COVER SHEET

TITLE: Investigating the Role of Kdm3b in the Acquisition of Pluripotency

AUTHOR'S NAME: Michael Diny

MAJOR: Biology

DEPARTMENT: N/A

MENTOR: Rupa Sridharan

DEPARTMENT: Cell & Regenerative Biology

MENTOR(2): Khoa Tran

DEPARTMENT(2): Cell & Regenerative Biology

YEAR: 2016

(The following statement must be included if you want your paper included in the library's electronic repository.)

The author hereby grants to University of Wisconsin-Madison the permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

RECEIVED
APR 29 2016

ABSTRACT

Investigating the Role of Kdm3b, a Histone Demethylase, in the Acquisition of Pluripotency

Induced Pluripotent Stem Cells are somatic cells that have been reprogrammed to acquire the properties of embryonic stem cells by the overexpression of various proteins. These somatic cells have acquired pluripotency by modifying their epigenome, the chemical marks on their DNA and histone tails. These epigenetic modifications are performed by various epigenetic enzymes. A family of histone demethylases, the Kdm3's (Kdm3a, Kdm3b, Kdm3c), are of particular interest to the Sridharan lab, as it has been shown that cells lacking Kdm3b have a significantly reduced reprogramming efficiency, but cells lacking Kdm3a and Kdm3c do not. The goal of my project was to determine the structural aspects of Kdm3b that result in its specific behavior in reprogramming. This was done by creating overexpression vectors that possess Kdm3a, Kdm3b, as well as two mutant Kdm3a/Kdm3b combinations, transducing these overexpression vectors into cells lacking the Kdm3b enzyme, and observing how efficiently these cells reprogram.

Michael Ding, Biology
Author Name/Major

Author Signature

4/28/16
Date

Rupa Sridharan / Cell & Regen Biology & Wisconsin Institute for Discovery
Mentor Name/Department

Mentor Signature

Investigating the Role of Kdm3b, a Histone Demethylase, in the Acquisition of Pluripotency

By:

Michael Diny

Department:

Cell & Regenerative Biology

Epigenetics Theme, Wisconsin Institute for Discovery

Submitted to the

CALS Honors Program

University of Wisconsin – Madison

PI: Rupa Sridharan

2016

Abstract

Induced Pluripotent Stem Cells are somatic cells that have been reprogrammed to acquire the properties of embryonic stem cells by the overexpression of various proteins. These somatic cells have acquired pluripotency by modifying their epigenome, the chemical marks on their DNA and histone tails. These epigenetic modifications are performed by various epigenetic enzymes. A family of histone demethylases, the Kdm3's (Kdm3a, Kdm3b, Kdm3c), are of particular interest to the Sridharan lab, as it has been shown that cells lacking Kdm3b have a significantly reduced reprogramming efficiency, but cells lacking Kdm3a and Kdm3c do not. The goal of my project was to determine the structural aspects of Kdm3b that result in its specific behavior in reprogramming. This was done by creating overexpression vectors that possess Kdm3a, Kdm3b, as well as two mutant Kdm3a/Kdm3b combinations, transducing these overexpression vectors into cells lacking the Kdm3b enzyme, and observing how efficiently these cells reprogram.

Introduction & Literature Review

Transcription factor-mediated reprogramming of somatic cells into induced pluripotent stem cells (iPSCs) is achieved through the overexpression of several proteins. The transcription factors that mediate this process are referred to as the Yamanaka factors, OSKM: Oct4, Sox2, Klf4, and c-Myc. When cells overexpress these 4 transcription factors, it triggers the expression of a range of proteins that ultimately lead to the cells acquiring pluripotency. iPSCs are nearly identical to embryonic stem cells and share their unique pluripotency and self-renewal abilities. The reprogramming of somatic cells into iPSCs by introducing defined factors represents not only a conceptually new understanding of cell fate but also a practical way to generate cells for therapeutic applications. Many cellular and pathological processes can be analyzed using iPSC-derived models such that detailed mechanisms can be elucidated, new drug targets validated and clinically relevant small molecules developed[2, 5]. However, the reprogramming process is inefficient (<5%)[2], and the mechanism by which it occurs it not completely understood. It is known that the reprogramming process entails major changes to the epigenome and transcriptome of the cell, but there are gaps in knowledge of the molecular activity underlying these changes. In order to understand the molecular activity that occurs during reprogramming, it is important to understand the functions of the various enzymes that modify the epigenetic code during the process. Epigenetics is defined as heritable changes in gene expression that does not involve changes to the underlying DNA; essentially, epigenetics is a change in phenotype without a change in genotype. Epigenetic enzymes regulate gene expression by altering methylation of DNA or modifying histone tails, which changes chromatin structure in a way that affects the accessibility of genes. Thus, these DNA/histone modifying epigenetic enzymes have important roles in regulating gene expression and regulating cell fate and identity. The

overarching goal of the Sridharan lab is to understand the epigenetic changes that occur during the reprogramming of somatic cells to induced pluripotent stem cells and the mechanism of how these epigenetic changes are established.

Conducting research on the epigenetic modifications that occur during the reprogramming process is difficult due to the low percentage of somatic cells that successfully convert to iPSCs; thus, when studying the reprogramming process, the Sridharan uses pre-iPSCs instead of somatic cells. Pre-iPSCs are stalled intermediates of the reprogramming process, showing gene expression patterns in between that of somatic cells and iPSCs[1]. These pre-iPSCs are thus useful in researching the molecular events that occur in the late stages of reprogramming. In a recent publication from the Sridharan lab, it has been shown that combining ascorbic acid (AA) and 2i (inhibitors of Mitogen activated protein kinase & Glycogen synthase kinase) synergistically enhances the reprogramming of pre-iPSCs to iPSCs, increasing successful reprogramming up to 80%[1] (Figure 1). It is known that AA can influence several epigenetic enzymes, including Tet enzymes, which regulate DNA methylation, and Jumonji-domain (JmjC) enzymes, which regulate histone methylation[1,2]. Studies by *Klose et al.* looked into the function of these JmjC domain enzymes and found them to be the largest group of histone demethylases, including the Kdm3 family, defined by catalyzing the demethylation of Lysine residues on histone tails through an oxidation reaction requiring iron Fe(II) and α-ketoglutarate[3]. Several studies show that JmjC enzymes have roles in maintaining pluripotency and affect reprogramming efficiency[1,5,6]. Albeit receiving increased research attention in the last 9 years since their discovery, several JmjC proteins remain uncharacterized, gaps persist in understanding substrate recognition and the integration of JmjC proteins into signaling pathways is just emerging.

While trying to uncover the molecular events that cause AA and 2i to synergistically enhance reprogramming, I assisted in using a siRNA interference protocol to deplete various histone demethylase enzymes in pre-iPSCs to determine which enzymes were crucial to reprogramming. The result was that pre-iPSCs with lower levels of the enzyme Kdm3b showed significant decrease in conversion to iPSCs upon treatment with AA+2i, indicating that Kdm3b is crucial for reprogramming[1, 5] (Figure 2). The Kdm3 subfamily of JmjC histone demethylases targets Histone 3 Lysine 9 mono- and di-methylation, and include Kdm3a, Kdm3b and Kdm3c[4]. Surprisingly, in our siRNA knockdown, cells lacking Kdm3a and Kdm3c only showed a modest decrease in reprogramming efficiency in the AA+2i system (Figure 2). This is interesting, because upon analyzing the amino acid sequence of the Kdm3 subfamily, we found that these enzymes contain minimal variations in their functional enzymatic domains (JmjC domain and zinc finger), but much greater differences in uncharacterized domains, length and binding partners[4] (Figure 3). Information on their 3D structure, subdomain structure and substrate/cofactor interaction is still lacking. Our discovery suggests that although the Kdm3 subfamily has nearly identical enzymatic domains and shares substrate specificity, they have differing activity in reprogramming. The cause of this difference is unknown; it could be due to different DNA binding, different protein-protein interaction or different protein structure.

The focus of my research project was to determine if cells lacking Kdm3b can have their phenotype rescued when transduced with a Kdm3b overexpression vector, and to determine the structural aspects of Kdm3b that lead to its differing behavior in reprogramming with AA+2i. In order to determine the structural domains that are crucial for Kdm3b's specific function, structure-function studies were performed. I created four vectors using a pCDH-Neomycin-IRES overexpression vector backbone: pCDH-Kdm3a, pCDH-Kdm3b, and two mutants of

Kdm3b. One mutant contained the N-terminal of Kdm3a (non-enzymatic end) and the C-terminal of Kdm3b (enzymatic end). The other mutant contained the N-terminal of Kdm3b (non-enzymatic end) and the C-terminal of Kdm3a (enzymatic end). The purpose of these mutants was to determine if the N-terminal, the C-terminal, or both were crucial for Kdm3b function. The N-terminal of Kdm3b is over 1000 base pairs longer than the N-terminal of Kdm3a, and it is possible that this part of the enzyme is important for chaperone binding or substrate recognition, leading to the differing behavior of these enzymes.

Once the pCDH-vector constructs were created and verified by DNA sequencing, lentivirus was created for each of the constructs. Lentivirus has become a major tool for gene delivery in mammalian cells. Lentiviral vectors give the ability to mediate potent transduction and stable expression into dividing cells *in vitro*. When transduced with lentivirus, cells integrate the vector into their genome and express it. The lentivirus was used to transduce: 1) Clone 8: Pre-iPSCs that were depleted of Kdm3b via the Crispr-Cas9 endonuclease system, and 2) Clone 3: Pre-iPSCs that express normal amounts of Kdm3b. Clone 8 does not express Kdm3b and is thus expected to reprogram at a lower efficiency, according to previous experiments done by the Sridharan lab. Clone 3 was sent through the same Crispr-Cas9 protocol as Clone 8, but for these cells the Crispr-Cas9 system did not successfully knock out Kdm3b, and these cells thus express normal (wildtype) amounts of protein. Because Clone 3 is a wildtype, we expect it to reprogram at a higher efficiency. Gene knockout and Kdm3b expression for Clone 3 and Clone 8 were confirmed by Surveyor Assay and Western Blot (Figure 4).

Experimental design for reprogramming Clone 3 and Clone 8 had the following conditions: wildtype cells not transduced with anything as a control, Clone 3/8 not transduced with anything, Clone 3/8 transduced with the pCDH- overexpression vector alone (to verify that the pCDH- vector itself did not have an effect on reprogramming), Clone 3/8 transduced with pCDH-Kdm3a, Clone 3/8 transduced with pCDH-Kdm3b, and Clone 3/8 transduced with the two mutant constructs – 7 conditions in total. An illustration of this experimental design can be seen in Figure 10. After transduction, these cells were put on Neomycin selection for 5 days, as the PCDH- overexpression vector used had Neomycin resistance. This selection would kill off cells that did not successfully integrate and express the pCDH-overexpression vector into their genome. However, the selection is not 100%, so after 5 days of selection, colonies were picked from the selection plate. These colonies were then plated: 1) onto plates for growth, and 2) onto coverslips for immunofluorescence staining. Cells plated on coverslips were stained for V5 to confirm successful integration and expression of the PCDH-construct, as it contained a V5 epitope.

The cells that had positive V5 staining were sent through reprogramming in either ES Media alone or ES Media with AA + 2i. After reprogramming in ES Media/AA+2i Media for 10 days, the cells were analyzed using Flow Cytometry to measure the percent of cells that successfully reprogrammed to iPSCs. This experiment will tell us 1) if Kdm3b knock out cells can have their phenotype / reprogramming efficiency rescued when transduced with a Kdm3b overexpression vector, 2) what effect overexpression of Kdm3a has on reprogramming in Kdm3b knockout cells and wildtype cells, and 3) if the non-enzymatic domains of Kdm3a/Kdm3b play an important role in enzyme function.

Methods

Tool Generation: Creating Overexpression Vectors

The first thing to be done in this project was to create the overexpression vectors needed for the reprogramming experiments. This tool generation involves multiple steps: PCR amplification of Kdm3a, Kdm3b, and the N/C terminals of Kdm3a/Kdm3b, ligating these purified PCR products into the cloning site of the pCDH- overexpression vector, transforming the pCDH-construct into Stellar competent cells, inoculating, mini-prepping and running a diagnostic digest on the transformations, DNA sequencing, and finally maxi-prepping the final product.

In order to amplify each insert, primers had to be designed to amplify the appropriate Kdm3a/Kdm3b sequence from pMX-Kdm3a and pMX-Kdm3b already in possession by the Sridharan lab. Primers were designed to amplify full length Kdm3a/Kdm3b, and primers were designed to amplify the N/C Terminals of Kdm3a/Kdm3b. Each of these primers contained a region that overlapped with the cloning site of the pCDH- vector, in order to allow the two pieces to properly anneal. In addition, primers contained a V5 tag on the 5' end, for use in Immunofluorescence staining to identify the induced protein. Once primers were designed, Phusion enzyme was used for PCR amplification. Each insert had differing PCR conditions that I had to optimize in order to get the desired product. Successful PCR conditions for Kdm3a were HF Buffer, +DMSO, at 55°C. Successful PCR conditions for Kdm3b were HF Buffer, +DMSO at 58°C. PCR conditions for the N-terminal of Kdm3a were GC Buffer, +DMSO at 65°C. PCR conditions for the C-terminal of Kdm3a were HF Buffer, +DMSO at 62°C degrees. PCR conditions for the N-terminal and C-terminal of Kdm3b were HF Buffer, +DMSO at 58°C and

60°C, respectfully. Extension times were ran according to size, 30 seconds per kb of DNA. PCR amplifications were run on 1% agarose gels containing Ethidium Bromide and a 1kb DNA ladder. The gels were imaged in an ImageQuant 4000 imaging machine. If PCR amplification was successful, as interpreted by the amplified DNA being the correct size in correlation with the 1kb DNA ladder, the DNA was cut out of the gel under a UV lamp (Ethidium Bromide binds to the DNA and fluoresces under UV light). After successfully amplifying and cutting out the desired insert, gel purification was used to extract the DNA. An example of a successful PCR can be seen in Figure 9.

The two mutant constructs required an extra sew-together step that full length Kdm3a and Kdm3b did not require. Because the two mutants were: Kdm3a N-terminal/Kdm3b C-terminal and Kdm3b N-terminal/Kdm3a C-terminal, the two pieces had to be sewn together. When the primers were designed, they were designed so that the 3AN/3BC pieces had overlapping ends, and the 3BN/3AC pieces had overlapping ends, to allow proper annealing. This sew-in step involved allowing the two appropriate pieces to anneal at 55°C, followed by PCR amplification so as to have the entire product as one piece instead of two individual pieces. An illustration of this can be seen in Figure 5. PCR amplifications were run on 1% agarose gels containing Ethidium Bromide and a 1kb DNA ladder. After successfully amplifying and cutting out the desired mutant inserts, gel purification was used to extract the DNA.

The next step was to ligate each insert into the cloning site of the pCDH- overexpression vector. In order to do this, the pCDH- vector was digested with XbaI and BamHI restriction enzymes overnight in order to linearize the vector and expose the cloning site that overlapped with the designed insert. Once digested, the linearized pCDH- vector was ran on a 1% agarose gel containing Ethidium Bromide and a 1kb DNA ladder, followed by gel purification to extract

the DNA. Then, with each purified insert and the purified, linearized pCDH- vector, one insert and one vector were combined in a 10ul reaction containing in-Fusion HD Enzyme Mix at a ratio of 150ng insert:50ng vector and ligated at 55 degrees for 15 minutes. This was done for each construct. An illustration of this can be seen in Figure 6.

After ligating the insert and vector, the newly created construct was transformed into Stellar competent cells using standard bacterial transformation protocol. The Stellar cells were then plated on an Ampicillin plate, and left to grow overnight at 37 °C. Once colonies grew overnight, an average of 8 colonies were picked from the plate, innoculated in 3ml LB Broth +Ampicillin overnight, then mini-prepped using a Zyppy Miniprep kit to extract DNA from the bacteria. Once DNA was extracted from the picked colonies, a diagnostic digest was performed to see if any of the picked colonies were successfully transformed with the vector+insert. A diagnostic digest involved using ApE Plasmid Editor to find restriction enzymes that would cut the vector at specific, desired points in the DNA sequence. Once a restriction enzyme was found, the vector was digested in said restriction enzyme for 3 hours, then ran on a gel with a 1kb ladder in order to determine if the DNA produced by the bacteria was the same size as predicted by ApE Plasmid Editor. For example, the pCDH-Kdm3a construct was digested with restriction enzyme NcoI, and the expected product size (in base pairs) was 5276, 2352, 2222, 1410 and 574bp. An illustration of this can be seen in Figure 7. This procedure was repeated for pCDH-Kdm3b and the two mutant constructs. When a diagnostic digest looked to have the correct product (digested fragments matched what was predicted with ApE Plasmid Editor), DNA from the mini-prep was used to perform a sequencing reaction to ensure that the entire construct, vector and insert, were 100% accurate and did not contain any mutations. Once the results from sequencing verified a 100% accurate construct, the next step was to perform a maxiprep in order

to obtain a large amount of the vector+insert DNA for downstream experiments and lentiviral production.

Lentiviral Production of pCDH- Constructs

With each DNA-verified construct successfully Maxiprepped to a large amount of DNA, the next step was to use that DNA to create lentivirus. Lentivirus was created using standard lentiviral production protocol using general 293T cells and D10 Media. 293T cells were grown and passaged 3 times before transduction. 8 hours before transduction, cells were placed in D10 media containing 10mM HEPES. Cells were then transfected with 5µg of appropriate pCDH-construct maxiprep, 1µg of VsVg packaging plasmid, and 5µg of Δ8.9 envelope plasmid; as well as 450ul OPTI-Mem (no serum media) and 40µl of PEI. Cells were then left to grow for two days. After the two days, the virus-containing supernatant media was harvested, passed through a 0.45µm filter, transferred to a 15mL conical, and concentrated using LentiX Concentrator overnight. After the lentivirus was concentrated, the solution was spun down at 15,000 x gravity for 45 minutes, and the subsequent pellet was resuspended in 100ul PBS. The virus was then aliquoted into 10ul aliquots for future use, and stored at -80°C.

Cell Transduction

With lentivirus of each pCDH-construct created, the next step was to thaw, grow, and transduce cells with the lentivirus.

Clone 3 and Clone 8 were used to explore the effects of variable Kdm3 expression. Recall that Clone 8 is a pre-iPSC, Kdm3b Knock-Out cell line via the Crispr-Cas9 system, while Clone 3 is a pre-iPSC cell line expressing normal levels of protein. These cells had gone

through screening via both Surveyor Assay and Western Blot to confirm appropriate Kdm3b expression (Figure 4). Kdm3b Knock-Out cells have been shown to reprogram at a reduced efficiency in AA+2i. As a control, Wildtype pre-iPSCs expressing normal levels of protein and not transduced with any lentivirus were reprogrammed alongside Clone 3/Clone 8. These cells have been shown to reprogram at a much higher efficiency in AA+2i.

The clones were first thawed in a 10cm dish, and grown until healthy and confluent. Once confluent, the cells were trypsonized, spun down at 300 x gravity for 3 minutes, resuspended to 6mL, and plated evenly in a 6-well dish, 1 mL per well. Each individual well of the dish was transduced with 30µl of lentiviral construct: one well transduced with pCDH-Empty, one with pCDH-Kdm3a, one with pCDH-Kdm3b, one with pCDH-3AN/3BC, one with pCDH-3BN/3AC, and one that was not transduced with anything. Note that there was a 7^{th} well with Wildtype pre-iPSCs as a control. The Wildtype control was not transduced with anything. An illustration of this can be seen in Figure 10. After 18 hours of exposure to virus, the cells were put on selection media – ES Media with 500µg/mL Neomycin. The cells that were not transduced with any virus and the Wildtype did not get put on selection as they did not have Neomycin resistance. Cells were left on selection for 5 days, replacing selection media every other day. After 5 days of selection, the cells that were alive were ones that successfully integrated the pCDH- construct into their genome, and thus had acquired Neomycin resistance and survived the 5 day selection.

However, the selection process is not 100%; thus, after the 5 day selection, a fraction of the cells were plated and fixed onto coverslips for Immunofluorescence staining. These cells were stained for DAPI (control) and V5. Staining for V5 confirmed that the cells successfully integrated and expressed the pCDH- construct, as there was a V5 tag on it (Figure 8).

Reprogramming

The cells that survived the 5 day selection and tested positive for V5 were then sent into reprogramming, following the reprogramming procedures by *Tran et al.* This involved growing the cells in in ES Media containing Ascorbic Acid (50μg/ml) and 2i (1μM, 3μM). As a control, cells were also reprogrammed in ES Media alone to compare how efficiently the cells reprogrammed without AA+2i. We expect the cells grown in AA+2i media to reprogram more efficiently. On Day 0, the start of the reprogramming experiment, the cells were trypsonized, spun down, resuspended, counted on a hemocytometer, and plated at 200,000 cells per well. On day 4, the cells were again trypsonized, spun down, resuspended, counted on a hemocytometer, and plated at 200,000 cells per well to prevent overcrowding. Media was replaced every other day. On Day 10 the experiment was complete, and the cells were sent through Flow Cytometry.

Results Analysis

Nanog is the protein used to determine if cells have acquired pluripotency, as Nanog is a protein only expressed by stem cells. Thus, cells that express Nanog have successfully reprogrammed to iPSCs. The cells we used had been engineered to express a Green Fluorescent Protein attached to Nanog. With GFP attached to Nanog, we could visualize Nanog under the fluorescent microscope and with Flow Cytometry. Nanog-GFP positive colonies were first counted under the fluorescent microscope to get a rough estimate as to how many Nanog positive colonies there were. Then, cells were analyzed using Flow Cytometry. Flow Cytometry is a more accurate way to analyze how many cells truly express Nanog, and gives a percentage of cells that express Nanog compared to the whole cell population. In Flow Cytometry, single cells are passed through a laser, and if the cell fluoresces as it passes through the laser, a computer

records it. Thus, cells that had successfully reprogrammed and express Nanog-GFP were counted as they passed through the laser.

Results & Discussion

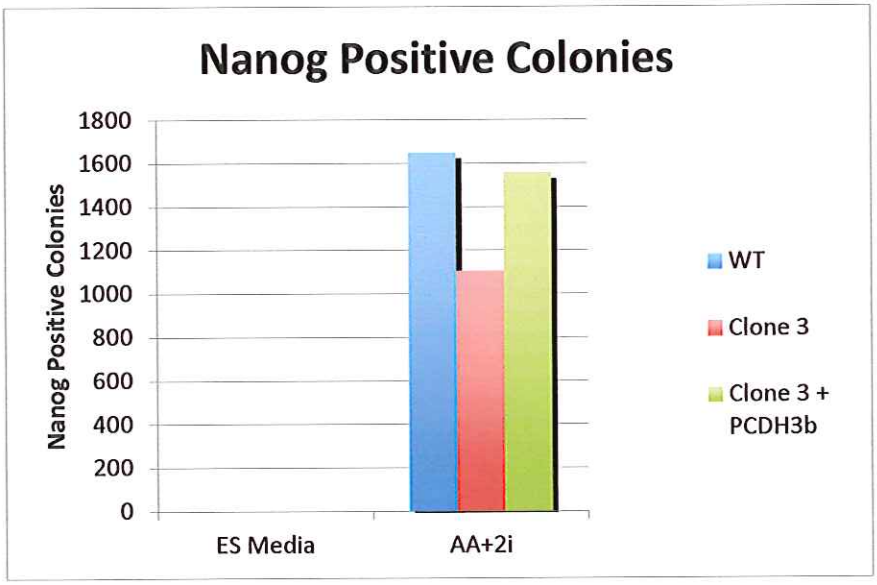
Lack of time hindered my ability to finish every aspect of this project. I was unable to complete all reprogramming experiments for Clone 3 and Clone 8, as I was unable to generate all of the overexpression vectors/lentiviral constructs early enough in the year to do so. Reprogramming Clone 3 & Clone 8 with all the lentiviral constructs will be completed by the undergraduate student who will take over my project, as the results of this experiment are of interest to the Sridharan lab.

However, I was able to reprogram Clone 3 and Clone 8 with pCDH-Kdm3a and pCDH-Kdm3b. Clone 3 was transduced with pCDH-Kdm3b, while Clone 8 was transduced with both pCDH-Kdm3a and Kdm3b. These clones were transduced with the overexpression vectors, put on selection, and sent through reprogramming as described above. The only difference being, they were not transduced with the mutant constructs or pCDH-Empty.

Clone 3 Results

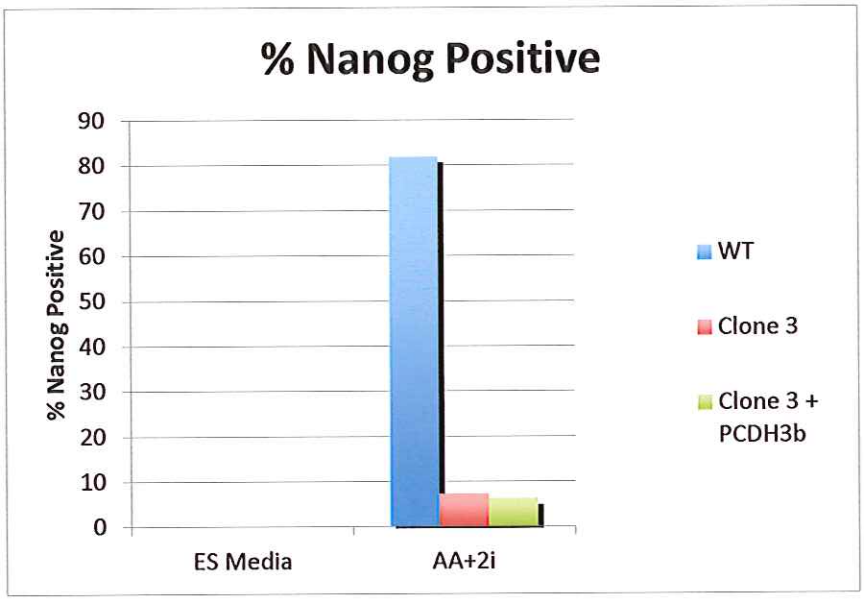
Recall that Clone 3 is a cell line of pre-iPSCs that went through the Crispr-Cas9 system with Clone 8, but had unsuccessful Kdm3b targeting, and thus expressed normal (wildtype) levels of protein. The reprogramming for Clone 3 had the following conditions: Wildtype, Clone 3 not transduced with anything, and Clone 3 transduced with pCDH-Kdm3b. These cells were transduced, selected for and reprogrammed as described above. On Day 10 these cells were counted under the fluorescent microscope and sent through Flow Cytometry. The results are as follows:

Nanog Positive Colonies	Wildtype	Clone 3	Clone 3 + pCDH-Kdm3b
ES Media	0	0	0
AA+2i	1646	1104	1554



A more accurate way to determine the percent of cells that truly converted to iPSCs and express Nanog is by using Flow Cytometry. These Kdm3b knockout cells were also counted using Flow Cytometry, and the results are as follows:

% Nanog Positive	Wildtype	Clone 3 Alone	Clone 3 + pCDH-Kdm3b
ES Media	0.095	0.063	0.042
AA+2i	81.8	7.17	6.13



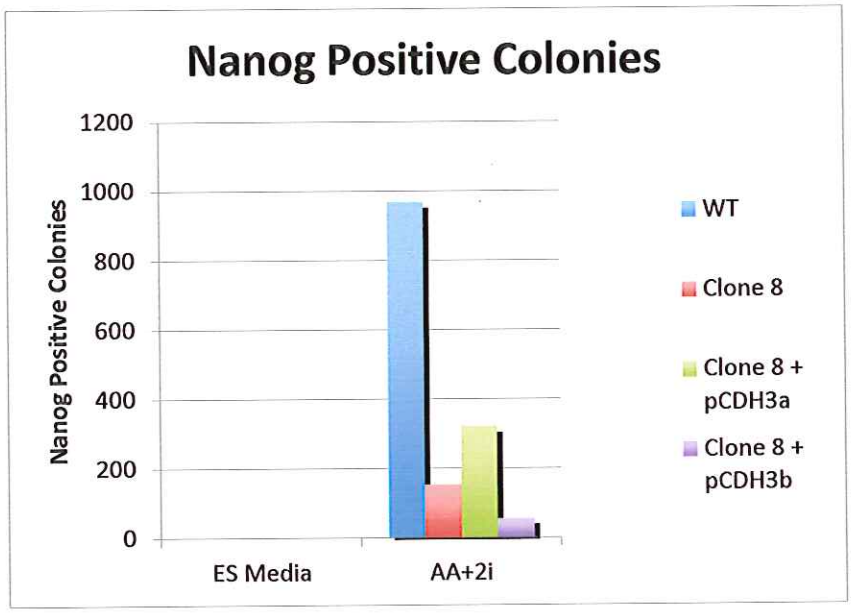
The results from both colony counts and Flow Cytometry confirmed that cells reprogrammed in ES Media alone reprogram at a much lower efficiency, if at all, compared to cells reprogrammed in the AA+2i media.

It is clear that the number of Nanog positive colonies counted under the microscope does not align with the actual percent of cells that express Nanog as determined by Flow Cytometry. Under the fluorescent microscope, the Nanog positive colonies counted were over 1000 for each condition. They were relatively similar, within a few hundred of each other. This suggests that all three of these conditions, Wildtype, Clone 3, and Clone 3+pCDH-Kdm3b, reprogrammed at a similar efficiency. This makes sense, since Wildtype and Clone 3 are both pre-iPSCs that express normal endogenous amounts of protein. However, when sent through Flow Cytometry, we see that less than 10% of Clone 3 cells truly reprogrammed, while Wildtype reprogrammed at the expected 80%. It took us many months to figure out why these cells were not reprogramming as expected. When it was confirmed that Clone 8 was a successful Kdm3b knock-out cell line and Clone 3 was an unsuccessful Kdm3b knock-out cell line, these cells were grown in large plates and passaged many times to be frozen down for future use. It was discovered that perhaps the multiple passages had altered the cells capability to reprogram in AA+2i. Cells from early passages reprogram as expected, but cells from the later passages have a reduced reprogramming efficiency. These Clone 3 cells were from a late passage, and that is why their reprogramming efficiency is much lower than the expected 80%. This experiment will need to be repeated with: 1) all of the pCDH- constructs, and 2) Clone 3 cells from an early passage, in order to get concrete results. Results from this experiment will tell us how pre-iPSCs that overexpress various Kdm3 enzymes/mutants reprogram in AA+2i.

Kdm3b Knock-Out pre-iPSCs: Clone 8 Results

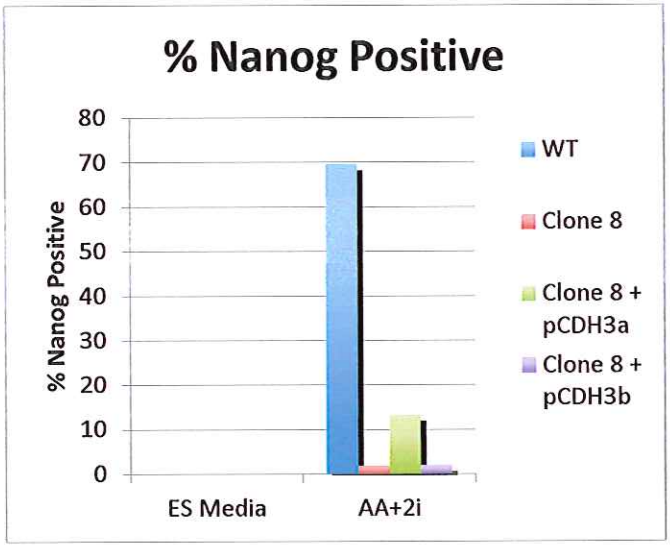
I was also able to complete the reprogramming experiment for Clone 8. Recall that Clone 8 is a cell line of pre-iPSCs that had successful Kdm3b knock-out VIA the Crispr-Cas9 system. The reprogramming experiment for Clone 8 had the following conditions: A wild type not transduced with anything, Clone 8 not transduced with anything, Clone 8 transduced with pCDH-Kdm3a, and Clone 8 transduced with pCDH-Kdm3b. These cells were transduced, selected for and reprogrammed as described above. On Day 14 these cells were counted under the fluorescent microscope and sent through Flow Cytometry. The results are as follows:

Nanog Positive Colonies	Wildtype	Clone 8	Clone 8 + pCDH-Kdm3a	Clone 8 + pCDH-Kdm3b
ES Media	0	0	0	0
AA+2i	966	153	320	55



A more accurate way to determine the percent of cells that truly converted to iPSCs and express Nanog is by using Flow Cytometry. These Kdm3b knockout cells were also counted using Flow Cytometry, and the results are as follows:

% Nanog Positive	Wildtype	Clone 8	Clone 8 + pCDH-Kdm3a	Clone 8 + pCDH-Kdm3b
ES Media	0.1	0.016	0	0.012
AA+2i	69.4	1.75	13.2	1.85



It is clear to see that cells reprogrammed in ES Media alone reprogrammed much less efficiently, if at all, compared to cells reprogrammed in ES Media with AA+2i.

Results from both colony counts and Flow Cytometry indicate that cells transduced with pCDH-Kdm3a reprogrammed at a higher efficiency than Clone 8 alone and Clone 8 transduced with pCDH-Kdm3b, but lower than Wildtype cells. This indicates that overexpression of Kdm3a has some phenotype rescue effect.

Since these cells are Kdm3b knock-outs, we expect reprogramming efficiency to increase when these cells are transduced with the Kdm3b overexpression vector and begin expressing Kdm3b again. However, it is clear that this is not the case. These cells were confirmed to be expressing Kdm3b once transduced with pCDH-Kdm3b VIA Immunofluorescence staining for V5 (Figure 8). A possible explanation as to why Clone 8 + pCDH-Kdm3b could not have its phenotype rescued is because the amount of Kdm3b being expressed due to our overexpression vector was much different than the amount of Kdm3b endogenously expressed by the cell. If the amount of Kdm3b expressed due to our overexpression vector is vastly lower than the amount of Kdm3b expressed endogenously by the cell, this could account for why the cell is still inefficient at reprogramming even when introduced to the pCDH-Kdm3b overexpression vector. A way to test this hypothesis would be by doing an rtPCR to see the real time mRNA expression levels of Wildtype, Clone 8 and Clone 8 + pCDH-Kdm3b, and comparing the amount of Kdm3b mRNA expressed between these three groups.

An additional explanation as to why Kdm3b KO cells could not have their phenotype rescued when transduced with a Kdm3b overexpression vector is our attachment of a V5 tag to the 5' end of Kdm3b. When primers were designed, they were designed to make Kdm3b have a V5 epitope at its 5' end for use in Immunofluorescence staining to confirm expression of the

induced protein. It is possible that the 5' tip of Kdm3b is crucial for its function, and thus by placing a V5 tag there, Kdm3b is unable to perform its function.

It is interesting to see that cells transduced with pCDH-Kdm3a have an increase in reprogramming efficiency. This could indicate that overexpression of Kdm3a above endogenous levels leads to Kdm3a overcompensating its function and performing the same function as Kdm3b. This is plausible, since Kdm3a and Kdm3b have the same H3K9 mono and di-methyl target.

This experiment will need to be repeated with all the pCDH- constructs. Results from this experiment will tell us: 1) If Kdm3b knock-out cells can have their phenotype rescued / reprogramming efficiency increased when introduced to a Kdm3b overexpression vector, and 2) How overexpressing various Kdm3 enzymes/mutants affects Kdm3b knock-out cells reprogramming in AA+2i.

Tables & Figures

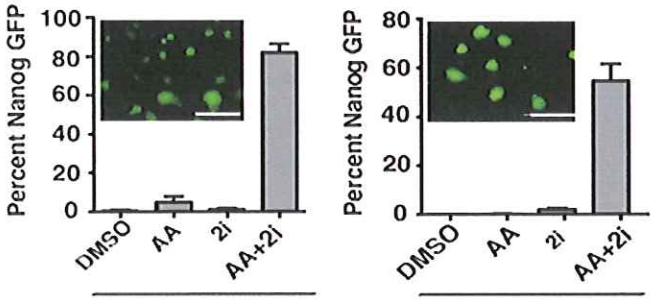


Figure 1: This figure illustrates the increase in successful reprogramming of pre-iPSCs to iPSCs when treated with AA+2i. Nanog is a gene that is only expressed in iPSCs, so Nanog positive colonies represent true iPSCs.

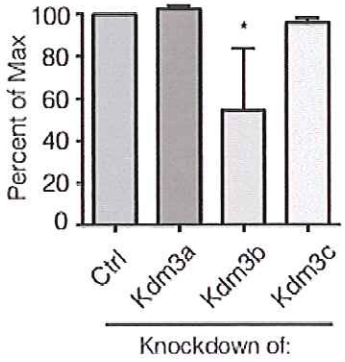


Figure 2: This figure shows that a decrease in Kdm3b decreases reprogramming efficiency in AA+2i significantly. Note that this decrease is not seen for Kdm3a or Kdm3c, which is surprising as the Kdm3s share H3K9me1/2 substrate specificity.

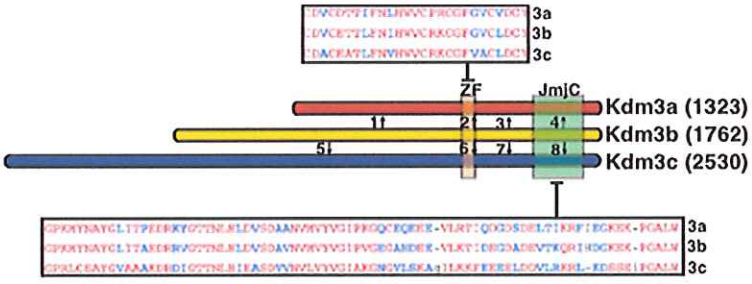


Figure 3: This figure shows the functional differences between the Kdm3 subfamily. There are slight variations in amino acid sequence in the functional zinc finger and JmjC domains. However, there are greater differences in non-functional domains, as well as large differences in length.

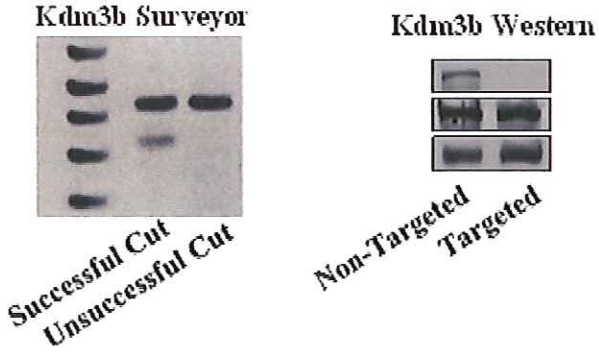


Figure 4: This figure shows the successful knockout of Kdm3b using the Crispr/Cas9 system. On the left, the surveyor assay shows successful homozygous cutting of the Kdm3b gene. On the right, the western blot shows no Kdm3b present in the cell.

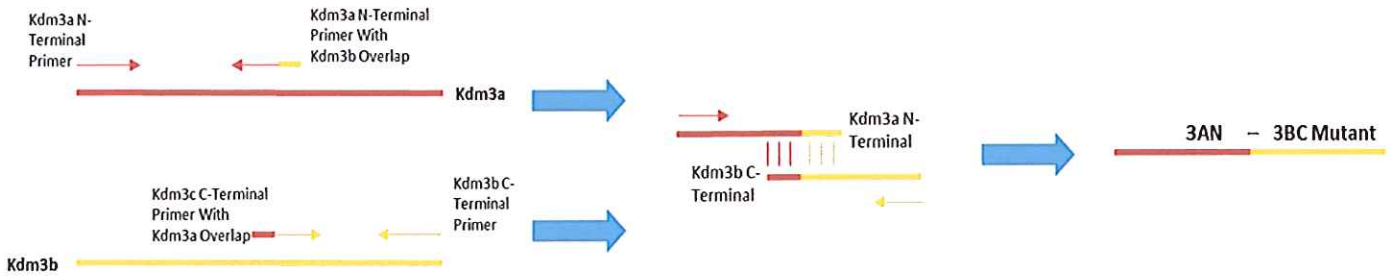


Figure 5: This figure illustrates the construction of a mutant construct, Kdm3a N-terminal + Kdm3b C-terminal. First, the N-Terminal is amplified from Kdm3a, designed to have a Kdm3b overlap on its 3' end. Then, the C-terminal is amplified from Kdm3b, designed to have Kdm3a overlap on its 5' end. These two pieces are then annealed at 55°C, and then amplified into one piece.

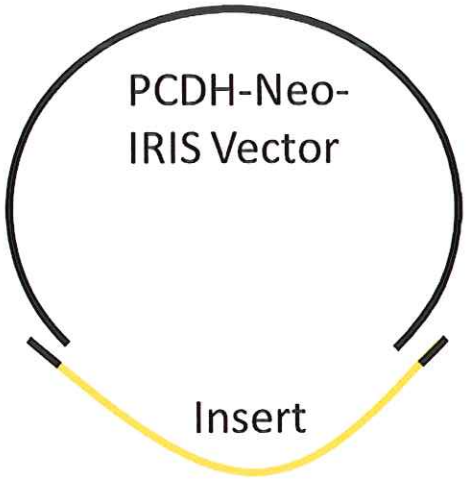


Figure 6: This figure illustrates the ligation of the pCDH- overexpression vector and a purified PCR insert. Vector/insert were combined at a ratio of 150ng vector:50ng insert with inPhusion HD enzyme mix and allowed to ligate for 15 minutes at 55°C.

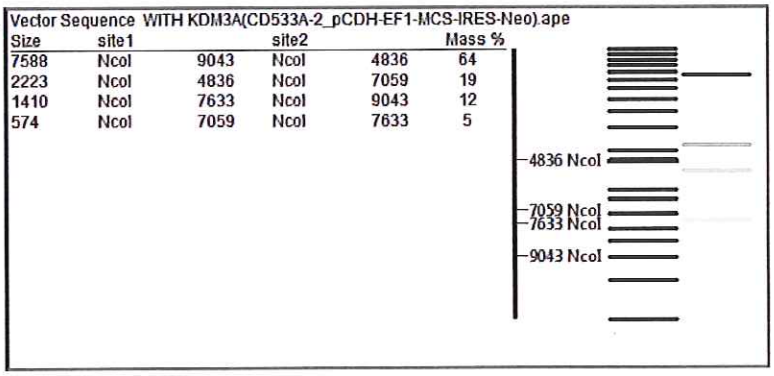


Figure 7: This figure illustrates the use of ApE Plasmid Editor to predict how restriction enzyme NcoI will cut pCDH--Kdm3a as a diagnostic digest. pCDH--Kdm3a is transformed into bacterial cells, the cells are grown for 18 hours, and then DNA is harvested from the cells using a Zyppy Miniprep kit. This DNA is then digested in NcoI, and compared to the digest predicted by ApE.

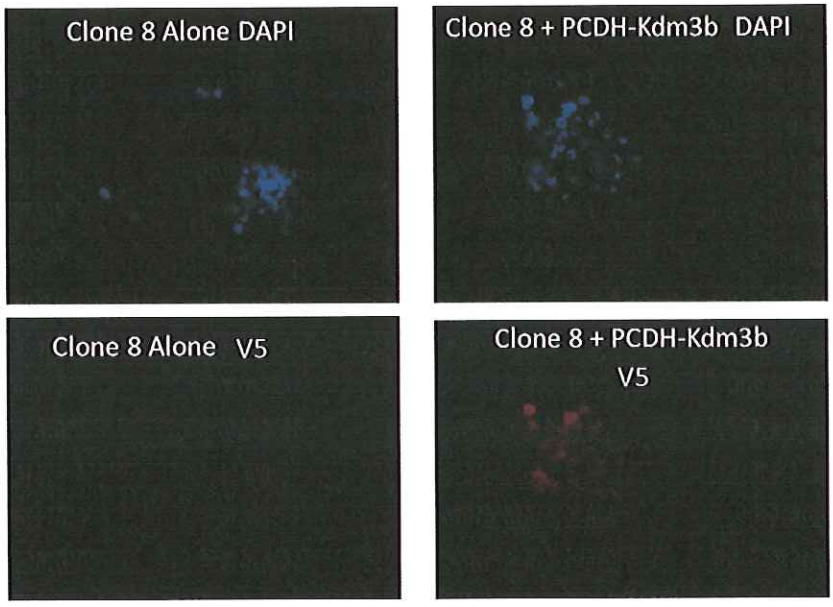


Figure 8: These are Immunofluorescence staining pictures captured from the fluorescent microscope. On the left is Clone 8 alone, stained with DAPI and V5. As you can see, Clone 8 alone does not express V5/Kdm3b. On the right is Clone 8 transduced with pCDH--Kdm3b, stained with DAPI and V5. There is visible V5/Kdm3b staining.

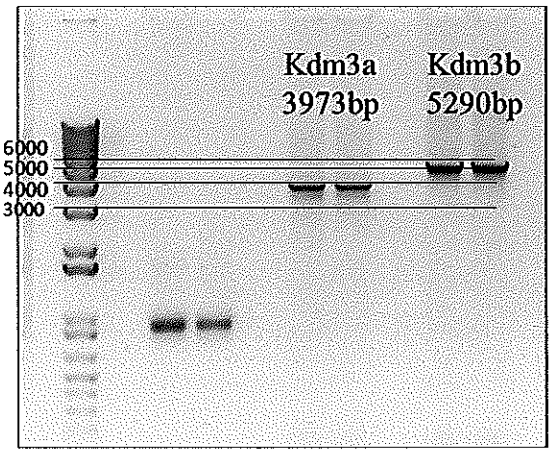


Figure 9: This is an example of a successful PCR. The expected sizes of Kdm3a and Kdm3b line up with the 1kb DNA ladder. These would be cut out of the gel under a UV lamp, then the DNA would be gel extracted.

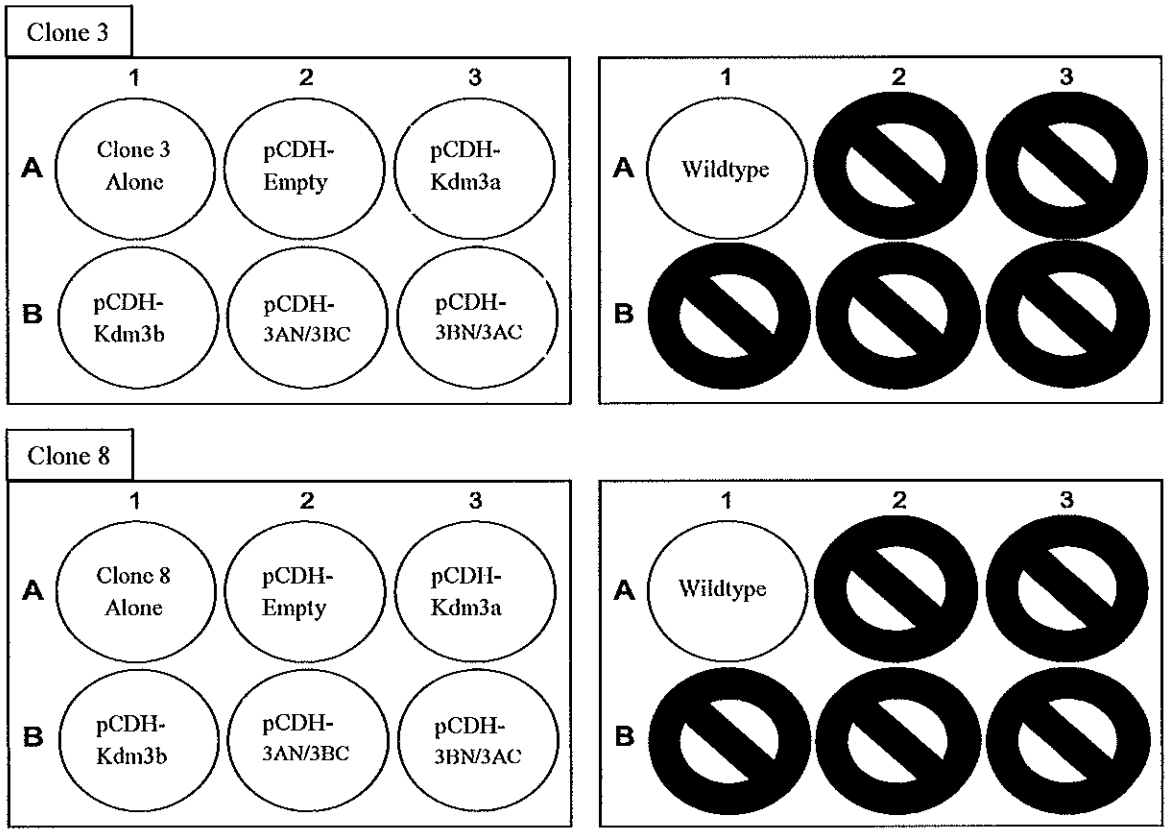


Figure 10: This figure illustrates the experimental design for the lentiviral transduction & reprogramming experiment.

References

1. Tran, K. et al. Collaborative rewiring of the pluripotency network by chromatin and signal modulating pathways. *Nature Communications.* (2014).
2. Jackson, S. A. & Sridharan, R. Peering into the black box of reprogramming to the pluripotent state. *Curr. Pathobiol. Rep.* 1, 129–136 (2013).
3. Klose, R. et al. JmjC-domain-containing proteins and histone demethylation. *Nature Review Genetics.* (2006).
4. Brauchle, M. et al. Protein complex interactor analysis and differential activity of Kdm3 subfamily members towards H3K9 methylation. *Plos One.* (2013).
5. Chen, J. et al. H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nature Genetics.* (2013).
6. Rotilli, D. Mai, I. et al. Targeting Histone Demethylases, A New Avenue for the Fight against Cancer. *Genes & Cancer.* (2011).
7. Kasioulis, I. Kdm3a lysine demethylase is an Hsp90 client required for cytoskeletal rearrangements during spermatogenesis. *Mol Biol Cell.* (2014).
8. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas system. *Science Magazine.* (2013).
9. Oost, J. New tool for genome surgery. *Molecular Biology.* (2013).