5-2011

Generation of human iPS cells using direct reprogramming proteins

Nick Blazic
University of Arkansas, Fayetteville

UNIVERSITY OF ARKANSAS
BIOLOGICAL AND AGRICULTURAL ENGINEERING DEPARTMENT

Generation of Human iPS Cells Using Direct Reprogramming Proteins

---

Nick Blazic

4/28/2011

Acknowledgements

I would personally like to thank Dr. Kaiming Ye for providing me with the opportunity/funding me to work in his lab for the summer of 2010 through his NSF REU grant, as well as allowing me to work in the lab throughout the year using a variety of different kits and conduct experiments. I am grateful for the time I was able to spend in the lab throughout the last year, and I learned a lot about how research is conducted, as well as how difficult research work can be.

I am also thankful to the graduate students/ Ph.D students that assisted me in the lab; Lingyun Zhou, Lu Zhang, and Jithesh Veetil. They taught me a lot about how to use the different lab kits and how to run specific lab experiments, as well as assisting me in designing different sequences and primers that were used in the experiments.

Finally, I would like to thank my other two committee members, Dr. Carl Griffis and Dr. Julie Carrier for the time that they spent reviewing my thesis paper and for the insights into my project that they will provide.

Table of Contents

Introduction

Human induced pluripotent stem (hiPS) cells present the opportunity to advance tissue engineering and regenerative medicine for treating many diseases such as diabetes, Parkinson's disease, spinal cord injury, and cancer. In recent years, stem cell research has made progress in advancing cell therapy techniques. Similar to human embryonic stem (hES) cells, hiPS cells can differentiate into the three germ layers, which include mesorderm, ectoderm, and endoderm. Thus, they have the potential of being differentiated into all types of cells that are required for tissue repair and regeneration. The advantage that hiPS cells have over hES cells is that they can be generated from patients' own cell sources. For example, they can be generated from somatic cells such as foreskin, skin cells, etc. Therefore, their utility for cell therapy could potentially eliminate the immune rejection that is inherent to hES cell-based cell therapies under current medical settings. Of particular interest, hiPS cell generation does not require the use of human embryos, which eliminates the moral and ethical concerns about hES cell research. However, the generation of hiPS cells from somatic cells still poses significant technical challenge. The efficiency of hiPS cell generation has been extremely low in the lab. While the use of viral-vectors for inducing the nuclear reprogramming has shown remarkable benefits in increasing the nuclear reprogramming efficiency, the use of oncogenes in these processes could potentially lead to the development of cancers in recipients who receive the transplantation of hiPS cell-derived cells during cell replacement therapy. To address these issues, a number of non-viral vector strategies have been explored and developed recently. Among these techniques, the use of transcription factors for nuclear reprogramming has been drawn lots of attentions. In order to use these transcription factors for nuclear reprogramming, one will have to be able to produce them

in the laboratory. In this work, we explore the feasibility of producing these transcription factors from *Pichia pastoris*.

Technological Research Background:

In 2006, the first successful generation of hiPS cells was documented by Takahasi and Yamanaka, who used retrovirus vectors encoding four transcription factors, Oct4, Sox2, Klf4, and C-myc to reprogram mouse adult fibroblast cells into mouse iPS (miPS) cells (1). In 2007, Takahashi and colleagues generated hiPS cells from adult human dermal fibroblast cells using retroviral transduction of the four transcription factors (2). To date, viral transduction has been the most successful method in generating iPS cells, but viral transduction raises concerns on their safe use in clinics. Therefore, a better technique for hiPS cell generation is highly desired. The problem with viral transduction is the use of oncogenes for nuclear reprogramming, which could lead to possible unpredictable genetic dysfunction such as cancer. Takahashi and colleagues were able to differentiate iPS cells into beating cardiomyocyte cells, which is a huge advancement in the biomedical engineering field (2). Yu and colleagues were able to generate hiPS cells from human somatic cells using nuclear transfer to deliver four transcription factors, *OCT4*, *SOX2*, *NANOG*, and *LIN28* (3). IPS cell technology potentially could overcome two important obstacles associated with hES cells, such as immune rejection and ethical concerns (4).

Along similar research directions, many efforts have been made to develop non-viral hiPS cell induction approaches. Kaji and colleagues used a single multi-protein expression vector composed of the four transcription factors to generate hiPS cells from human and mouse fibroblasts (5). However, the efficiency was still not high enough to deem the technique a success. Melton's research team was able to generate hiPS cells from human fibroblasts using only two of the known transcription factors, Oct4 and Sox2, along with valproic acid. The

purpose of their experiment was to try to generate iPS cells without the use of the oncogenes, Klf4 and c-myc. However, the efficiency was low, as only one colony was formed from nearly 100,000 cells, which is less than 0.001% efficiency (6).

One discovery made the development of non-viral hiPS cell induction possible. Just recently, various cell penetrating peptides (CPPs) have been discovered and can help to deliver proteins into targeted cells (7, 8). By linking the desired protein to these cell penetrating peptides, such as a poly-Arginine peptide, the protein is able to enter the cell by a mechanism using translocation into the cytoplasm**.** A company called STEMGENT (Cambridge, MA) has created and is marketing four reprogramming recombinant Human Protein Sets with each having one of the four required transcriptional factors (9). The STEMGENT proteins were produced from *E. coli* in inclusion bodies, refolded and purified. However, these proteins have not been proven to work efficiently on human somatic cells, as the researchers generated iPS cells from mouse fibroblasts. The efficiency of hiPS cell generation with these proteins was only 0.001 % (9). Further optimization of the generation of iPS cells is required before the technology can be actually used for generating hiPS cells and used successfully in a clinical setting.

Scope of Project:

Accordingly, hiPS cells are very promising for seeking cures to many diseases. In light of their tremendous promise, the existing techniques for generating hiPS cells suffer from several disadvantages, such as the requirement of using viral vectors for nuclear reprogramming of somatic cells (skin) to form hiPS cells. The possibility was explored of using a non-viral approach to reprogram somatic cells into hiPS cells and to then further differentiate them into glucose-responsive, insulin-secreting cells for islet transplantation. This project proposes to

develop a *Pichia pastoris* protein expression system to produce the four transcriptional factors needed for somatic cell reprogramming. *P. pastoris* is yeast that is commonly used as a heteroprotein expression system. *P. pastoris* has been used in the lab and hopefully will be able to produce proteins at a higher efficiency than *E.coli* is able to do. The reasoning behind using *P. pastoris* is that a hypothesis has been developed that the transcriptional factors required from induction of iPS cells can be functionally produced using *P. pastoris*, which will allow for post-translational modifications of proteins to take place and would improve efficiency of iPS cell generation. The high degree of glycosylation capacity of the *Pichia* protein expression system will allow for production of fully functional transcription factors that are essential to induce the reprogramming of somatic cells into hiPS cells. The advantage that *P. pastoris* has over *E. coli* is that *E. coli* is unable to glycosylate proteins, The four transcriptional factors that are needed for nuclear reprogramming are Sox2, Oct4, Klf4, and C-myc. Using these recombinant proteins, we will investigate whether applying these four transcriptions factors to human foreskin cells could lead to efficient reprogramming of foreskin cells into iPS cells.

Materials / Methods and Results:

1- Experimental Design

The first step of this work was to acquire the DNA sequence encoding these four transcription factors. The NCBI blast was used to determine the sequences of Sox2, Oct4, Klf4, and C-myc. NCBI blast is a website that allows for a gene sequence/ protein to be searched, and if a match is found within the database of an uploaded plasmid, the entire sequence will be given on the website. Once all four of the transcription factor sequences were determined, the next step was to insert them into a software program called DSgene (Accelrys, Inc. San Diego,CA).

DSgene allowed for different gene sequences to be inputted and translated into the amino acid sequence. Since the amino acid sequences were known from the STEMGENT plasmids(9), the NCBI(National Center for Biotechnology Information) blast sequences were double- checked using DSgene and compared to the known STEMGENT amino acid sequences of Sox2, c-myc, Klf4, and Oct4. The reason that NCBI blast was used was because STEMGENT is a company that wanted to sell their plasmids for profit, so they did not want to give the sequence of the transcription factors away. The translated sequences matched the amino acid sequences of each of the STEMGENT transcription factor sequences, so it was concluded that the NCBI sequences were correct. Figure 1 and 2 below document the sequences that were discovered through NCBI blast, as well as the corresponding amino acid sequences that were outputted from DSgene and compared to the given STEMGENT amino acid sequences. The importance of checking the sequences was due to the fact that STEMGENT did not give the base pair sequences, but rather gave the amino acid sequences. By translating the NCBI blast sequences into amino acids and comparing to the STEMGENT, it was determined if the sequences found were correct or not.

SOX2 original blasted sequence from NCBI

atgctcccggccgccatggcggccgcgggaattcgattccgcatgtacaacatgatggag
acggagctgaagccgccgggcccgcagcaaacttcgggggcggcggcggcaactccacc
gcggcggcggccggcggcaaccagaaaaacagcccggaccgcgtcaagcggcccatgaat
gccttcatggtgtggtcccgcgggcagcggcgcaagatggcccaggagaaccccaagatg
cacaactcggagatcagcaagcgcctgggcgccgagtggaaacttttgtcggagacggag
aagcggccgttcatcgacgaggctaagcggctgcgagcgctgcacatgaaggagcacccg
gattataaatacccggcccccggcggaaaaccaagacgctcatgaagaaggataagtacacg
ctgcccggcgggctgctggccccggcggcaatagcatggcgagcgggtcggggtgggc
gccggcctgggcgcgggcgtgaaccagcgcatggacagttacgcgcacatgaacggctgg
agcaacggcagctacagcatgatgcaggaccagctgggctacccgcagcacccgggcctc
aatgcgcacggcgcagcgcagatgcagcccatgcaccgctacgacgtgagcgccctgcag
tacaactccatgaccagctcgcagacctacatgaacggctcgcccacctacagcatgtcc
tactcgcagcagggcacccctggcatggctcttggctccatgggttcggtggtcaagtcc
gaggccagctccagccccctgtggttacctcttcctcccactccagggcgccctgccag
gccggggacctccgggacatgatcagcatgtatctccccggcgccgaggtgccggaaccc
gccgcccccagcagacttcacatgtcccagcactaccagagcggcccggtgcccggcacg
gccattaacggcacactgcccctctcacacatgtga

Oct4 original blasted sequence NCBI

atggcgggacacctggcttcggatttcgccttctcgcccccctccaggtggtggaggtgat
gggccaggggggccggagccgggctgggttgatcctcggacctggctaagcttccaaggc
cctcctggagggccaggaatcgggccgggggttgggccaggctctgaggtgtgggggatt
cccccatgcccccgccgtatgagttctgtgggggggatggcgtactgtgggccccaggtt
ggagtggggctagtgccccaaggcggcttggagacctctcagcctgagggcgaagcagga
gtcggggtggagagcaactccgatggggcctccccggagccctgcaccgtcacccctggt
gccgtgaagctggagaaggagaagctggagcaaaaccccggaggagtcccaggacatcaaa
gctctgcagaaagaactcgagcaatttgccaagctcctgaagcagaagaggatcaccctg
ggatatacacaggccgatgtggggctcaccctgggggttctatttgggaaggtattcagc
caaacgaccatctgccgctttgaggctctgcagcttagcttcaagaacatgtgtaagctg
cggcccttgctgcagaagtgggtggaggaagctgacaacaatgaaaatcttcaggagata
tgcaaagcagaaaccctcgtgcaggcccgaaagagaaagcgaaccagtatcgagaaccga
gtgagaggcaacctggagaatttgttcctgcagtgcccgaaacccacactgcagcagatc
agccacatcgcccagcagcttgggctcgagaaggatgtggtccgagtgtggttctgtaac
cggcgccagaaggcaagcgatcaagcagcgactatgcacaacgagaggattttgaggct
gctgggtctccttctcaggggaccagtgtcctttcctctggcccagggccccatttt
ggtaccccaggctatgggagccctcacttcactgcactgtactcctcggtccctttccct
gaggggaagcctttccccctgtctccgtcaccactctgggctctcccatgcattcaaac

Klf4 original blasted sequence from NCBI

atggctgtcagcgacgcgctgctcccatctttctccacgttcgcgtctggcccggcggga
agggagaagacactgcgtcaagcaggtgccccgaataaccgctggcgggagggagctctcc
cacatgaagcgacttcccccagtgcttcccggccgcccctatgacctggcggcggcgacc
gtggccacagacctggagagcggcggagccggtgcggcttgcggcggtagcaacctggcg
cccctacctcggagagagaccgaggagttcaacgatctcctggacctggactttattctc
tccaattcgctgacccatcctccggagtcagtggccgccaccgtgtcctcgtcagcgtca
gcctcctcttcgtcgtcgccgtcgagcagcggccctgccagcgcgccctccacctgcagc
ttcacctatccgatccgggccgggaacgacccgggcgtggcgccgggcggcacgggcgga
ggcctcctctatggcagggagtccgctccccctccgacggctccctttcaacctggcggac
atcaacgacgtgagcccctcgggcggcttcgtggccgagctcctgcggccagaattggac
ccggtgtacattccgccgcagcagccgcagccgccaggtggcgggctgatgggcaagttc
gtgctgaaggcgtcgctgagccgcccctggcagcgagtacggcagcccgtcggtcatcagc
gtcagcaaaggcagccctgacggcagccaccccggtggtggtggcgccctacaacggcggg
ccgccgcgcacgtgcccccaagatcaagcaggaggcggtctcttcgtgcacccacttgggc
gctggaccccctctcagcaatggccaccggccggctgcacacgactttcccccttgggacgg
cagctccccagcaggactaccccgaccctgggtcttgaggaagtgctgagcagcagggac
tgtcaccctgccctgccgcttcctcccggctccatcccacccgggggcccaattaccca
tccttcctgcccgatcagatgcagccgcaagtcccgccgctccattaccaagagctcatg
ccacccggttcctgcatgccagaggagcccaagccaaagagggaagacgatcgtggccc
cggaaaaggaccgccacccacacttgtgattacgcgggctgcggcaaaacctacacaaag
agttcccatctcaaggcacacctgcgaacccacacaggtgagaaaccttaccactgtgac
tgggacggctgtggatggaaattcgcccgctcagatgaactgaccaggcactaccgtaaa
cacacggggcaccgcccgttccagtgccaaaaatgcgaccgagcattttccaggtcggac
cacctcgccttacacatgaagaggcatttttaa

C-MYC original blasted sequence from NCBI

ATGCCCCTCAACGTTAGCTTCACCAACAGGAACTATGACCTCGACTACGACTCGGTGCAG
CCGTATTTCTACTGCGACGAGGAGGAGAACTTCTACCAGCAGCAGCAGAGCGAGCTG
CAGCCCCCGGCGCCCAGCGAGGATATCTGGAAGAAATTCGAGCTGCTGCCCACCCCGCCC
CTGTCCCCTAGCCGCCGCTCCGGGCTCTGCTCGCCCTCCTACGTTGCGGTCACACCCTTC
TCCCTTCGGGGAGACAACGACGGCGGTGGCGGGAGCTTCTCCACGGCCGACCAGCTGGAG
ATGGTGACCGAGCTGCTGGGAGGAGACATGGTGAACCAGAGTTTCATCTGCGACCCGGAC
GACGAGACCTTCATCAAAAACATCATCATCCAGGACTGTATGTGGAGCGGCTTCTCGGCC
GCCGCCAAGCTCGTCTCAGAGAAGCTGGCCTCCTACCAGGCTGCGCGCAAAGACAGCGGC
AGCCCGAACCCCGCCCGCGGCCACAGCGTCTGCTCCACCTCCAGCTTGTACCTGCAGGAT
CTGAGCGCCGCCGCCTCAGAGTGCATCGACCCCTCGGTGGTCTTCCCCTACCCTCTCAAC
GACAGCAGCTCGCCCAAGTCCTGCGCCTCGCAAGACTCCAGCGCCTTCTCTCCGTCCTCG
GATTCTCTGCTCTCCTCGACGGAGTCCTCCCCGCAGGGCAGCCCCGAGCCCCTGGTGCTC
CATGAGGAGACACCGCCCACCACCAGCAGCGACTCTGAGGAGGAACAAGAAGATGAGGAA
GAAATCGATGTTGTTTCTGTGGAAAAGAGGCAGGCTCCTGGCAAAAGGTCAGAGTCTGGA
TCACCTTCTGCTGGAGGCCACAGCAAACCTCCTCACAGCCCACTGGTCCTCAAGAGGTGC
CACGTCTCCACACATCAGCACAACTACGCAGCGCCTCCCTCCACTCGGAAGGACTATCCT
GCTGCCAAGAGGGTCAAGTTGGACAGTGTCAGAGTCCTGAGACAGATCAGCAACAACCGA
AAATGCACCAGCCCCAGGTCCTCGGACACCGAGGAGAATGTCAAGAGGCGAACACACAAC
GTCTTGGAGCGCCAGAGGAGGAACGAGCTAAAACGGAGCTTTTTTGCCCTGCGTGACCAG
ATCCCGGAGTTGGAAAACAATGAAAAGGCCCCCAAGGTAGTTATCCTTAAAAAAGCCACA
GCATACATCCTGTCCGTCCAAGCAGAGGAGCAAAAGCTCATTTCTGAAGAGGACTTGTTG
CGGAAACGACGAGAACAGTTGAAACACAAACTTGAACAGCTACGGAACTCTTGTGCGTAA

Figure 1: Each of the four transcription factors were found using NCBI blast, and included an ATG start codon.

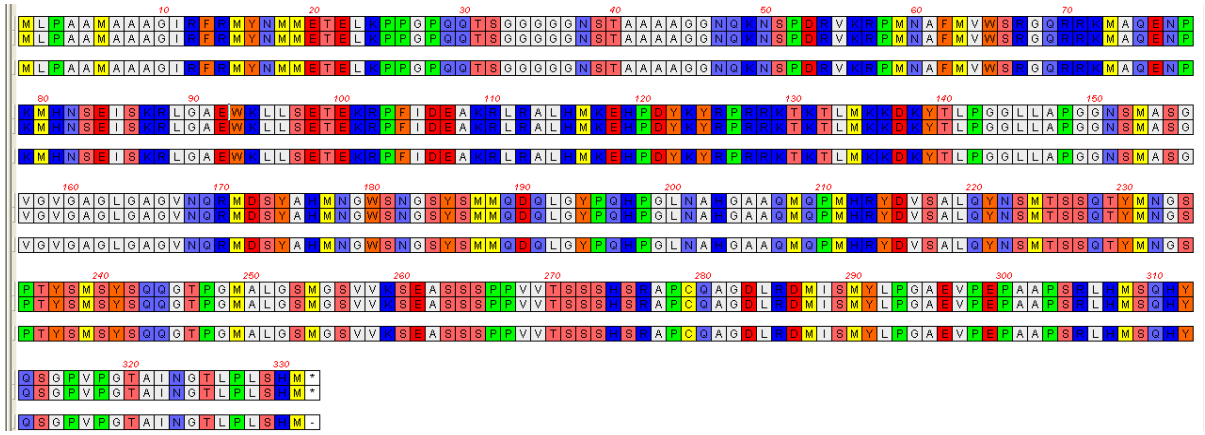


Figure 2: Sox2 Amino Acid Comparison

Figure 2 illustrates the translation of the NCBI database Sox2 sequence versus the given STEMGENT Sox2 amino acid sequence. The first row of each pair denotes the NCBI sequence translated to the amino acid sequence while the second row denotes the STEMGENT amino acid sequence. Refer to the Appendix Figure 1-3 for the other three transcription factor amino acid comparisons.

All four of transcription factors' DNA sequences were translated to their amino acid sequences and compared to the known STEMGENT amino acid sequences The next step was to subclone these DNA sequences into a *P. pastoris* protein expression plasmid.

2- Modification/Design of each of the four transcription factors:

The first modification that was done was to remove the ATG codon from each of the sequences. This ensured that there was not a double ATG codon within the plasmid once the genes were inserted into the *P. pastoris* plasmid. ATG is the start codon for initiation of translation. Another step was to remove a stop codon so that a his-tag can be added to the proteins for facilitating the protein purification.

3-Selection of Plasmid:

In the lab, different *P. pastoris* plasmids were available for cloning. There are two main plasmids used for gene cloning in the lab. These two plasmids- purchased from Invitrogen (Carlsbad, CA) are pPicZ-A and pPic9. The pPicZ-A plasmid allows for intracellular production of the desired protein/ proteins from *P. pastoris*, while pPic9 secretes the gene of interest/ protein by utilizing the α- factor secretion signal of *P. pastoris.* Both contain a multiple cloning site (MCS) that allowed for the desired sequence/ insert to be inserted into the plasmid MCS site and produced by *P. pastoris*. Another difference between pPicZ-A and pPic9 is that pPicZ-A contains a C-terminal 6X His tag which can be used for protein purification purposes. The pPicZ-A plasmid was chosen as the *P. pastoris* expression plasmid because it contained a 6X His tag that allowed for protein purification.

The pPicZ-A plasmid contains an ATG site at the AOX-1 site, which is the start of translation. The gene of interest, which in this case was one of the four transcription factors, was

designed to be inserted into the MCS of the plasmid. The MCS contains several different restriction enzyme sites for subcloning. Figure 3 depicts the pPicZ-A plasmid map for subcloning Sox2 gene.

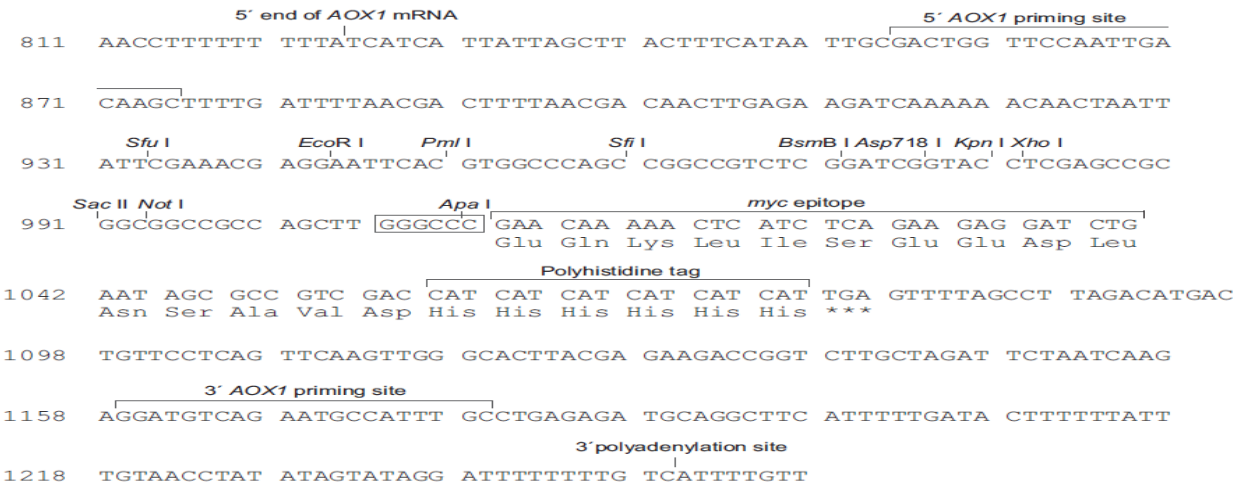


Figure 3: pPicZ-A Plasmid Map of MCS site

4- Selection of Enzymes

After the plasmid was chosen, the next step was to select restriction enzyme sites for subcloning. EcoRI and NotI were chosen because Sox 2 gene does contain these two restriction enzyme sites. The NotI site is 5' GCGGCCGC 3' and the EcoRI site is 5' GAATTC 3'. The enzyme digestion of these two sites generates two sticky ends, which is optimal for ligation. Once the restriction enzyme sites were chosen, they were checked to ensure that it was possible to run a double digestion by using the New England BioLabs double digest finder. This double digest finder allows for two enzymes to be chosen for a double digestion. For EcoRI and NotI HF, the recommended buffer was NEBuffer 4 with BSA at 37°C. An EcoRI site was inserted on the N-terminal side of each transcription factor sequence and a NotI site inserted on each C-

terminal end. Now that the enzyme sites were added to each transcription factor, the next step was to add the Arginine tag and linker sequence to each sequence.

5- Addition of Arginine tag and linker

Once the sequence was designed, an Arginine tag and linker was added to each transcription factor sequence on the 3' (C-terminal) end. This addition allowed for the produced protein to penetrate the cell membrane of target cells and enter the nucleus, where reprogramming of the cell into an iPS cell can begin. Without the addition of the poly-Arginine (11R) tag, the protein would not be able to enter the cell. This Arginine tag acts as a CPP, similar to Ayman and Melikov's research work (7,8). The Arginine tag has a large positive charge, and is able to penetrate the cell membrane of cells using a translocation mechanism. The Arginine tag must have a linking sequence to allow for the tag to work properly. Below is the linker and Arginine tag that was found using the given sequence from STEMGENT:

Arginine Tag linker sequence: GAGAGCGGCGGCGGTGGTTCTCCGGGG

Arginine Tag: CGCCGTCGCCGCCGTCGCCGTCGTCGCCGTCGC

Now that all of the additions had been made to each of the four transcription factors, the DSgene software was used again to run a test trial of enzyme digestion on my sequence and *P. pastoris* plasmid. This simulation allowed for the digested transcription factors to be inserted into the digested plasmid in the MCS site. Once this was done, each plasmid was checked to ensure that no frame shift had occurred. A frame shift causes possible gene mutations to occur, so if any frame shift occurs, such as a codon shift or deletion, this must be accounted for by adding or removing extra base pairs. The software was then used again to translate the sequence into the amino acid sequence. If the sequence was the same as before, no changes needed to be

made to the sequence designs. Figure 4 below shows the Sox2 sequence with the EcoRI and NotI site added, while figure 5 shows the plasmid pPicZ-A with the Sox2 sequence inserted.

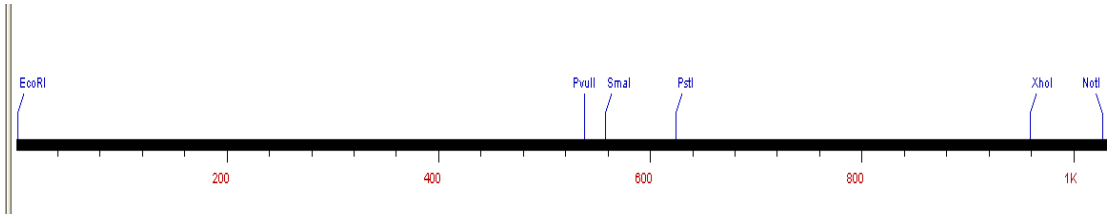


Figure 4: Sox2 sequence with enzyme sites added

EcoRI was added to the N-terminal end of the Sox2 sequence, while NotI was added to the C-terminal end of the Sox2 sequence.

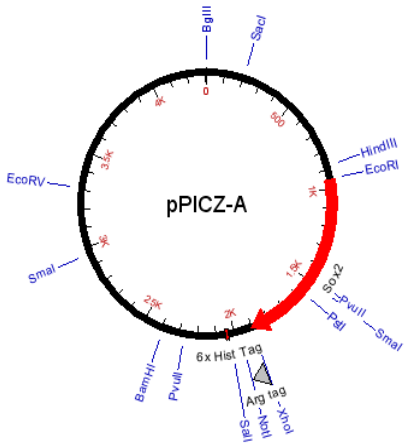


Figure 5: pPicZ-A plasmid with Sox2 inserted

The above figure shows the Sox2 designed sequence correctly inserted within the MCS site of the pPicZ-A plasmid between the EcoRI site and the NotI site. Refer to the Appendix figures 4-6 to observe the three other transcription factors correctly inserted within the pPicZ-A plasmid.

6- Codon Optimization:

Since *P. pastoris* does not use the same codon as human does, the codons were optimized for effective protein expression. Different methods for optimization were researched, including constructing oligonucleotides and overlapping PCR. Codon optimization allowed for more efficient translation to occur, which in turn provides higher protein generation efficiency. Forward and reverse nucleotides were constructed using DSGene software for each of the four transcription factors by following the protocol that Marlatt et al. used to optimize their codon sequence (10). Marlatt et al. constructed forward and reverse oligonucleotide sequences of a known sequence to be able to optimize the sequence for *E-coli.* The same protocol was followed for each of the four known transcription factor sequences, but instead of optimizing for *E.-coli,* each sequence was optimized for *P. pastoris* expression.

In total, 46 oligonucleotides were constructed for the Sox2 transcription factor. The other three transcription factors were all different base pair lengths, so it took a different number of oligonucleotides for each transcription factor. Each transcription factor had forward and reverse compliment oligonucleotides. Marlatt's protocol used overlapping reverse compliment oligonucleotides that would allow for annealing and ligation to occur (10). To be able to create optimal oligonucleotides for each transcription factor, I followed a basic protocol, which can be found within the Appendix section 1. Figure 6 below shows the Sox2 oligonucleotide sequence that was created by utilizing the DSgene software and the codon optimization protocol.

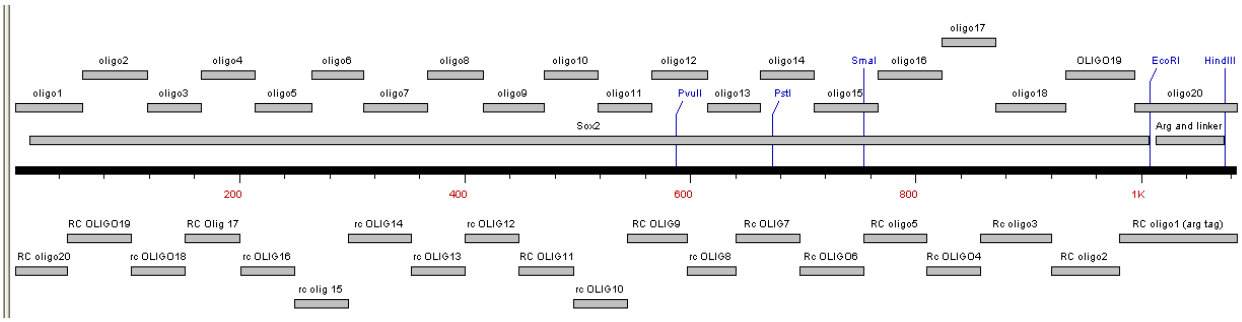


Figure 6: Sox2 Forward and Reverse Compliment Oligonucleotides

The forward Sox2 oligonucleotides are located on the top side of the dividing line, while the reverse compliment Sox2 oligonucleotides are on the bottom side of the line. The two oligonucleotide sequences have overlapping sequences. Refer to the Appendix figures 7-9 for the three other transcription factor oligonucleotide sequences that were designed.

After looking at the different alternatives and cost of ordering the oligonucleotides, it was better to allow the company Genscript (Piscataway, NJ) to optimize and create the designed Sox2 sequence than to custom order 46 oligonucleotides. By having Genscript create and optimize the sequence, it saved valuable time and money. After checking the sequence one final time, the Sox2 sequence was sent to Genscript and the company created and sent me the optimized completed sequence, which contained the desired inserted enzyme sites and c-terminal Arginine tag/linker.

Genscript used a cloning vector called pUC57 that allowed for sequences to be created and inserted in the MCS site of the vector. The pUC57 vector was delivered with the desired created sequence inside of the MCS site of the plasmid. Genscript took the assembled sequence, optimized the codons for *P. pastoris* expression, and inserted the Sox2 sequence inside the MCS site between the EcoRI site and the KpnI enzyme site. The final sequence length of the designed

Sox2 transcription factor was 1038 base pairs. Refer to the Appendix figure 10 for the final ordered optimized sequence of Sox2 from Genscript.

7- pPicZ-A Experiments:

The first experiment was to prepare LB broth and LB-agar (Appendix table 1 and 2). In order to grow pPicZ-A (Invitrogen, Carlsbad, CA), this was the type of broth and agar that was used. The goal of this experiment was to grow *E. coli* and have the *E. coli* pick up the pPicZ-A plasmid through the transformation process. The reason that this was done was to create more pPicZ-A plasmids and have it readily available for the subsequent cloning. Refer to the Appendix table 1 and 2 for the synthesis of Low Salt LB broth / agar protocol.

***8- E-coli/* pPicZ-A cloning, Glycerol Stock, and plasmid purification/elution:**

Refer to the Appendix section 2 for the complete protocol that was used for cloning the pPicZ-A plasmid. Once the pPicZ-A plasmid was successfully cloned, the plasmid DNA had to be purified/ eluted using a kit called Miniprep Promega Pure Yield Plasmid (micro centrifuge) (Madison, WI). Refer to Appendix section 3 for the complete protocol that was followed for plasmid purification.

9- Spectrophotometry Analysis:

Now that the plasmid DNA had been extracted, spectrophotometry was conducted to determine the concentration of the DNA plasmids. The software program that was used is called Gen5 from the company BioTek (Winooski, VT). The protocol that was followed for spectrophotometry consisted of 2 μl samples in each well plate, with one plate used as the control. The control used for each spectrophotometry (SynergyMx, BioTek, Winooski,

Vermont) run was distilled water. Refer to the results section of the paper for the spectrophotometry data obtained for the pPicZ-A DNA.

10- Enzyme digestion protocol

After DNAs were purified and their purity was validated by a spectrophotomer, the next step was to construct the protein expression plasmid. To do this, a protocol was used and some calculations were conducted that can be found in Appendix section 4.

11- Gel Electrophoresis

Once the enzyme digestion process had begun, 0.8% agarose gel was prepared (Appendix section 5). This gel is used for gel electrophoresis to check for the band lengths after the enzyme digestion had been completed. Refer to the Appendix section 5 to find the protocol used to prepare 0.8% agarose gel.

12- Gel Electrophoresis to check the size of the bands after enzyme digestion

Once the enzyme digestion is completed, the next step was to load each of the enzyme digested solutions into a separate lane. The protocol and images of the digested pPicZ-A plasmid are located in appendix section 6 indicates the steps that were taken when running gel electrophoresis. Once that the correct plasmid was verified, the next step in the process was to run experiments for the Sox2 plasmid DNA that was created by Genscript.

13- Dilution of plasmid SOX2 DNA

The first step that was taken when the Sox2 template DNA arrived was to dilute the plasmid. The Sox2 insert was delivered in a vector called PUC57. Instructions to dilute the plasmid came enclosed. Twenty μL of nuclease free water was added to the plasmid DNA,

which resulted in the concentration of the plasmid DNA to be 4——. A 1-1000 dilution followed by a 1-10 dilution were then performed, which resulted in a final concentration to 20—.

14- Transformation Protocol

Since the Sox2 template DNA was subcloned in the vector PUC57 (Genscript), LB agar plates supplemented with 100 μg/mL ampicillin (Gold Bio, Inc.) were prepared. The PUC57 vector contains an ampicillin resistance gene marker; an existing transformation protocol from Dr. Ye's lab was followed for each transformation that was run throughout the many experiments this protocol is found in the appendix section 7.

15- Glycerol Stock of Genscript Sox2 plasmid

Just as was done with the pPicZ-A plasmid, the same procedure described earlier (appendix section 2) was used to obtain a glycerol stock of the Sox2 plasmid. The glycerol stock ensures that there are reserves if something were to go wrong with future experiments.

16- Sox2 plasmid DNA extraction/ purification from *E.coli* Culture.

To purify and extract the Sox2 plasmid DNA, the earlier-described plasmid miniprep system protocol was followed (Appendix section 3) to isolate and purify the Sox2 plasmid DNA. Once the Sox2 plasmid DNA was purified, spectrophotometry was conducted to quantify the DNA concentration. The data that was derived from spectrophotometry-based procedure is found in the Discussion section of the paper. The next step of the experiment is to run a double digestion reaction of the Sox2 DNA and the pPicZ-A DNA, followed by gel extraction and ligation.

17- Enzyme Digestion of PPicZ-A and Sox2 using EcoRI/ NotI HF

The earlier protocol for enzyme digestion (appendix section 4) was followed and made three tubes from the three successful plasmid isolations of Sox2 plasmid DNA and one tube of pPicZ-A. One difference that occurred was that BSA (bovine serum albumin (New England BioLabs, Ipswich, MA) was added to the reaction because BSA was needed for the double digestion of EcoRI and NotI HF.

18- Qiagen MinElute Gel Extraction Kit using a Micro Centrifuge

Now that pPicZ-A and Sox2 have both been digested, the gel needed to be extracted and eluted/purified using the protocol from the Qiagen MinElute Gel Extraction Kit. The protocol for this process is described in the appendix section 8. Once the the DNA was purified and eluted, spectrophotometry was conducted to determine the concentration of the DNA.

19- Spectrophotometry of Sox2 and pPicZ-A after enzyme digestion

To quantify the concentration of DNA after enzyme digestion, spectrophotometry was performed. Even though a low DNA concentration was obtained through this experiment, the experiment was continued to see if any colonies could be grown after transformation. As described in the following section, transformation was never successful throughout my entire experiment. *E. coli* colonies were never able to grow on the agar plates after performing ligation/transformation. Another enzyme digestion was conducted later, and better spectrophotometry results were obtained but transformation failed as well.

20- Ligation protocol using Roche rapid DNA ligation kit

Ligation was conducted after each enzyme digestion of the backbone and Sox2 plasmid DNA. As described earlier, ligation allows for the sticky ends of the digested products to anneal and stick together, forming one single final plasmid. The final product after ligation should be

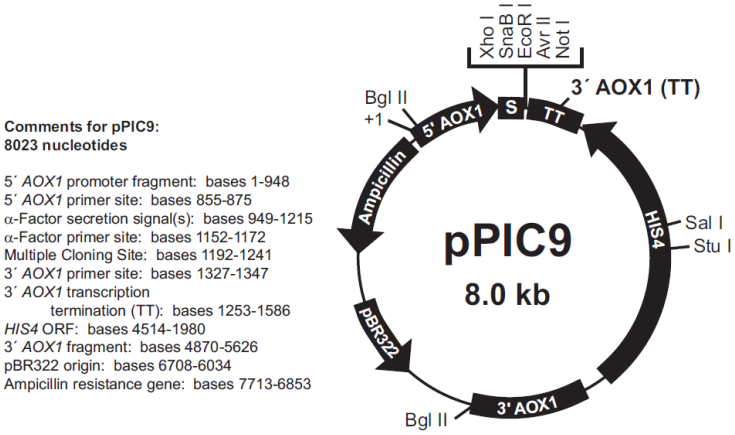
the pPicZ-A plasmid with the Sox2 insert located within the MCS site of the plasmid. The Rapid DNA Ligation Kit used for ligation was obtained from Roche (Indianapolis, IN). For the ligation mixture, the common procedure is to add a 3:1 or 2:1 insert to backbone ratio for ligation to occur most optimally. The final volume of the ligation reaction should be 21 µL. Refer to table 4 in the appendix to observe the reaction mixture used for the pPicZ-A/ Sox2 ligation.

After the ligation was completed, the next step that was to transform the newly digested Sox2 into the pPicZ-A DNA that was digested using the same restriction enzymes. The protocol described earlier for transformation (Appendix section 7) was used, but no colonies grew in any of the agar plates. These agar plates were Zeocin+ because pPicZ-A has a Zeocin resistance. Since transformation failed, a new strategy was developed for my research project.

21- Plasmid / Experimental Design Strategy Change

pPic9 protein expression vector was also used in this work. The plasmid pPic9 allowed for the secretion of proteins of interest and a better chance of glycosylation/ phosphorylation, which was a big part of the reason that *P. pastoris* was initially chosen as the expression vector. Now that the cloning strategy has changed, primers must be designed to allow for the Sox2 sequence to be inserted inside of pPic9. Since pPic9 has a different MCS site than pPicZ-A, the DSgene program was used for design of the new Sox2 sequence. Another problem that was addressed was that pPic9 did not contain a 6X Histidine tag like pPicZ-A, which is needed for purification of the protein. To accommodate these changes, primers were designed that contained a c-terminal Histidine tag. PCR allowed for the new Sox2 sequence to be created and then inserted into the pPic9 plasmid. The last step that was taken was to ensure that the newly designed sequence did not cause a reading frame shift and that translation occurred correctly starting at codon 949 in the MCS site of pPic9. Figure 7 below shows the plasmid map of pPic9

from Invitrogen that was used to understand the MCS site within pPic9, as well as the antibiotic resistance of pPic9 (Ampicillin).



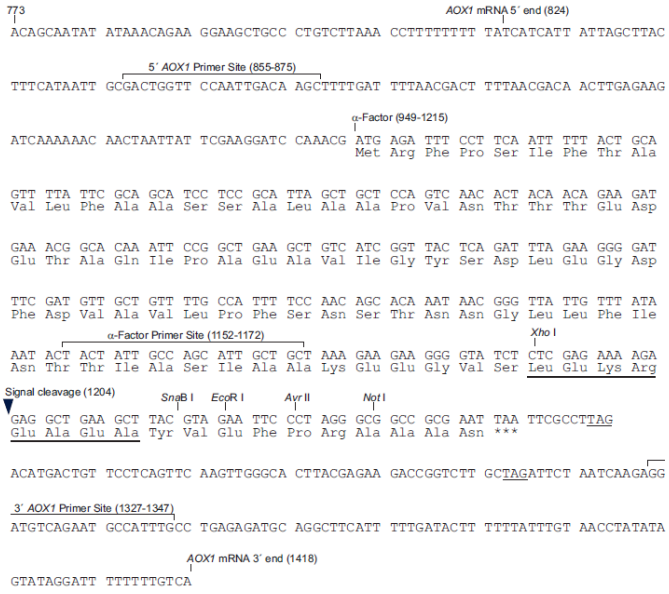


Figure 7: Plasmid map of pPic9 and MSC site

The MCS site in the above figure is located between the XhoI site and the NotI site. Therefore, the two restriction enzyme sites EcoRI and NotI need to be added to the primer sets, since the same enzymes were used for the template design of Sox2.

For the initial primer design, a graduate student in the lab, Lingyun Zhou, suggested a change of the enzyme site from NotI to AvrII because NotI had low cleavage efficiency. For this strategy to work, a forward and reverse compliment primer was designed to incorporate the new reading frame. The template of Sox2 can be found in the appendix figure 10 and the initial primer design for PCR is shown below in figure 8. Primers were ordered from a company called Eurofins MWG Operon, and they shipped the designed primers to the lab.

Forward primer with EcoRI site from Eurofins MWG Operon

5'-GCGCGAATTCTATAATATGATGGAGACCGAACTTAAACCACCAGG- 3'

Additional base pairs (hanger) were added to the beginning of the sequence to aid in binding of the sequence.

Reverse Compliment primer with AVRII site from Eurofins MWG Operon

5'- CAGTCACCTAGGATGATGATGATGATGATGTCTTCTTCTTCTTCTTCT -3'

Figure 8: Primer design #1

A six base pair hanger site was added before the NotI site to aid in sequence binding. CAGTCA is a common sequence added before NotI. Figure 9 below shows the pPic9 plasmid with the Sox2 PCR amplified sequence inserted successfully into the plasmid.

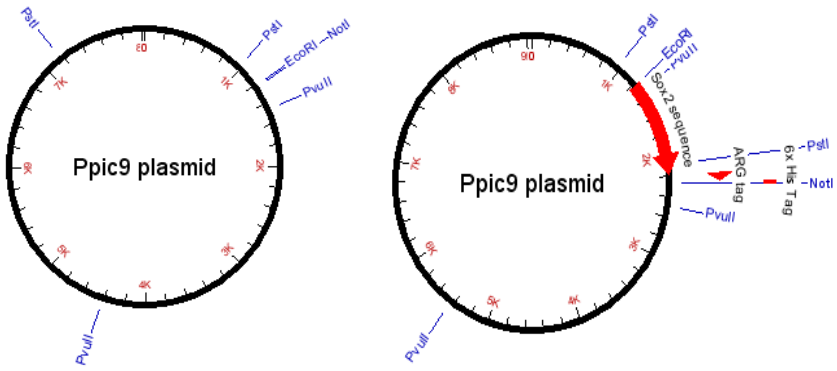


Figure 9: pPic9 plasmid and pPic9 plasmid with Sox2 PCR product inserted

22- Make LB Agar and LB Broth

The first step taken in the pPic9 plasmid design process was to prepare LB broth and LB agar plates for use in transformation and overnight cell cultures. The two Appendix tables 5 and 6 describe the process used to create LB broth and LB agar.

Agar plates were then prepared, and pPic9 colonies were observed by following the same protocol that was used for pPicZ-A. The same protocol that was used for the pPicZ-A plasmid purification (Appendix section 2,3) was carried out, such as creating a pPic9 glycerol stock, purification/elution of the plasmid DNA, and spectrophotometry to quantify the pPic9 DNA concentration obtained from the PureYield plasmid purification protocol. Now that the pPic9 plasmid DNA had been produced and quantified for later use, the next step was to run PCR using the primer that was ordered from Eurofins MWG Operon.

23- PCR protocol

For the first primer that was ordered, I followed the protocol and added the required amount of water to dilute each primer to 100uM. After diluting to 100— , I conducted a 1/100

dilution to get a final concentration for each primer of 10 —. After diluting the primers, the template DNA from Genscript was also diluted to a final concentration of 10 — Now that the correct concentrations of primer and template DNA was achieved, the PCR reaction mixture was created and is referenced in the Appendix section 9.

Once the PCR reaction mixture was prepared, the mixture was placed inside a PCR machine and a specific cycle was inputted to amplify the desired Sox2 plasmid DNA sequence from the DNA temperates. The PCR cycling instructions are located in the Appendix section 9, table 8.

After PCR was completed, the Sox2 PCR product was checked using gel electrophoresis protocol described earlier in appendix section 6. However, only 5 µL of the Sox2 product was used to confirm that the Sox2 sequence was generated from PCR. Once it was confirmed, the next step was the QiAquick PCR purification kit protocol (Appendix section 9) followed by the enzyme double digestion protocol that was described earlier in Appendix section 4, this time EcoRI and AVRII were used as the two enzymes. Once enzyme digestion was completed, the same protocol was used to perform gel extraction, ligation, and transformation as found in the Appendix. However, after three tries using this primer design, I was never able to grow any colonies that contained the ligated pPic9/ Sox2 plasmid. Different steps included changing the ligation ratio and calculating the molecular weight of the insert versus the backbone, but nothing produced any colonies. The next step that was taken was to design a new primer and redo all of the above experimental steps/ protocol.

Primer Design #2 Reverse Compliment with NotI site from Eurofins MWG Operon

5’ ATAAGAATGCGGCCGCATGATGATGATGATGATGTCTTCTTCTTCTTCTTCT -3’

Figure 10: Primer Design #2

The above primer was designed as the reverse compliment to the earlier forward primer, the only changes that were made was to change the enzyme site back to NotI from AvrII and to add a base pair sequence in front of the NotI site (ATAAGAAT). The reason that this was done was because it was decided that NotI was located in the original template so something may have gone wrong with the PCR reaction when using AvrII. The same forward primer was used from earlier and the newly designed reverse compliment primer was ordered from Eurofins MWG Operon. When the primer arrived, it was diluted to 100uM by adding 433 uL. From here, the same dilution process was used as described earlier. The same experiments were repeated from the last primer design, such as PCR, gel electrophoresis, PCR purification, spectrophotometry, enzyme digestion, gel extraction, ligation, and transformation. This process once again did not result in any colonies forming after repeating the above steps from PCR to transformation multiple times, so two new primer sets were designed for another trial.

The two primers were designed using the same process as before, but this time the length of the primers were reduced and Dr. Ye helped in the design of the primers. The final primer designs are shown below. The Tm’s for each primer were calculated to be 49°C, and this made the PCR cycling process optimal since both of the primer’s Tm were the same.

Forward Sox2 Primer #3

5’ GCGCGCGAATTCATGTATAATATGATGGAG 3’

Reverse Compliment Sox2 Primer #3

5’ATCATCGCGGCCGCAATGTGATGGTGATGGTGATGTCTTCTTCTTCTTCTTCTTCT’3

Figure 11: The designed primers were ordered from Eurofins MWG Operon and the same experiments were performed with a few modifications. For the PCR cycling reaction, the annealing temperature was changed from 60°C to 49°C and the extension time was changed to 20 seconds. All of the experiments through transformation were conducted using the same protocol, with the exception of ligation, in which a new kit was used. This kit was called Fermentas Rapid DNA Ligation Kit. The protocol used for this kit is found in Appendix section 10.

Discussion

Although the results obtained throughout this research experiment were not as expected, the enzyme digestion and spectrophotometry results show band lengths and data are positive. The first spectrophotometry results are for the pPicZ-A DNA and digested DNA concentrations, shown below in table 1.

Table 1: Spectrophotometry of pPicZ-A plasmid DNA

	Optical Density			
Reaction	**260**	**280**	**260/280**	**Concentration (ng/µL)**
1-pPicZ-A	0.85	0.51	1.861	233.295
2-pPicZ-A	0.76	0.45	1.865	192.173

The optimal 260/280 ratio for digested DNA products is 1.8-2.0. The ratio for both of the isolated colonies was in the optimal range, so the pPicZ-A plasmid was amplified successfully. Table 2 shows the spectrophotometry results obtained from the Sox2 plasmid DNA.

Table 2: Sox2 Plasmid DNA Spectrophotometry

	Optical Density			
Reaction	**260**	**280**	**260/280**	**Concentration (—)**
1-Sox2	0.417	0.222	1.876	416.59
2-Sox2	.486	.261	1.860	486.332
3-Sox2	.465	.248	1.875	465.366
4-Sox2	.013	.009	1.419	12.5

The above spectrophotometry Sox2 recipes were created using the transformation protocol from section 14 in the methods section of the paper. Clearly, the fourth Sox2 reaction did not digest correctly, since very little DNA was obtained and the ratio was poor. The other three reaction tubes showed positive results, as the 260/280 ratio was between 1.8- 2.0 and a high DNA concentration was achieved. The reason that the fourth reaction was so different may be due to the type of spreading that was done for the agar plate. For the fourth plate, a different technique was used called smearing, which may have caused problems with the *E.coli*. Figure 12 below shows the gel electrophoresis image that was seen under fluorescence after enzyme digestion.

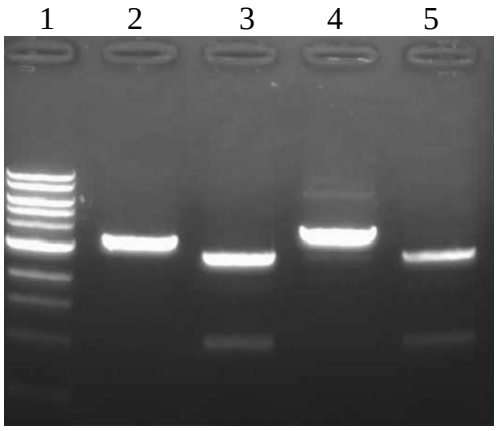


Figure 12: Enzyme Digestion of pPicZ-A and Sox2 using EcoRI and NotI HF

This was the first enzyme digestion of the pPicZ-A plasmid and Sox2. Since pPicZ-A is a 3.3kb long plasmid, the sequence was correctly digested because only a small portion of pPicZ-A was cut from the plasmid. Using EcoRI and NotI HF, the plasmids were cut, and resulted in the needed band lengths. The 1kb DNA ladder in lane 1 is used as a standard for ascertaining the band lengths. Lane 2 is the pPicZ-A digested plasmid that is about 3.3 kb, while lanes 3,4, and 5 are the different Sox2 digested plasmids. Transformation would later fail, so pPic9 was chosen as a new plasmid. Once the enzyme digestion reaction was complete and the gel was extracted/ purified, the DNA concentration for both the pPicZ-A and Sox2 DNA was quantified. Table 3 shows the results obtained from this enzyme digestion spectrophotometry.

Table 3: Sox2/ pPicZ-A enzyme digested (NotI/EcoRI) spectrophotometry

	Optical Density			
Reaction	**260**	**280**	**260/280**	**Concentration (—)**
Sox2	0.009	0.0	3.95	9.101
pPicZ-A	0.011	0.005	2.233	11.215

Based on the optical density data above, the enzyme digestion of Sox2 did not work correctly, since the 260/280 ratio was almost 4.0 and needed to be closer to 2.0. However, after another experiment, similar data was generated and it was concluded that something was wrong with enzyme digestion of the Sox2 sequence. This is the reason that PCR was used to try to produce the optimal sequence that could be used for ligation/ transformation. Table 4 documents the spectrophotometry results obtained for the four pPic9 isolated colonies.

Table 4: pPic9 plasmid DNA spectrophotometry

	Optical Density			
Reaction	**260**	**280**	**260/280**	**Concentration (—)**
pPic9 #1	0.114	0.061	1.873	113.94
pPic9 #2	0.13	0.068	1.909	130.023
pPic9 #3	0.101	0.052	1.957	101.3
pPic9 #4	0.105	0.054	1.93	104.762

These results obtained from the pPic9 colonies indicate that the process went correctly and pPic9 plasmid DNA was amplified using *E.coli.* Now that the pPic9 plasmid DNA was quantified, the next step was to run PCR on the Sox2 template using the first primer and to then run gel electrophoresis to check the band length of the sequence.

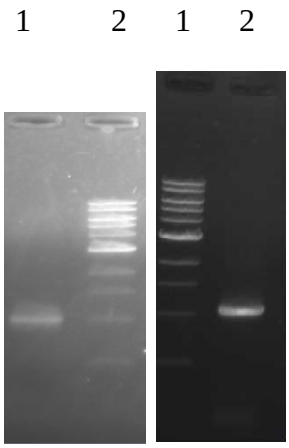


Figure 13: PCR product confirmation from two separate PCR reactions

For the image on the left, lane 2 indicates the 1kb DNA ladder, while lane 1 shows the Sox2 band length, which was about 1050 base pairs. The Sox2 sequence was confirmed since it was the correct length according to gel electrophoresis. The above image on the right is another Sox2 PCR product confirmation, but in this image lane 2 indicates the Sox2 product. Once the PCR product was confirmed, spectrophotometry of the PCR product was quantified below in table 5.

Table 5: Sox2 PCR

	Optical Density			
Reaction	**260**	**280**	**260/280**	**Concentration (—)**
Sox2	0.037	0.019	1.893	36.636
Sox2	0.035	0.019	1.893	35.164

The above results were positive since the 260/280 ratio is optimal and a high concentration was generated. Since it seemed that PCR was successful, the next step was to run

a double digestion on pPic9 and Sox2. Figure 14 below shows the gel electrophoresis image from enzyme digestion using NotI HF and EcoRI as the two enzymes.

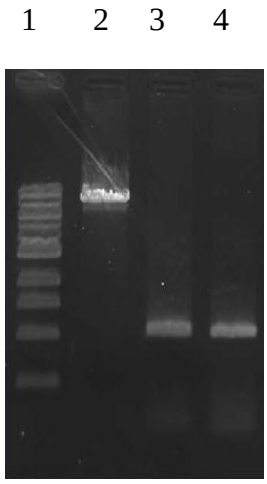


Figure 14: Enzyme digestion of ppic9 and Sox2 using NotI and EcoRI

Lane 1 was the 1kb DNA ladder, while lane 2 was the pPic9 plasmid DNA, which is about 8 kb in length. The Sox2 sequence was in both lanes 3 and 4. The band lengths were deemed to be correct based on the above gel image, and were then extracted, purified, and ligated. However, just like all of the other attempts, transformation into *E.coli* would later fail. Since transformation failed 3 times using this first primer design, a new primer that was designed earlier in the paper was attempted using the same above technique. Table 6 and 7 below shows the spectrophotometry results obtained after enzyme digestion and gel extraction/ purification from two different experiments using the same Sox2 PCR product.

Table 6: Spectrophotometry Results Sox2 Enzyme Digested using EcoRI- NotI HF

	Optical Density			
Reaction	**260**	**280**	**260/280**	**Concentration (—)**
pPic9	0.022	0.012	1.863	21.889
Sox2	0.044	0.022	1.948	43.692

Table 7: Spectrophotometry Results Sox2/ pPic9 Enzyme Digested using EcoRI- NotI HF

	Optical Density			
Reaction	**260**	**280**	**260/280**	**Concentration (—)**
Sox2	0.027	0.013	2.08	27.074
pPic9	0.019	0.006	3.019	18.692

Both table 6 and 7 above show results obtained from enzyme digestion of two separate Sox2 PCR products. Both of the transformations failed after ligation/transformation into *E.coli*, which was the reason a new primer was designed. After looking at the data, it is still not concluded why transformation did not work, since the concentration was inserted at a 3:1 molar ratio of insert to backbone, just as the ligation protocol asked for. However, when working with experiments that have never been done before, results are not always what one would expect them to be.

Below are a more images and data generated from all of the experiments. Many of the gel images and spectrophotometry data tables were left out of this report, due to so many experiments being run and since most of the images were the same. All of the gel electrophoresis images confirmed the correct DNA sequences, but for some reason

transformation failed every time. In total, three different constructed primer pairs were used to amplify the Sox2 sequence from the ordered template DNA. Even though the correct Sox2 sequence seemed to be amplified each time, transformation was never successful. The spectrophotometry data below also was statistically positive. In science, it is sometimes difficult to explain why something does not work the way that one would expect.

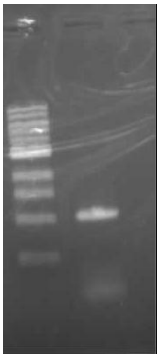


Figure 15: Sox2 PCR product confirmation

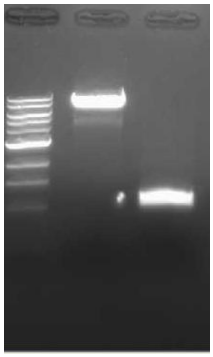


Figure 16: Sox2 Ppic9 enzyme digestion results

Table 8: Sox2 PCR Results

	Optical Density			
Reaction	**260**	**280**	**260/280**	**Concentration (—)**
Sox2	0.109	0.059	1.85	109.101

Table 9: Sox2 pPic9 Enzyme Digestion

	Optical Density			
Reaction	**260**	**280**	**260/280**	**Concentration (—)**
pPic9	0.028	0.011	2.449	27.65
Sox2	0.03	0.015	2.008	30.491

Table 10: Sox2 PCR Product

	Optical Density			
Reaction	**260**	**280**	**260/280**	**Concentration (—)**
Sox2	0.1	0.091	1.552	25.922

Table 11: Enzyme Digestion of Ppic9 -Sox2 PCR Product Spectrophotometry

	Optical Density			
Reaction	**260**	**280**	**260/280**	**Concentration (—)**
pPic9	0.021	0.01	2.217	21.198
Sox2	0.009	0.003	2.926	9.229

Accomplishments:

As documented throughout this paper, no colonies were ever formed that contained the Sox2 sequence inserted in the pPic9 plasmid. The same basic experiments were performed using the protocol described in the Appendix, but different primers were used, as well as different ligation ratios. Nothing was able to produce any different results. The main thing learned throughout this process was that research work is not as easy as it first seems. It requires much work and background understanding, as well as a lot of time spent in the lab. Even though it was frustrating not being able to obtain results, it was still a rewarding experience being able to work in the lab and gain hands- on experience working with plasmid design and PCR. One problem that may have occurred it that my original Sox2 template DNA may have been designed incorrectly, but it is hard to be able to know this without redesigning the sequence and trying to obtain new results.

The original plan was to create four different proteins from the four different transcription factor sequences, and not even one was able to be created. It goes to show that lab work, especially research work that has never been done before, takes a lot of time and things may not go according to plan. An important lesson learned through the research work was learning how to adjust on the fly and analyze what may be going wrong. Overall, many hours were spent in the lab and much work was done, but no significant results were achieved except for learning the steps of real world research. At the beginning of the project, I thought that research work in the lab was going to be a relatively easy process, but learned that research work is difficult and requires much experience and effort. I have a new respect for people that do research lab work for a career after my undergraduate research work.

Future Work

There is work that could be done on the project for the future. If this project could be started over, more time would be taken in designing the Sox2 and other transcription factor sequences to ensure that they could be inserted into the desired plasmid. The idea is still feasible to use *Pichia pastoris* to increase the protein generation efficiency. A problem that may have occurred is that the sequence was causing a frame shift/ mutation, which did not allow the *E.coli* to pick up the plasmid, so no colonies formed. Many steps are involved in DNA subcloning, and each step requires optimization, relying upon experience and skills that need to be accumulated through hard work and training. Even though the research work conducted never produced desired results, I learned how to use a lot of different kits and was able to gain valuable experience working in a research lab. This experience is something that will be valuable for my future endeavors and I am thankful to have been given the opportunity to work in the lab.

References

1) **Takahashi, K., and Yamanaka, S. (2006)** *Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors.* Cell. 126: 663-676.

2) **Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007)** *Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors*. Cell. 131: 1-12.

3) **Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., Slukvin, I.I., and Thomson, J.A. (2007)** *Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells*. Sciencexpress: 1-4.

4) **Yamanaka, S. (2009)** *A Fresh Look at iPS Cells*. Cell. 137: 13-17.

5) **Kaji, K., Norrby, K., Paca, A., Mileikovsky, M., Mohseni, P., and Woltjen, K. (2009)** *Virus free induction of pluripotency and subsequent excision of reprogramming factors*. Nature. 458: 771-775.

6) **Danwei Huangfu, Kenji Osafune, Rene´ Maehr, Wenjun Guo, Astrid Eijkelenboom1, Shuibing Chen, Whitney Muhlestein, and Douglas A Melton. (2008)** *Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2.* Nature Biotechnology. 26: 1269-1275.

7) **Ayman, E-S., Futaki, S., and Harashima, H. (2009)** *Delivery of Macromolecules Using Arginine- Rich Cell – Penetrating Peptides: Ways to Overcome Endosomal Entrapment.* AAPS Journal. 11: 13-22.

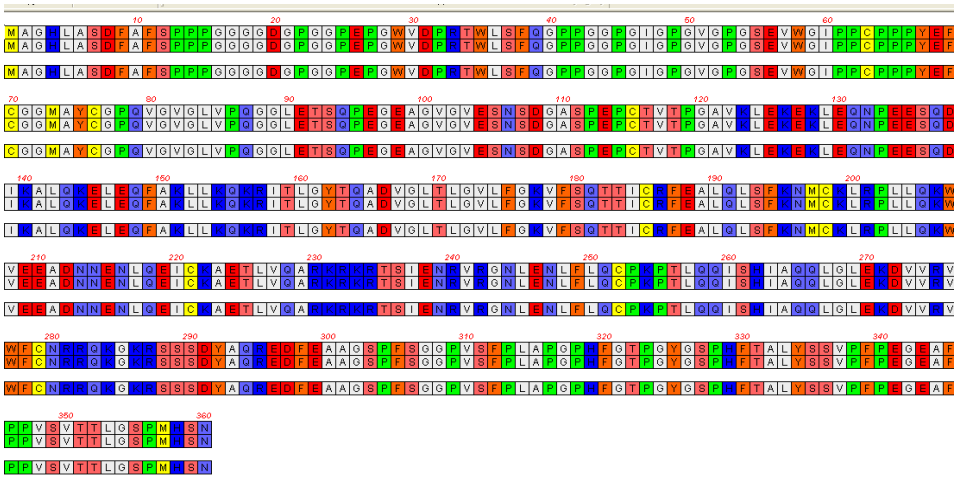
8) **Melikov, K. and Chernomordik, L.V. (2005)** *Arginine- rich cell penetrating peptides: from endosomal uptake to nuclear delivery*. Cell Molecular Life Science 62: 2739-2749.

9) **Zhou, H., Wu, S., Joo, J.Y., Zhu, S., Han, D.W., Lin, T., Trauger, S., Bien, G., Yao, S., Zhu, Y., Siuzdak, G., Scho¨ ler, H.R., Duan, L., and Ding, S. (2009)** *Generation of Induced Pluripotent Stem Cells Using Recombinant Proteins*. Cell Stem Cell. 4: 381-384.

10) **Marlatt, N.M., Spratt, D.E., Shaw, G.S.** *Codon optimization for enhanced Escherichia coli expression of human S100A11 and S100A1 proteins.* Protein Expression and Purification. 73: 58-64.

Appendix:

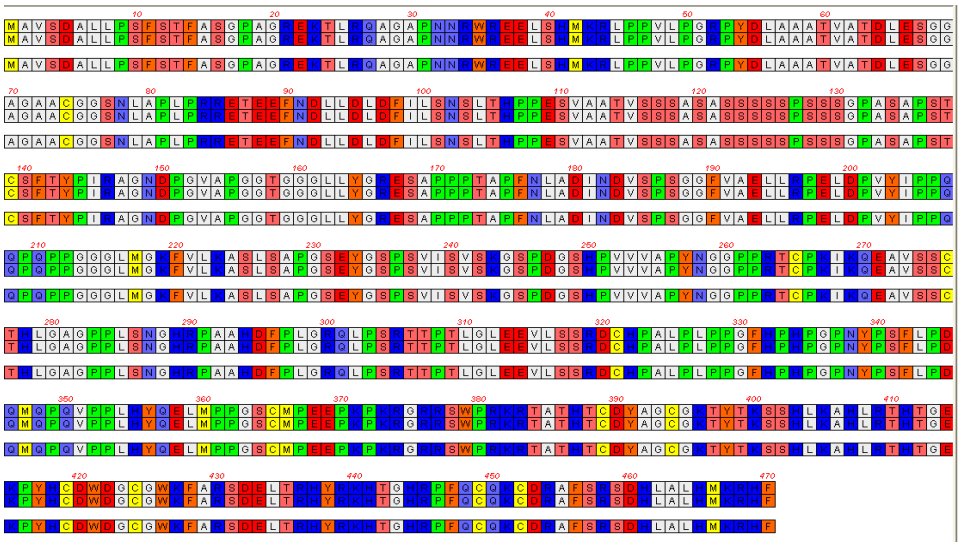
Figure 1: OCT4 Translation Comparison



Line 1: Oct4 NCBI translated to amino acid sequence

Line 2: Oct4 STEMGENT amino acid sequence

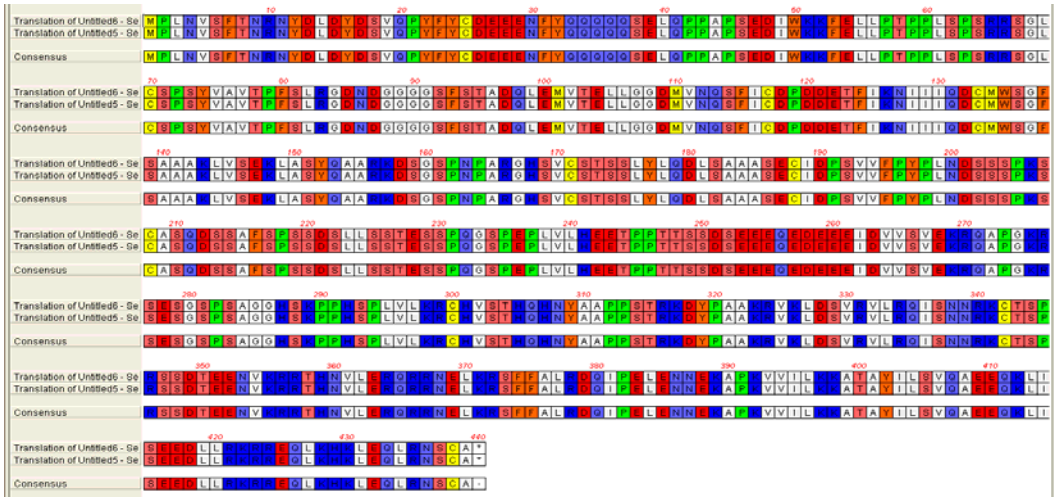
Figure 2: KLF4 Translation Comparison



Line 1: Klf4 NCBI translated to amino acid sequence

Line 2: Klf4 STEMGENT amino acid sequence

Figure 3: C-myc Translation Comparison



Line 1: C-mycf NCBI translated to amino acid sequence

Line 2: C-myc STEMGENT amino acid sequence

Figure 4:pPicZ-a with Klf4 inserted, C-myc, and Oct4 inserted

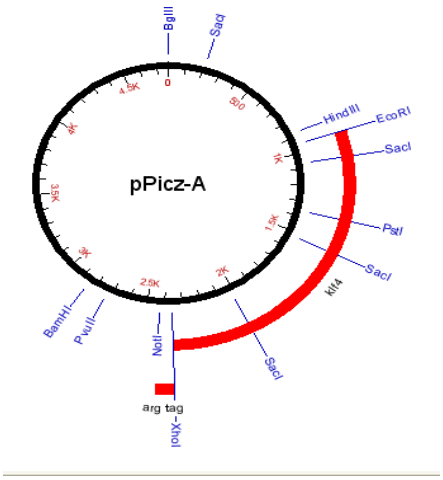


Figure 5: pPicZ-a with C-myc inserted

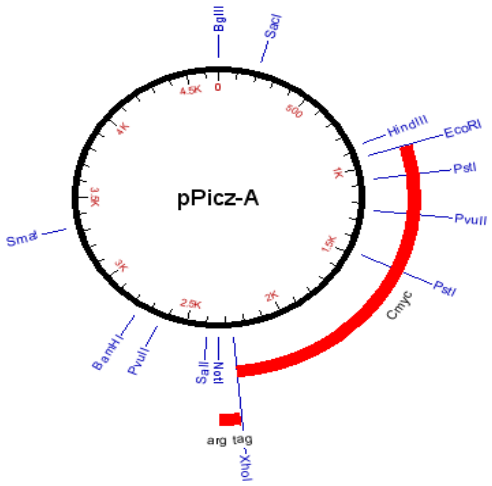


Figure 6: pPicZ-a with Oct4 inserted

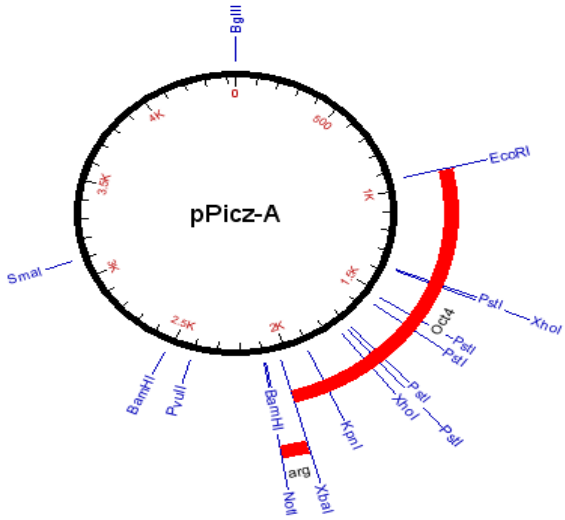


Figure 7: C-mycForward and Reverse Compliment Oligonucleotides

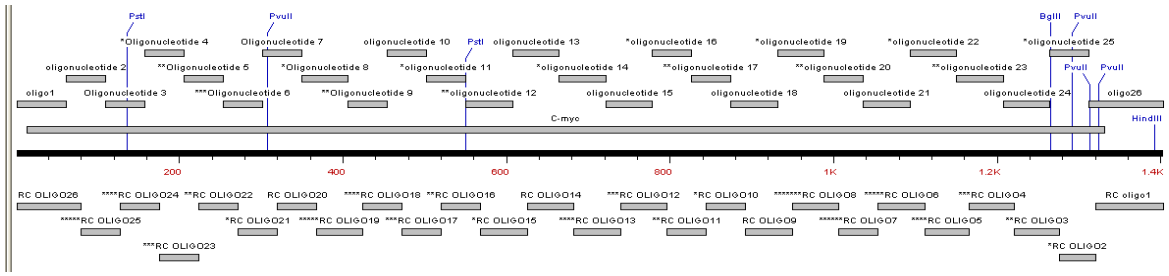


Figure 8: Oct4 Forward and Reverse Compliment Oligonucleotides

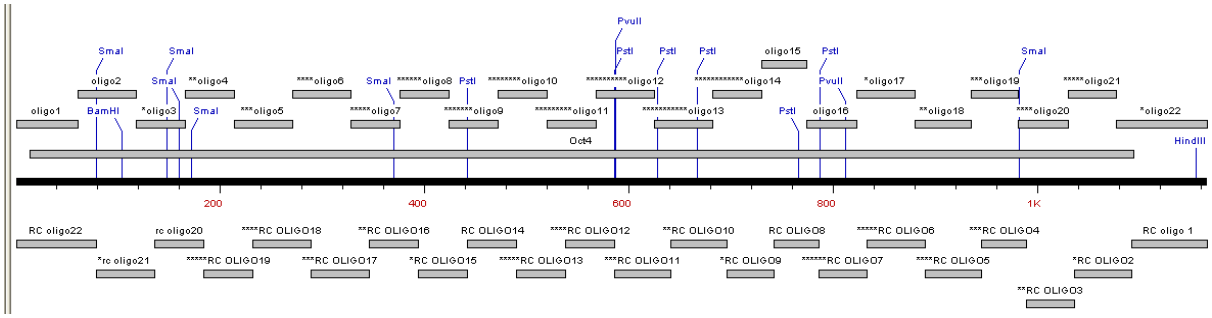
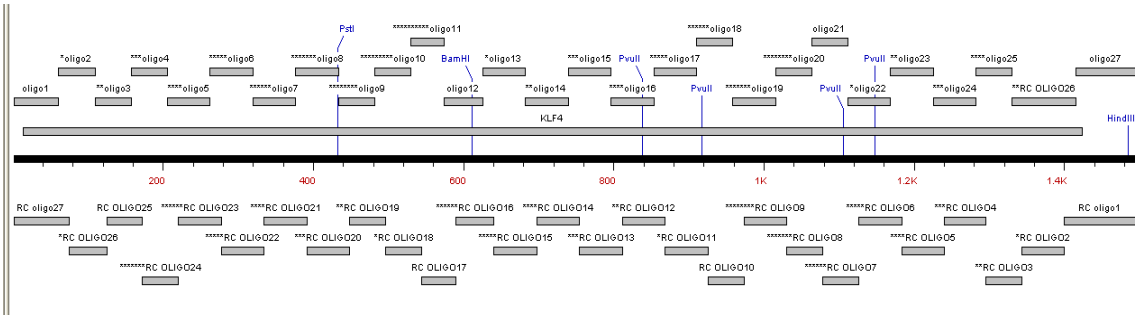


Figure 9: Klf4 Forward and Reverse Compliment Oligonucleotides



Section 1: Codon optimization protocol

- Using online website called Genius.de, insert known sequence into codon optimization software for *P.pastoris*
- Create oligonucleotides of about 60 nucleotides each in length (5'-3' direction).
- Label each oligonucleotide with a number
- Create reverse compliment oligonucleotides that have 15 base pairoverlap(start the first oligonucleotide shorter so that there is an overhang and overlap, allowing annealing and ligation to occur.
- Using ADDGENE software for the enzyme digestion, ensure that the sequence will not cause a frame shift once inserted into the plasmid.

Figure 10: Optimized Sox2 DNA sequence from Genscript

5'GAATTCATGTATAATATGATGGAGACCGAACTTAAACCACCAGGACCACAACAAACTTCAGGAGGA
GGAGGAGGAAATAGTACCGCAGCAGCAGCTGGTGGAAACCAAAAGAACTCTCCAGATAGAGTTAAAA
GACCTATGAACGCTTTTATGGTCTGGTCAAGAGGTCAAAGAAGAAAGATGGCTCAGGAAAACCCAAA
AATGCATAATTCCGAGATTTCAAAGAGATTGGGAGCTGAATGGAAATTGCTTTCTGAAACAGAGAAGA
GACCTTTCATCGATGAGGCAAAAAGATTGAGAGCTCTTCATATGAAGGAACACCCAGACTACAAGTAC
AGACCTAGAAGAAAGACTAAGACCTTGATGAAGAAAGACAAGTACACCCTTCCAGGTGGATTGCTTG
CCCCTGGTGGAAACAGTATGGCATCTGGTGTTGGAGTCGGTGCCGGATTGGGTGCAGGAGTTAACCAA
AGAATGGATTCTTACGCTCATATGAACGGTTGGTCCAATGGATCTTACTCCATGATGCAAGACCAGTT
GGGTTATCCACAGCATCCTGGTCTTAACGCCCACGGAGCTGCCCAAATGCAGCCAATGCACAGATACG
ACGTTTCAGCATTGCAATATAACAGTATGACATCTTCCCAGACTTACATGAATGGTTCTCCAACTTACT
CAATGAGTTATTCTCAACAGGGTACTCCTGGAATGGCTTTGGGTTCCATGGGATCAGTTGTCAAATCCG
AGGCCTCAAGTTCTCCACCTGTTGTCACTTCCTCAAGTCATTCAAGAGCTCCATGTCAAGCCGGAGATT
TGAGAGACATGATTAGTATGTACCTTCCAGGAGCTGAAGTTCCAGAGCCTGCAGCTCCTAGTAGATTG
CATATGTCTCAACACTATCAGTCCGGTCCAGTCCCTGGAACCGCCATCAATGGTACATTGCCTCTTTCT
CACATGCTCGAGGAGAGTGGAGGAGGAGGTTCACCAGGAAGAAGAAGAAGAAGAAGAAGAAGAAGA
AGAAGAGGCGGCCGCGGTACC3'

Table 1:Low Salt LB Broth Recipe- 300 ml

BactoTryptone (pancreatic digest of casein) Ref 211705	*10g X 0.3 =* ***3grams***
NaCl 99.85% DNASE RNASE protease free MW 58.44	*5g X 0.3 =* ***1.5 grams***
Yeast Extract REF 212750	*5g X 0.3=* ***1.5 grams***
Distilled H2O	*950ml X 0.3 =* ***300mL***

Table _____: Synthesis of Low Salt LB Agar plates with Zeocin

Tryptone	*10g X 0.3 =* ***3grams***
NaCl 99.85% DNASE RNASE protease free MW 58.44	*5g X 0.3 =* ***1.5 grams***
Yeast Extract REF 212750	*5g X 0.3=* ***1.5 grams***
Distilled H2O	*950ml X 0.3 =* ***300mL***

- If making agar, add 15g/L across organics Agar pure powder code 400402500 (after pH is adjusted to 7.5) of agar to the solution (for 300ml, add 15x.3= 4.5g) …
- Add Zeocin(25 ug/ml) to agar, pour into plates. The concentration of Zeocin (code 327300010) was 1 ug/ml, so I added 150 uLZeocin to the 300ml agar solution

Section 2:

Take *E-coli* stock with pPicZ-a, and inoculate streak onto Agar-Zeocin plate

- Place plate/plates into 37 degree Celsius incubation machine overnight (16.5 hours)
- Inoculated into 5ml Low Salt LB broth tubes the following morning
- Took agar plates out of 37 degrees Celsius and checked to see if colonies had grown.
- If colonies grown, pipetted 5ml of broth into tubes and added 2.5 uL of Zeocin to each tube.
- Pick out single colony from agar plate and circle bottom of plate, using a toothpick, pick up cell and place toothpick into broth tube
- Wrap tubes with tinfoil and place on shaker at 37 degrees Celsius, 250 rpm, overnight
- After overnight shake, stopped shaker machine to see if the color of the broth changed. If the color changed, experiment was successful and colonies grew.

- Placed broth tubes in 4 degrees Celsius until ready to do Miniprep Pure Yield Plasmid Protocol.

Glycerol Stock Creation of pPicZ-a

Now that the pPicZ-a plasmid has been successfully cloned using *E-coli,* a glycerol stock can be made to allow for long term storage of the plasmid/ *E-coli*. To do this, a glycerol stock solution was created using the following protocol.

Materials: 80% glycerol solution, *E-coli*with plasmid in broth(*E- coli* strain used was DH5α),

- 300 uL*E-coli*-broth solution and 100 uL of glycerol solution -4 total 5ml centrifuge tubes
- Add glycerol first, then *E-coli* containing plasmid.
- Mixed the tubes by pipetting up and down 3-5 times, Stored in-80 degrees Celsius

Section 3:Miniprep Pure Yield Plasmid protocol using a micro centrifuge:

Now that I had successfully created a glycerol stock, the rest of the broth containing the *E-coli* and pPicZ-a plasmid could be purified using the Miniprep Pure Yield Plasmid protocol shown below. I followed all of the following steps and was able to successfully obtain the genomic plasmid DNA of pPicZ-a.

The first step that must be done when using this kit is to prepare one Wizard® SV Minicolumn assembly for each lysate. Each minicolumn assembly consists of a Wizard® SV Minicolumn and a Collection Tube. Label the Collection Tube and place the Wizard® SV Minicolumn assembly in a microcentrifuge tube rack. I prepared two tubes, since I had two different pPicZ-a plasmids created. From here, I followed the protocol that is listed below from the kit manual.

- Transfer the entire sample lysate from the 1.5ml microcentrifuge tube to a Wizard® SV Minicolumn assembly.

- Place the Wizard® SV Minicolumn assembly containing the sample lysate into a microcentrifuge and spin at 13,000 × *g* for 3 minutes to bind the genomic DNA to the Wizard® Minicolumn. If some lysate remains on the column after the initial spin, spin again for 1 minute at 13,000 × *g*.
- Remove the Wizard® SV Minicolumn from the Minicolumn assembly and discard the liquid in the Collection Tube. Replace the Wizard® SV Minicolumn into the Collection Tube.
- Verify that ethanol has been added to the Wizard® SV Wash Solution as described in Section 3.A.
- Add 650µl of Wizard® SV Wash Solution to each Wizard® SV Minicolumn assembly.
- Centrifuge at 13,000 × *g* for 1 minute.
- Discard the liquid in the Collection Tube and replace the Wizard® SV Minicolumn into the empty Collection Tube.
- Repeat Steps 5–7 three times for a total of four washes of the Wizard® SV Minicolumn.
- After the last wash, empty the Collection Tube and reassemble the Wizard® SV Minicolumn assembly. Centrifuge at 13,000 × *g* for 2 minutes to dry the binding matrix.
- Remove the Wizard® SV Minicolumn and place in a new labeled 1.5ml micro-centrifuge tube for elution. Add 250µl of room temperature Nuclease-Free Water to the Wizard® SV Minicolumn. Incubate for 2 minutes at room temperature.
- Place the Wizard® SV Minicolumn/elution tube assembly into the centrifuge and spin at 13,000 × *g* for 1 minute.
- Remove the Wizard® SV Minicolumn/elution tube assembly from the centrifuge. Add another 250µl of Nuclease-Free Water to the Wizard® SV Minicolumn and incubate at

room temperature for 2 minutes. Place the Wizard® SV Minicolumn/elution tube assembly into the centrifuge and spin at 13,000 × *g* for 1 minute.

Total elution volume will be approximately 500µl.Remove the Wizard® SV Minicolumn and discard. Cap the elution tube containing the purified genomic DNA and store at –20°C.

Section 4: Plasmid Enzyme digestion protocol

- Identified two enzymes that each cut one time on the plasmid pPicZ-a
- The enzymes used were EcoRI and BamHI HF
- I checked to ensure that double digestion was possible for these two enzymes
- The result of digestion should be two bands with one band having a length 3 kb and another band with 500 base pair length due to the cuts created by these enzymes

Must know the final concentration of the plasmid DNA that is going to be used, as well as the amount of enzyme volume needed for the reaction to occur.

Sample Calculation of Enzyme Digestion reaction:

Based on the spectrophotometry results obtained, the amount of uL needed for enzyme digestion can be calculated. For enzyme digestion to work optimally there needs to be about 1ug of the plasmid present in the reaction. Below is a sample calculation of the process to calculate the needed amounts of each ingredient for enzyme digestion.

$230\frac{ng}{ul} * (\frac{1ug}{1000ng}) = 0.235$ ug/uL -Need about 4uL of pPicZ-a plasmid #1

$192n\frac{ng}{ul} * (1ug/1000ng) = 0.192$ –Need about 5uL of pPicZ-a plasmid #2

Now that I have calculated the amount of plasmid needed, the next step was to calculate the needed amount of enzymes for the enzyme digestion reaction. First, check the units for each

needed enzyme. EcoRI has 20,000$\frac{U}{mL}$ and BamHI HF has 20,000$\frac{U}{mL}$. The optimal amount of enzyme for digestion is 10U. Therefore, another simple calculation was done to determine the correct amount of enzymes to add for the reaction.

$$20{,}000\ \frac{U}{mL} * \left(\frac{1mL}{1000\ uL}\right) = 20\frac{U}{uL}$$

Since there is 20$\frac{U}{uL}$, I multiplied the concentration by 0.5 uL to get 10 U of enzyme. Therefore, only 0.5uL of both EcoRI and BamHI HF were needed for the double digestion reaction to occur optimally. I added2 uLof the NEBuffer4 solution that was given by the NEB double digestion finder. The NEBuffer4 solution is 10x concentrated, so this is the reason that 2uL of buffer solution was added to the reaction solution. The final solution volume needed to be 20uL, so a specific amount of autoclaved water was added to each reaction. All of the above reagents were pipette into a 1.5 mL centrifuge tube and then centrifuged for 10 seconds.

Table 2 and 3 below show the restriction digestion reactions for each of the two pPicZ-a solutions.

Table 2: Tube #1 Enzyme Digestion reaction with EcoRI and BamHI HF

Tube #1	**uL needed**
NE Buffer4	2
pPicZA #1	4
BamHI HF	0.5
EcoRI	0.5
Autoclaved H2O	13
Final Reaction Volume	20

Table 3: Tube #2 Enzyme Digestion reaction with EcoRI and BamHI HF

Tube #2	**uL needed**
NE Buffer4	2
pPicZA #2	5
BamHI HF	0.5
EcoRI	0.5
Autoclaved H2O	12
Final Reaction Volume	20

The enzymes should be kept on ice and added to the reaction mixture last. This allows for the enzymes to function correctly. Once all of the reagents are added to the solution, the tubes were placed in a 37 degree Celsius water bath for 1 hour, which allowed enzyme digestion to occur.

Section 5:Preparation ofAgarose 0.8% Gel

- Collect 1x TAE Buffer and agarose gel. (Products used: 50x TAE Buffer – FisherBiotech Product BP1332-4, Agarose – GibCoultra pure™ Agarose Catalog 15510-027).
- For 0.8% agarose gel, 0.8 grams of agarose (measured on Denver Instrument's XP-300 scale) for every 100 mL of sterile distilled water.
- Microwave mixture in a flask for 45-60 seconds, or until all agarose is dissolved.
- For ethidium bromide staining, add 5uL of ethidium bromide after microwaving.
- Place gel in a casting tray and insert well comb, which creates different lanes/holes.
- Allow gel to cool for no less than 45 minutes to ensure proper gelling.

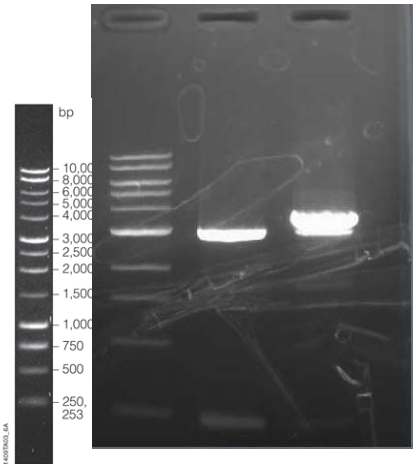
Section 6: Gel Electrophoresis protocol:

Gel electrophoresis was performed using the mini gel migration tank from National Labnet Co., Inc.

- Add proper amount of 1X TAE buffer to well of tank to cover the gel once placed in the tank. ~400 mL
- Place cooled gel into the tank.
- Add 4 uL of 6x loading dye into each 20uL sample, making final volume of samples 24uL
- Place 1kb DNA Ladder (7uL) into the left well and samples in the wells that follow to the right. (DNA ladder from New England BioLabs® Inc. Quickload® 1kb DNA Ladder N0468s).
- Run the gel electrophoresis at 100 Volts for 30 minutes.
- The gel was imaged in the Biorad Gel Documentation System. PDQUEST was used to view the gel.

Images of the gel are below in Figure 11

Figure 11: pPicZ-a digested with EcoRI and BamHI HF



Based on the above figure, the correct fragments were generated from the enzyme digestion. There is a 500 base pair fragment and a 3000 base pair fragment, meaning that the cloning of pPicZ-a was successful.

Section 7: Transformation protocol:

- Thaw a tube of NEB 5-alpha Competent *E. coli* cells on ice for 10 minutes.
- Add 2 uL of the plamid containing SOX2 (concentration of 20pg/ul) to 25 uL pre-thawed GC5 competent cells, mix by flicking the tube
- Incubate 30 min on ice
- Heat shock at 42°Cfor 30 seconds
- Place on ice for 5 minutes
- Add 500 uL room temperature SOC media
- Transfer the solution to a 5 mL culture tube ,incubate @ 37°C for 1 hr, 225 rpm shaker
- Spread 50-100uL of the cells onto warmed Agar plate (Ampicillin +) and incubate at 37°C overnight.

- Observe for colonies the following morning.
- If colonies formed, pick single colonies and inoculate into broth containing Ampicillin resistance and incubate individual colonies again overnight on shaker.
- For each chosen colony, use 5mL tube, add 5mL LB broth, 5uL Ampicillian (100ug/mL)
- Using a toothpick, pick single colony, drop toothpick into broth with Ampicillin+, cap the tube.
- Place the tubes into 4° C until ready for shaking at 250 rpm, 37°Covernight

Section 8: QiagenMinElute Gel Extraction Protocol

- Excise the DNA fragment from the agarose gel with a clean, sharp scalpel.
- Minimize the size of the gel slice by removing extra agarose.
- Weigh the gel slice in a colorless tube. Add 3 volumes of Buffer QG to 1 volume of gel (100 mg or approximately 100 µl). For example, add 300 µl of Buffer QG to each 100 mg of gel. For >2% agarose gels, add 6 volumes of Buffer QG. The maximum amount of gel slice per spin column is 400 mg; for gel slices >400 mg use more than one MinElute column.
- Incubate at 50°C for 10 min (or until the gel slice has completely dissolved). To help dissolve gel, mix by vortexing the tube every 2–3 min during the incubation.
- After the gel slice has dissolved completely, check that the color of the mixture is yellow (similar to Buffer QG without dissolved agarose).
- Add 1 gel volume of isopropanol to the sample and mix by inverting the tube several times. For example, if the agarose gel slice is 100 mg, add 100 µl isopropanol. Do not centrifuge the sample at this stage.

- Place a MinElute column in a provided 2 ml collection tube in a suitable rack.
- To bind DNA, apply the sample to the MinElute column, and centrifuge for 1 min. For maximum recovery, transfer all traces of sample to the column. The maximum volume of the column reservoir is 800 µl. For sample volumes of more than 800 µl, simply load and spin again.
- Discard the flow-through and place the MinElute column back in the same collection tube.
- Add 500 µl of Buffer QG to the spin column and centrifuge for 1 min.
- Discard the flow-through and place the MinElute column back in the same collection tube.
- To wash, add 750 µl of Buffer PE to the MinElute column and centrifuge for 1 min.
- Discard the flow-through and centrifuge the MinElute column for an additional 1 min at ≥10,000 x *g*.
- Place the MinElute column into a clean 1.5 ml microcentrifuge tube.
- To elute DNA, add 10 µl of Buffer EB (10 mMTris·Cl, pH 8.5) or water to the center of the membrane, let the column stand for 1 min, and then centrifuge for 1 min.
- Store DNA at –20°C as DNA may degrade in the absence of a buffering agent.

Table 4:pPicZ-a/ Sox2 Ligation Mixture

DNA dilution buffer	2uL
Backbone (pPicZ-a)	3ul (had 11ng/uL)
Insert (Sox2)	6 ul (had 9ng/uL)
MIX WELL	

Ligase Buffer 2x	10 uL
Ligase 5U/ul	1uL
Final Reaction Volume	21uL
Mix Well and incubate at RT for 20 mins	

Table 5: LB Broth Synthesis

Distilled H2O	***1 Liter***
LB broth	***25 grams***

Table 6: LB Agar Synthesis

Distilled H20	***300 mL***
LB agar	***12 grams***

Once the reagents were added, both the LB agar and broth were autoclaved and were then ready to be used. Ampicillin was prepared earlier and 300uL of Ampicillin was added to the agar once it had cooled to 55°C.

Section 9: PCR reaction

Table 7: PCR Reaction Mixture:

Reagent/ concentration	Volume added
H20	32.5uL
5x HF Buffer	10uL
dNTP (10mM)	1uL
Primer Forward (10uM/uL)	2.5 uL
Primer Reverse (10uM/uL)	2.5 uL
Template DNA (10ng/uL)	1uL
Phusion Polymerase	0.5 uL

Final Reaction Volume	50 uL

Table 8: PCR Cycling instructions/ Steps

Process	**Temp**	**Time**	**Cycles**
Initial Denaturation	**98 degrees C**	**30 seconds**	**1 cycle**
Denaturation	**98**	**5-10 seconds**	**30 seconds**
Annealing	**60**	**20 seconds**	**30**
Extension	**72**	**30 seconds**	**30**
Final Extension	**72**	**7 min**	**1**
Hold	**4 degree C**	**Hold**	**1**

QIAquick PCR Purification Kit Protocolusing a micro centrifuge

This protocol is designed to purify single- or double-stranded DNA fragments from PCR and other enzymatic reactions (see page 8). For cleanup of other enzymatic reactions, follow the protocol as described for PCR samples or use the new MinElute Reaction Cleanup Kit. Fragments ranging from 100 bp to 10 kb are purified from primers, nucleotides, polymerases, and salts using QIAquick spin columns in a microcentrifuge.

• All centrifuge steps are at ≥10,000 x *g* (~13,000 rpm) in a conventional tabletopmicrocentrifuge.

1. Add 5 volumes of Buffer PB to 1 volume of the PCR sample and mix. It is not necessary to remove mineral oil or kerosene.

For example, add 500 µl of Buffer PB to 100 µl PCR sample (not including oil).

2. Place a QIAquick spin column in a provided 2 ml collection tube.

3. To bind DNA, apply the sample to the QIAquick column and centrifuge for 30–60 s.

4. Discard flow-through. Place the QIAquick column back into the same tube.

Collection tubes are re-used to reduce plastic waste.

5. To wash, add 0.75 ml Buffer PE to the QIAquick column and centrifuge for 30–60 s.

6. Discard flow-through and place the QIAquick column back in the same tube.

Centrifuge the column for an additional 1 min at maximum speed.

IMPORTANT: Residual ethanol from Buffer PE will not be completely removed unless the flow-through is discarded before this additional centrifugation.

7. Place QIAquick column in a clean 1.5 ml microcentrifuge tube.

8. To elute DNA, add 50 µl Buffer EB (10 mMTris·Cl, pH 8.5) or H2O to the center of theQIAquick membrane and centrifuge the column for 1 min. Alternatively, for increased DNA concentration, add 30 µl elution buffer to the center of the QIAquick membrane, let the column stand for 1 min, and then centrifuge.

IMPORTANT: Ensure that the elution buffer is dispensed directly onto the QIAquick membrane for complete elution of bound DNA. The average eluate volume is 48 µl from 50 µl elution buffer volume, and 28 µl from 30 µl elution buffer.

Elution efficiency is dependent on pH. The maximum elution efficiency is achieved between pH 7.0 and 8.5. When using water, make sure that the pH value is within this range, and store DNA at –20°C as DNA may degrade in the absence of a buffering agent. The purified DNA can also be eluted in TE (10 mMTris·Cl, 1 mM EDTA, pH 8.0), but the EDTA may inhibit subsequent enzymatic reactions.

Section 10: Fermentas Rapid DNA Ligation Kit (#k1422)

1. Thoroughly mix the 5x rapid ligation buffer prior to use.

2. Add the following to micro centrifuge tube.

Linearized vector DNA	10-100 ng
Insert DNA (at 3:1)	Varies
5x Rapid Ligation Buffer	4 uL
T4 DNA Ligase, 5 U/uL	1 uL
Water, nuclease free	To 20 uL
Total	20 uL

3. Vortex and spin briefly to collect drops.
4. Incubate the mixture at 22 C for 15 min
5. Use 2-5 uL of the ligation mixture for transformation.