utilized to clone Rheo R and Rheo A.



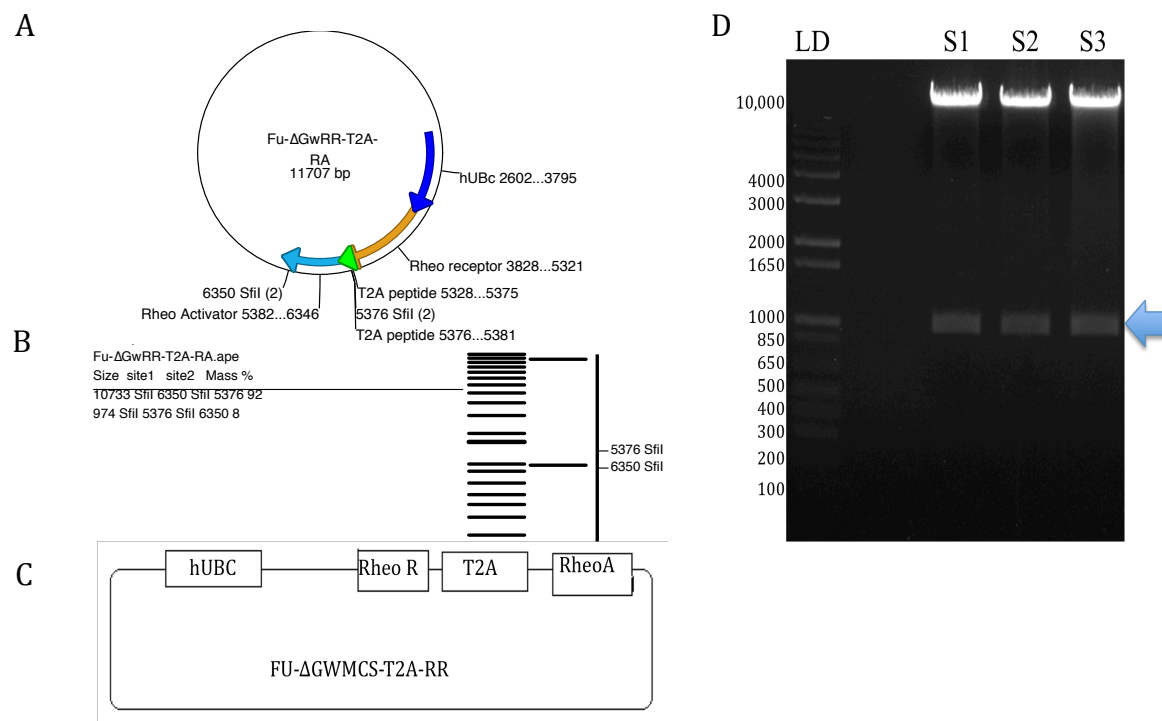
**Figure 10. (A)** NheI and Hpa I digestion of FU-ΔGWMCS-T2A vector in 0.8% agarose gel. Digested vector was divided and loaded into four consecutive wells. Excised fragments were 18bp and were too small to be shown on the gel. All bands were excised and purified by gel extraction. **(B)** NheI and StuI digestion of TOPO vector containing Rheo R in 0.8% agarose gel. The arrow shows successful excision of Rheo R from the vector. Two bands were excised and were purified using gel extraction.



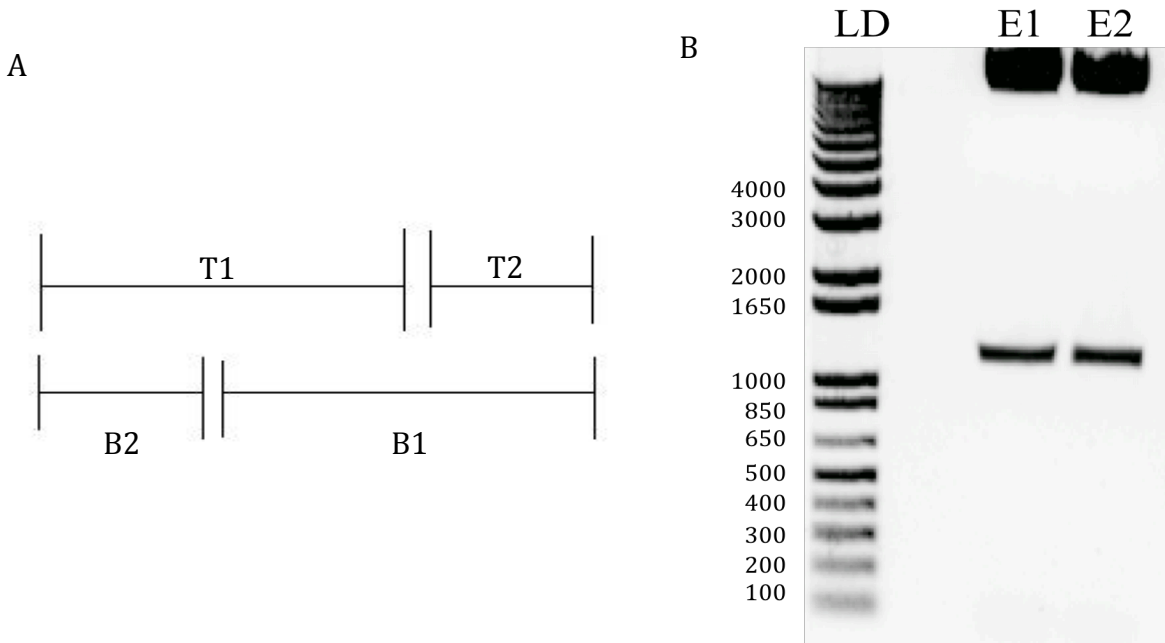
**Figure 11. (A)** Schematic drawing of FU-ΔGWRR-T2A indicating the position of Rheo R and Smal restriction enzyme recognition sites. **(B)** The map of Smal digestion of FU-ΔGWRR-T2A. Excised fragments are 7017, 2514, 1018, and 195 bp. **(C)** Schematic drawing of the vector, FU-ΔGWRR-T2A, after the insertion of Rheo R. **(D)** Smal digestion of FU-ΔGWRR-T2A after its genomic DNA extraction from HB101. Digested vector was separated by electrophoresis in 0.8% agarose gel. Digested vector was divided and loaded into three consecutive wells (1~3). V is the undigested vector for the positive control and C is the unligated HB101 vector. Proper insertion of Rheo R into FU-ΔGWMCS-T2A vector is confirmed by excision of expected fragments of 7000, 2500bp, and 1000bp from the vector. 200bp fragments can also be observed at longer UV exposure times.



**Figure 12. (A)** Confirmation of Sfi I digest of Rheo A in pCR2.1-TOPO vector. Digested vector was load into two wells and was separated by electrophoresis in 0.8% agarose gel. The 1000bp fragment confirmed the excision of Rheo A from the vector. 1000bp fragments were excised from the gel and purified by gel extraction. **(B)** Confirmation of Sfi I digest of FU-ΔGWRR-T2A. Digested vector was loaded into five consecutive wells and were separated by electrophoresis in 0.8% agarose gel. FU-ΔGWRR-T2A only has two Sfi I site in close proximity and digestion of Sfi I enzyme linearized the construct. All bands were excised and purified by the gel extraction.

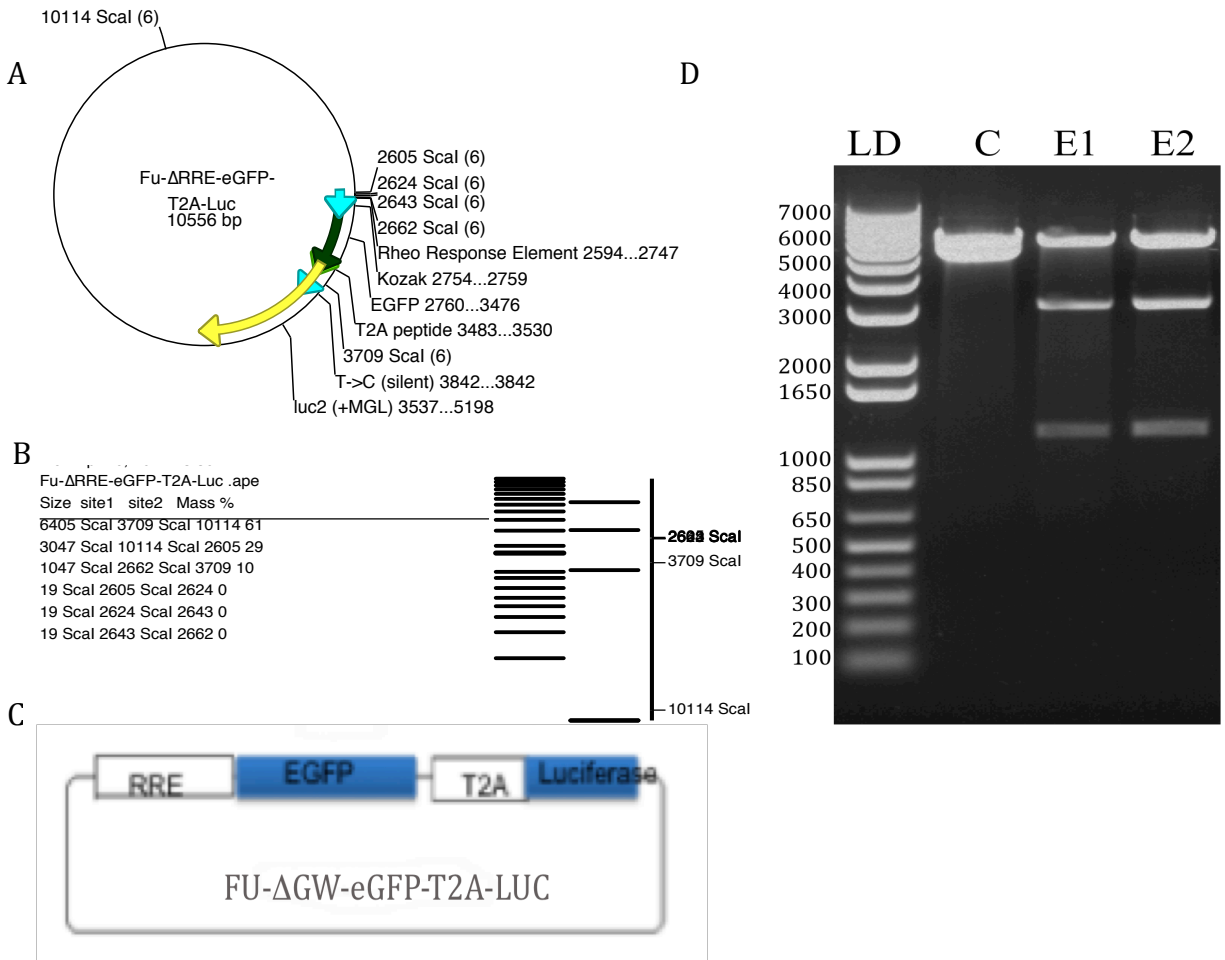


**Figure 13. (A)** Schematic drawing of FU-ΔGWRR-T2A-RA showing the position of Rheo R, Rheo A and Sfi I restriction enzyme recognition sites. **(B)** Map of Sfi I digestion of FU-ΔGWRR-T2A-RA indicating excised fragment of 974bp. **(C)** Sfi I digest of FU-ΔGWRR-T2A-RA after its genomic DNA extraction from HB101 by mini prep. Digested vector was loaded into three consecutive wells designated as S1, S2 and S3 and were separated by electrophoresis in 0.8% agarose gel. Proper insertion of Rheo A into FU-ΔGWRR-T2A vector was confirmed by excision of ~1000bp diagnostic bands indicated with arrow. **(D)** Schematic drawing of the vector, FU-ΔGWRR-T2A-RA, indicating the insertion of Rheo A into FU-ΔGWRR-T2A.



**Figure 14. (A)** Schematic drawing of RRE oligonucleotide hybridized to form RRE. Four different oligonucleotides, T1, T2, B1, and B2 were synthesized based on the gene sequences provided from Genbank. **(B)** *PacI* and *BamHI* digestion of FU-ΔGW-eGFP-T2A-LUC on 0.8% agarose gel. Digested vector was divided in half and was loaded onto two consecutive wells, designated as E1 and E2. 1300 base pair fragments were excised from the vector, which indicated the excision of hUBC. Vectors were excised from the gel and purified by gel extraction. Purified products were then ligated with RRE and were transformed into HB101 cells.



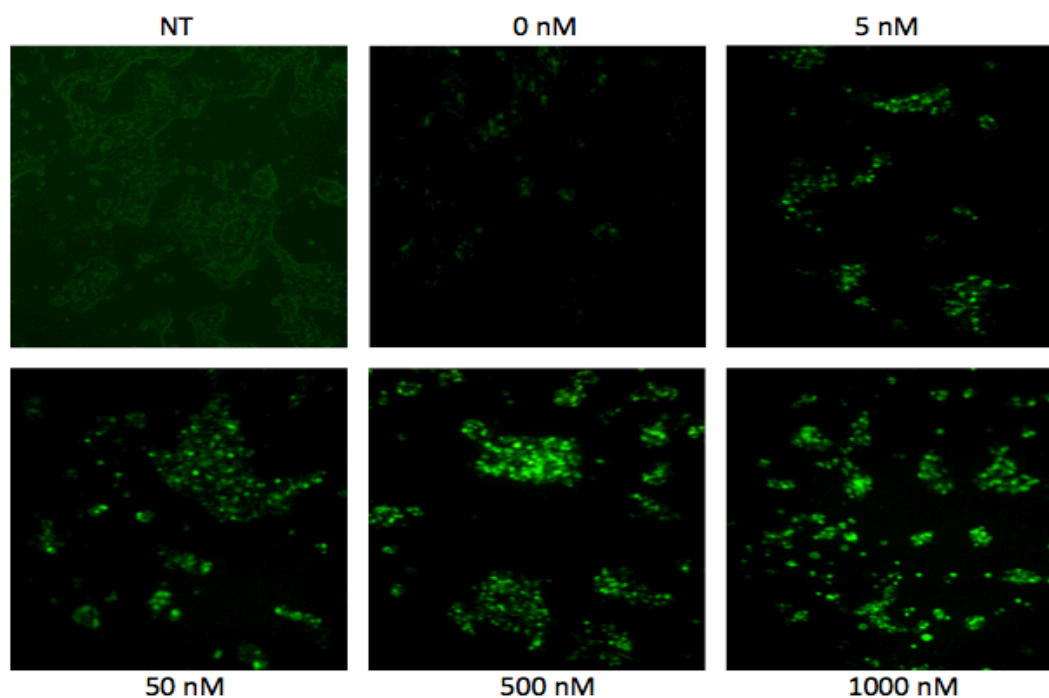


**Figure 15. (A)** Schematic drawing of FU-ΔRRE-eGFP-T2A-LUC vector and its ScaI restriction enzyme recognition sites. **(B)** Map of ScaI digestion of FU-ΔRRE-eGFP-T2A-LUC vector. Expected size of excised fragments are 6405, 3047, 1047 and 19 base pairs. **(C)** Schematic drawing of the vector, FU-ΔRRE-eGFP-T2A-LUC, indicating the insertion of RRE into FU-ΔGW-eGFP-T2A-LUC. **(D)** ScaI digestion of FU-ΔGW-eGFP-T2A-LUC after its genomic DNA extraction from HB101 by mini prep. Digested vector was divided and loaded into two wells designated as E1 and E2. Proper insertion of RRE into FU-ΔGW-eGFP-T2A-LUC was confirmed by observing excised fragments of about 1000 and 3000 bp.

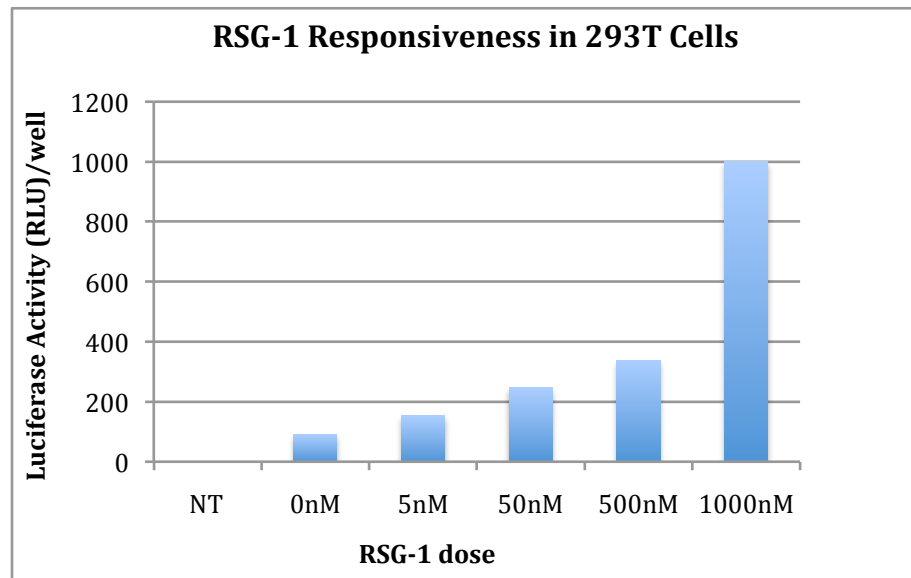
```
Matches(|):140
Mismatches(#):0
Gaps():0
Unattempted(.):0
```

2602 cggagtactgtcctccgagcggagtactgtcctccgagcggagtactgtcctccgagcggagtctctgtcctccgagcggag 2701  
 977 CGGAGTACTGTCTCCGAGCGGAGTACTGTCTCCGAGCGGAGTACTGTCTCCGAGCGGAGTACTGTCTCCGAGCGGAGTCTCTCTCCGAGCGGAG 1076  
 2702 actctagagggtatataatgggggccgtcagctactacca 2741  
 1077 ACTCTAGAGGGTATATAATGGGGGCCGTCAGCTACTACCA 1116

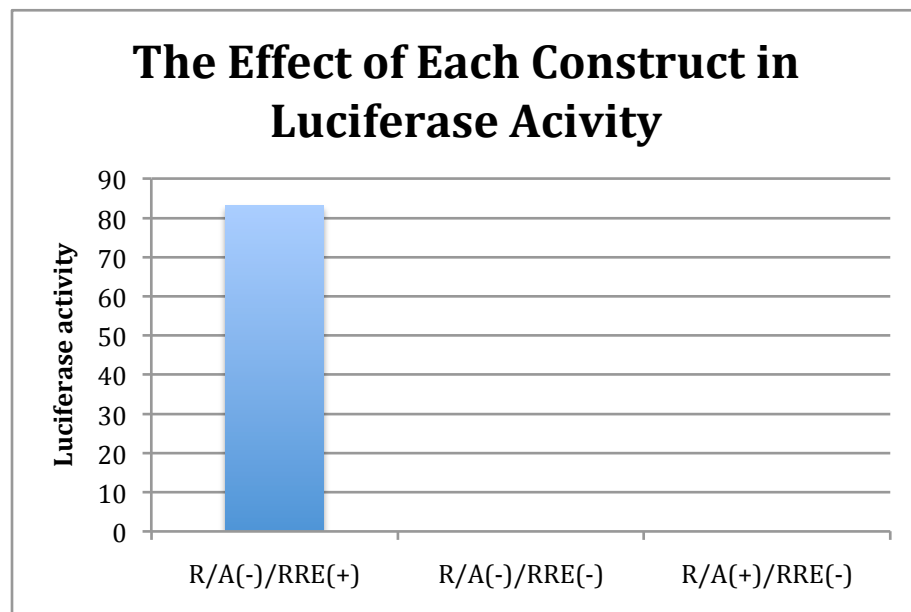
**Figure 16.** The sequences of RRE aligned with the published sequence data obtained from Genebank. No mismatches were found, confirming accurate nucleic acid hybridization and transformation into FU-ΔGW-eGFP-T2A-LUC construct.



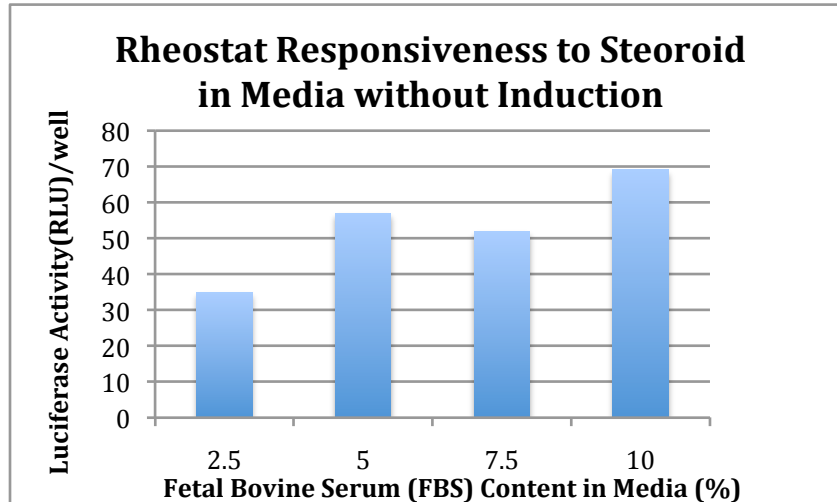
**Figure 17.** The eGFP activity observed by fluorescent microscope at 20X after 48 and 72 hours of the induction and transfection respectively. The eGFP activity was greater in cells induced with higher dose of the ligand, RSG-1. No eGFP activity was observed from the cells that was not transfected (NT). Deem eGFP activity was observed from the transfected cells that were not induced with the RSG-1 (0 nM).



**Figure 18.** RSG-1 Responsiveness in 293T cells. 293T cells were co-transfected with both FU- $\Delta$ RRE-eGFP-T2A-LUC and FU- $\Delta$ GwRR-T2A-RA constructs 12 hours prior to induction of four different ligand concentrations. Forty-eight hours after induction, cells were harvested for firefly luciferase assay. No luciferase activity was detected in negative control (NT) which both of the constructs were not transfected. The average value for each triplicate experiment is shown in the graph. The results show that luciferase activity increase in ligand dependent manner, but its activity was also detected in the non-induced sample (0 nM).



**Figure 19.** The Effect of each construct in luciferase activity. 293T cells were only transfected with either FU- $\Delta$ RRE-eGFP-T2A-LUC or FU- $\Delta$ GwRR-T2A-RA construct and induced with 500nM of RSG-1. Luciferase activity was only detected in sample that had RRE transfected. No luciferase activities were detected in other samples, which indicated that RRE can generate gene expression without its receptor mechanism.



**Figure 20.** RRE Responsiveness to FBS in media. Cells were cultured in four different FBS concentration and were transfected only with FU- $\Delta$ RRE-eGFP-T2A-LUC. RSG-1 was not added in order to expound RRE sensitivity to other endogenous nuclear receptor that may have been activated by growth hormone or other chemical compound present in FBS. Lower luciferase activity was detected in cells grown in lower concentration of FBS.

## Tables:

Rheo R-Forward	5'-ggctagcgatgccaccatgaagctactgtcttctatcgaa-3'	60°C
Rheo R-Reverse	5'-ctttaaggcctgcctcccccattccttg-3'	55°C
Rheo A-Forward	5'-cgggccctatgggcctaaaaagaagcg-3	47°C
Rheo A-Reverse	5'-gaggccgcattgggcctcagaaggtgattcaagcat-3'	57°C

**Table 1.** Primer sequences used to amplify Rheo R and Rheo A from pNERB-R1 vector and their melting temperature

[illegible]

**Table 2.** Sequences of RRE and its four oligonucleotide. Sequence of RRE was obtained from Genebank and its oligonucleotides were synthesized (Integrated DNA Technologies, www.idtdna.com). Four oligos were annealed by nucleic acid hybridization.

	#1	#2	#3	Average
NT	0	0	0	0
0 nM	96.275	89.785	88.677	91.579
5nM	149.375	155.065	155.29	153.093
50nM	243.336	245.931	250.872	246.713
500nM	328.758	350.347	337.529	338.878
1000nM	1042.816	968.462	1061.706	1004.328

**Table 3.** Triplicate of luciferase activity data obtained from the experiment “RSG-1 Responsiveness in 293T Cells”. Average of the three values were graphed in Figure 15.

## References

- 1 Murry, CE. Keller, G. Differentiation of embryonic stem cells to clinically relevant populations: Lessons from embryonic development. *Cell* 2008; 132:661-680
- 2 Plath K. William L. Progress in understanding reprogramming to the induced pluripotent state. *Nature* 2011; 12:253-265
- 3 Takahashi, K. Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006; 126:663-676
- 4 Tkahashi, K. Tanabe, K. Ohnuki, M. Narita, M. Ichisaka, T. Tomoda, K. Yamanaka, S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 2007; 131: 861-872
- 5 Tarnowsk, M. Sieron, A. Adult stem cells and their ability to differentiate. *Med SciMonit* 2006; 12:154-163
- 6 Inoue, H. Yamanaka, S. The use of induced pluripotent stem cells in drug development. *Clinical Pharmacology & Therapy* 2011, doi: 10.1038/clpt.2011.38
- 7 Okita, K. Yamanaka, S. Induction of pluripotency by defined factors, *Exp. Cell. Res.* 2010, doi: 10.1016/j.yexcr.2010.04.023
- 8 Nakagawa, M., Takizawa, N., Narita, M., Ichisaka, T. & Yamanaka, S. Promotion of direct reprogramming by transformation-deficient Myc. *Proc. Natl. Acad. Sci. USA* 107, 14152–14157 (2010).
- 9 Lee, G, Eirini P. Papapetrou, H. Modeling pathogenesis and treatment of familial dystonomia using patient-specific iPSCs
- 10 Huangfu, D. et al. Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol* 2008; 26, 795-797.
- 11 Wevers, L., Drevet, J. R., Lunke, M.D. & Iatrou, K. Insect Biochem. *Mol. Biol.* 2008; 25, 857-866.
- 12 Kumar, M.B, et al. A single point mutation in ecdysone receptor leads to increased ligand specificity: Implication for gene switch application. *PNAS* 2002; 99, 14710-14715
- 13 Carlson, G. R., Dhadialla, T.S., Hunter, R., Jansson, R.K., Jany, C. S. (2001) *Pest Manag. Sci.* 57, 115-119