Generation of Isogenic Pluripotent Stem Cell Lines

For Study of APOE, an Alzheimer's Risk Factor

by

Mary Frances Lakers

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved April 2017 by the
Graduate Supervisory Committee

David Brafman, Chair
Karmella Haynes
Xiao Wang

Arizona State University

May 2017

ABSTRACT

Alzheimer's disease (AD), despite over a century of research, does not have a clearly defined pathogenesis for the sporadic form that makes up the majority of disease incidence. A variety of correlative risk factors have been identified, including the three isoforms of apolipoprotein E (ApoE), a cholesterol transport protein in the central nervous system. ApoE ε3 is the wild-type variant with no effect on risk. ApoE ε2, the protective and most rare variant, reduces risk of developing AD by 40%. ApoE ε4, the risk variant, increases risk by 3.2-fold and 14.9-fold for heterozygous and homozygous representation respectively. Study of these isoforms has been historically complex, but the advent of human induced pluripotent stem cells (hiPSC) provides the means for highly controlled, longitudinal in vitro study. The effect of ApoE variants can be further elucidated using this platform by generating isogenic hiPSC lines through precise genetic modification, the objective of this research. As the difference between alleles is determined by two cytosine-thymine polymorphisms, a specialized CRISPR/Cas9 system for direct base conversion was able to be successfully employed. The base conversion method for transitioning from the ε3 to ε2 allele was first verified using the HEK293 cell line as a model with delivery via electroporation. Following this verification, the transfection method was optimized using two hiPSC lines derived from ε4/ε4 patients, with a lipofection technique ultimately resulting in successful base conversion at the same site verified in the HEK293 model. Additional research performed included characterization of the pre-modification genotype with respect to likely off-target sites and methods of isolating clonal variants.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

Figure Page

LIST OF TABLES

1. INTRODUCTION

1.1. Overview

Alzheimer's disease is a complicated, costly disease that scientific research has been attempting to understand for over a hundred years. In that time period, though the cause of sporadic Alzheimer's disease was not defined, a multitude of risk factors have been identified. The most impactful genetic risk factor was determined to be the three isoforms of Apolipoprotein E, a cholesterol transport protein in the central nervous system. Though the variants of ApoE have a dramatic effect on an individual's risk of developing Alzheimer's disease, the causal relationship is unknown. This lack of clarity is exacerbated by the interactions of other risk factors and the limitations of study models available. One solution that would isolate ApoE variants from any other risk factor is to use induced pluripotent stem cells. These cells can be genetically modified in vitro from one genotype to another, and then differentiated into the very neural cell types affected by Alzheimer's disease. Generation of such isogenic pluripotent cell lines would provide a well-characterized, easily controlled model for use in longitudinal study of the pathogenesis of Alzheimer's disease.

1.2. Organization

Chapter 2 discusses the current understanding of Alzheimer's disease, including the specific genetic risk factor APOE. Advances in genetic modification techniques are also discussed. Chapter 3 discusses the methodology employed to transfect cells in culture, assess genetic modification, and optimize the clonal line generation procedure. Chapter 4 describes the results generated by these methods. Chapter 5 summarizes the key accomplishments and suggests future work.

2. LITERATURE REVIEW

2.1. Alzheimer's Disease Characterization

Alzheimer's disease is a household name as the foremost cause of dementia. This complex neurodegenerative disorder places a massive burden on society in its economic and quality of life impact. More than ten percent of the population aged 65 and older has Alzheimer's disease. Based on the current incidence, a new patient develops Alzheimer's disease every 66 seconds, a rate that will only increase as the population ages. [1] This prevalence has a cost. Per "Changing the Trajectory of Alzheimer's Disease", a report issued by the Alzheimer's Association, costs associated with patient care alone are projected to rise to more than $1.1 trillion by 2050 if disease prevalence remains unchecked. [2] NIH funding allocated for supporting research rose to $910 million in 2016 from $503 million in 2012. This figure only includes the federal contribution to research funding, with additional costs incurred by the private sector, particularly when clinical trials are involved. [3] At the individual level, the patient is not the only victim of their cognitive decline. Caregivers are severely impacted by the progression of the disease. Over a third of Alzheimer's patients' caregivers are over the age of 65 and susceptible to their own difficulties that come with age. Over 15 million people acted as caregivers in 2016, exhausting time, money, and their own quality of life in aiding dementia patients. [1] Perhaps the greatest burden of Alzheimer's disease is that it has no cure. In 2014, Alzheimer's disease was reported as the sixth leading cause of death in the United States, and the only leading pathology-related cause to have an increased death rate from 2004 to 2014. Age-adjusted death rates declined during this period for all other non-injury causes, including heart disease, cancer, and stroke. [4]

Alzheimer's disease has been studied for more than a century without achieving an absolute understanding of the cause or development. The late-stage physical pathology has been well-documented, featuring intra- and extracellular proteinaceous aggregations and loss of functioning neurons in correlation with cognitive decline. The amyloid cascade hypothesis has been the leading causal suggestion for years, positing that the neurodegenerative pathway begins with the accumulation of extracellular plaques of the amyloid-β peptide. [5] While the normal function of this peptide has not been well understood, the pathologic aggregation has been the subject of intense study. This aggregation is dependent on the cleavage mechanisms that process the amyloid precursor protein (APP). In both pathways, the transmembrane APP is cleaved initially to form a membrane-bound fragment and a secreted fragment. The membrane-bound fragment is again cleaved by γ-secretase to separate the intracellular and extracellular domains. In the non-amyloidogenic pathway, the first cleavage via α-secretase results in a long initial secreted fragment, APPsα, and so leaves a short extracellular domain, Aβ40, to be released by the γ-secretase activity. In the amyloidogenic pathway, the first cleavage via β-secretase takes place closer to the N-terminus of APP, resulting in a shorter secreted fragment than APPsα and leaving a longer extracellular domain for later release as Aβ42. This alternate amyloid-β protein is more prone to aggregation, and as such is indicated as a key component in the pathogenesis of Alzheimer's disease. A variety of options have been suggested as to the form of amyloid-β that prompts the neurogenerative cascade, ranging from the originally-indicated senile plaques to soluble oligomers. [6] One specialized amyloid-β hypothesis suggests that amyloid-β oligomers form neurotoxic calcium ion channels in neurons, leading to cell death and the clinical neurodegenerative effects. [7] Other research complementary to that naming amyloid-β as the key component has demonstrated a positive correlation between synaptic health

and plasticity and secretion of APPsα in animal models. [8] However, the amyloid-β hypotheses have complications. Another common hypothesis as to the casual pathway of Alzheimer's disease features tau phosphorylation and aggregation.

While the pathology described above is characteristic of all cases of Alzheimer's disease, the disorder can be parsed into groups based on the age and heritability of development. Familial Alzheimer's disease (FAD) is characterized by inheritable mutations in APP, presenilin 1, or presenilin 2. These mutations predispose the APP pathway to production of Aβ42 and the resulting accumulation of amyloid-β plaques. FAD is typically early on-set, with patients commonly presenting symptoms in their 40s to early 60s. FAD is also relatively rare, accounting for less than 5% of all Alzheimer's disease cases. Sporadic Alzheimer's disease (SAD) comprises the remainder of Alzheimer's disease cases, and in contrast with FAD, does not correlate with any inheritable, deterministic genetic mutations. SAD is typically late on-set, presenting in patients older than 65. Given the lack of deterministic mutations, improved understanding and potential treatment of SAD requires characterization of the causal pathway that precipitates the known pathology.

2.2. Apolipoprotein E, a Genetic Risk Factor

In the effort to better characterize SAD and its cause, a variety of technologies converged to implicate Apolipoprotein E (APOE) as the single most correlative genetic risk factor. The early 1990s featured two parallel efforts at the Duke Alzheimer's Disease Research Center, one exploring genome-wide association studies (GWAS) to find correlative gene networks, and one studying the pathology of Alzheimer's disease, specifically amyloid-β plaques, to investigate protein function. The results of the former suggested a strong relationship between development of Alzheimer's disease and gene

expression from chromosome 19. [9] The results of the latter indicated that a protein matching the conserved regions of ApoE was present in amyloid-β plaques. After the connection was made between Alzheimer's disease and chromosome 19, and chromosome 19 and ApoE, more specific GWAS results strongly indicated a correlation between isoforms of ApoE and the prevalence of SAD.

ApoE is involved in a variety of pathways in the central nervous system, but has a primary role of cholesterol transport. Produced almost exclusively by astrocytes, the protein is comprised of a helical bundle at the N-terminus and a longer, hinged single helix at the C-terminus. (Figure 1) These two segments of ApoE are distinguished by their function, with the N-terminus binding to transported lipoproteins and the C-terminus interacting with receptor proteins. [10] There are three common variants of APOE; ε2, the protective allele; ε3, the wild-type allele, and ε4, the risk allele. ApoE ε3 is the most common allele in the general population, followed by ε4 and ε2. These variants are distinguished by two single-nucleotide polymorphisms (SNPs) at the 112 and 158 amino acid positions. (Table 1) The corresponding protein isoforms appear to differ in the interaction between the lipid-binding domain and the N-terminus. When an arginine residue is present at the 112 position as in the ε4 variant, the secondary structure allows for a salt bridge to form between the arginine residue at the 61 position and the glutamine residue at the 255 position on the lipid-binding domain. When a cysteine is present in this position as in the ε3 and ε2 variants, the secondary structure does not permit such a salt bridge. [11] The variation in structure between the isoforms is implicated as the cause of the variable pathogenic risk. The most common explanation for the variable risk was a suggestion that the ApoE isoforms have disparate capabilities of binding Aβ fragments for clearance, with ε4 being the least capable and thus allowing

for the extracellular aggregation of senile plaques. [12] However, a more recent hypothesis proposes that ApoE and Aβ are competing substrates in their interaction with cellular clearance pathway receptors, specifically low-density lipoprotein receptor-related protein 1 (LRP1) and LDL receptor (LDLR). [13] Whatever the causal relationship, it is clear that ApoE isoforms are involved in complex pathways.

Allele	Polymorphisms	General Population Frequency	AD Population Frequency	Effect on Risk
ε4	Arg112; Arg158	13.7%	36.7%	3.2-fold more likely (one allele); 14.9-fold more likely (two alleles)
ε3	Cys112; Arg158	77.9%	59.4%	Average (wild-type)
ε2	Cys112; Cys158	8.4%	3.9%	40% less likely

Table 1. Comparison of the three major alleles of APOE with respect to their prevalence and effect on the risk of developing Alzheimer's disease. *[14]*

Other environmental and lifestyle factors have been identified as affecting an individual's risk for developing Alzheimer's disease. Age is the most dramatic correlative factor, with prevalence of Alzheimer's disease increasing in older demographics. One risk factor study found that patient risk of developing Alzheimer's increases by an average of 1.5-1.8-fold in successive age groups spanning five years from age 60 to age 95. [15] Health factors have been of additional interest. Vascular risk factors in particular have been found to correlate with the prevalence of Alzheimer's disease, including midlife hypertension, diabetes, high cholesterol, and lifestyle risks, such as obesity and smoking. [16] Other neurological conditions have been shown to share a correlation with development of Alzheimer's disease as well, including post-traumatic stress disorder

(PTSD) and head injury. In a study of the aging veteran population, it was found that those who suffered from PTSD had a higher incidence of Alzheimer's disease of 1.77-2.31-fold relative to veterans without PTSD. [17] A meta-analysis of 32 studies investigating the link between head injury and the development of Alzheimer's disease later in life found an overall increased risk of 1.51-fold with respect to patients that did not experience head injury. [18] Beyond health factors, education levels have been indicated as a potential risk factor for developing AD. A compilation of international risk factor studies suggested that a lower education corresponds with a doubled risk of developing Alzheimer's disease. [19] Possible explanation for the impact of education suggests that an increased "cognitive reserve" provides additional synapses that can be used to maintain homeostasis when neurodegeneration occurs. [20] This great diversity of risk factors and their potential interactions emphasizes the difficulty in elucidating the effect of APOE without accounting for these factors.

2.3. Methods of Studying ApoE Isoforms

Researching neurological disorders presents many challenges, with one of the greatest being determining appropriate modeling methods. This is critical for comparison of the three isoforms of APOE, requiring a tightly controlled, longitudinal, and human-transferable model. The approach closest to the ground truth is analysis of primary human tissue, capturing the complexity of the pathology in question exactly as it presents in the nervous system with all known and unknown factors at work. Unfortunately, primary tissue in the central nervous system is typically available for harvest at time of death, and as such does not typically provide insight into the early stages, developmental pathway of disease, or effects of treatment. Reliance on primary human tissue also precludes explicit control of common risk factors, preventing study of

truly isolated variables as is needed to characterize the impact of APOE variants. Use of animal models rectifies some of these issues, allowing for more direct control and the ability to perform analyses over time. A variety of mouse models have already contributed to greater understanding of Alzheimer's disease. The positive correlation between the presence of $A\beta_{42}$ and the formation of the amyloid-β plaques, increased plaque formation with subject age, and the potential neurotoxicity of soluble amyloid-β oligomers were all identified and confirmed in various transgenic mouse models. [21] One drawback of animal models is the inability to model SAD. Since the root cause of SAD is still under investigation, generation of animal lines that express the Alzheimer's disease phenotype relies on overexpression of FAD mutations. [22] Even more specialized animal models are required to generate neurofibrillary tangles as that pathology is typically not seen in models that only rely on overexpression of Aβ42. [23] While such direct approaches are useful when analyzing downstream effects of the pathology of Alzheimer's disease, they do not allow for study of the natural cause as artificial causes are induced. Other cause for concern with animal models is the lack of translation of successful therapies from animal to human trials. This gap between the two systems indicates an incomplete approximation of the disease using the current modeling approaches. [24] Such a gap adds potential error to an already complex pathway.

The third approach available for studying the effect of APOE is the use of cultured human cells, specifically stem cells. Human pluripotent stem cells can be used for modeling tissues in a dish that are otherwise difficult to acquire. Previously, such cells could be obtained as embryonic stem cells from blastocysts generated through in vitro fertilization. This method met with scrutiny due to the limited cell numbers achieved and

the ethical dilemma posed by their source. In 2006, it was reported that terminally differentiated cells could be "reprogrammed" to a pluripotent state by exposing the cells to certain transcription factors (Oct4, Sox2, Klf4 and c-Myc). [25] Other methods have since been developed to both increase efficiency and reduce the lasting footprint of the conversion method. [26] Delivery of these four vectors via the Sendai virus has proven to be a robust solution as these recombinant vectors are not integrated into the host genome. Reprogramming with the Sendai virus is also reported with higher efficiency than other viral techniques. [27] Using such methods, human induced pluripotent stem cells (hiPSCs) can be generated from any living individual, retaining the genotype of the patient with the added abilities to self-renew in culture and to be differentiated into any desired cell type. Hypothetically, these cells would also express pathological phenotypes when differentiated into the tissues that are diseased in the patient. A variety of studies have demonstrated this modeling to be possible, from imitation of action potential changes expressed in cardiomyocytes derived from hiPSCs from patients with a mutation associated with cardiac arrhythmia, to recapitulation of defective neuronal differentiation and migration in hiPSCs from patients with familial dysautonomia. [28], [29] Familial Alzheimer's has also been successfully modeled using patient-derived hiPSCs with respect to generation of both amyloid-β aggregation and neurofibrillary tangles of hyperphosphorylated tau given known mutation of APP. [30] This method provides a platform for rigorous control of environmental factors, grants ease of access for analysis and longitudinal study, and removes the complications of using another species as the model basis. Given these hiPSCs as intrinsically clonal populations, there is the opportunity to use gene modification methods at the cellular level to create parallel, isogenic cell lines to isolate the APOE variant as the only distinguishing characteristic.

2.4. Genetic Modification Methods

Gene modification methods have evolved rapidly in recent history. The use of meganucleases, engineered versions of restriction enzymes, was one of the first adaptations of native DNA modification machinery for engineering purposes. Each meganuclease targets a large, typically unique recognition site in the genome and effects a double-stranded break (DSB) in the genome at the recognition site. Zinc finger nucleases (ZFNs) are combinations of protein monomers that each match three nucelotides. The combined targeting proteins can be conjugated to the FokI nuclease to create the same DSB described previously, activating endogenous repair pathways. ZFNs are limited in size by the heterogenous structure of the targeting proteins that limits the Transcription activator-like effector nucleases (TALENs) are similar to ZFNs in that they are comprised of ordered recognition protein domains and require conjugation to an endonuclease to effect the DSB, but have an advantage in that each targeting monomer only targets a single base. This allows for more precise design of targeting mechanisms; however, the conserved nature of the targeting monomers can complicate viral transduction methods of vector delivery. Both ZFNs and TALENs, while programmable, are time-consuming and expensive to manufacture given that each targeted sequence must be translated into the targeting mechanism in a specific, compatible sequence of monomers. The revolution in programmable targeting arrived in the form of Clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) proteins. These systems relied on DNA targeting through specific guide RNA strands and the resulting Waston-Crick base-pairing rather than a combination of protein monomers. This simplification of the targeting mechanism reduced the time and cost involved with design for a new target and made genetic modification more accessible.

Though the evolution of genetic modification methods has advanced rapidly with respect to the ease and specificity of the targeting mechanism, all four of the methods previously discussed create a DSB in the targeted DNA sequence and rely on the same repair mechanisms to silence or alter the gene. If the endogenous repair machinery solves the damaged DNA without additional intervention, the two strands are religated through a process called non-homologous end joining (NHEJ). In NHEJ, the native machinery attempts to join the broken strands by adding or removing nucleotides based on any overhangs present. These insertions and deletions, or indels, are useful in silencing genes by altering the sequence or even causing a frameshift mutation. The next step in modification techniques provided a method to not just silence genes, but to alter the sequence. In homology-directed repair, long homology arms are supplied that span the native genome sequence across the DSB, but with a desired mutation in between the matching sequences. The native machinery then uses these homology arms as a template for filling the gap rather than completing the break with random nucleotides, inserting the mutation into the repaired DNA. HDR has since been validated many times over as a method for supplying an altered sequence for the targeted gene, but the method still has many drawbacks. These include a high prevalence of off-target effects, low modification efficiency, and the additional work required to generate and deliver the homology arms.

In order to improve upon the efficiency and specificity of CRISPR/Cas9, a variety of techniques have been developed. One of these techniques, published by the David Liu lab at Harvard University, is centered around direct base conversion of the targeted nucleotide. [31] This improved base editor includes multiple functional groups added to a catalytically inactivated Cas9 protein. The first of these is the cytidine deaminase that converts the cytosine nucleotide to uracil. The second additional functional group, uracil

DNA glycosylase inhibitor, prevents the native base excision repair machinery from removing the RNA base and replacing it with the original cytosine. The third modification was restoration of activity to the HNH endonuclease domain which creates a single-stranded break in the off-target strand that, at this point, is still guanine. The damage prompts other endogenous repair machinery to remove the guanine through mismatch repair and replace it with adenine. Other machinery eventually replaces the uracil with thymine, completing the transition from cytosine to thymine without ever creating a DSB or relying on the low efficiency of HDR. The efficiency of this method is reported as 15-75% with less than 1% indel formation, compared to 0.5% conversion efficiency and 4.3% indel formation when standard CRISPR/Cas9 and HDR are used to create the mutation. Such precision in correcting point mutations brands this specialized base editor as an ideal tool for use in the generation of these isogenic cell lines.

3. METHODS

3.1. Verification in HEK293 Cell Line

Determination of the gene modification system included identification of the vectors, delivery method, and platform for verification. The human embryonic kidney cell line HEK293 was chosen as the verification model for the base modification system. This line is reliably cultured in vitro and was the platform used to demonstrate the efficacy of the modified CRISPR/Cas9, providing an opportunity to replicate the published results prior to attempting the direct base conversion in hiPSCs. Use of this model also allowed for optimization of delivery method parameters for cell survival and gene modification.

CRISPR/Cas9 systems have been successfully delivered in a variety of methods since their discovery. This research posed additional constraints that limited the delivery options available for use. Given that the intent of generating these isogenic pluripotent cell lines is to minimize the genetic footprint of the modification, stable integration of the Cas9 vector into the native genome was distinctly undesirable. This precluded use of common viral vectors as most function by inserting the transduced cassette into the host genome, altering the genotype of the cell line and potentially interrupting other vectors. Such permanent expression of the Cas9 protein could also exacerbate off-target effects as the protein would continue to be expressed. Lentiviral and adeno-associated viral vectors were among those rejected in favor of transient expression. [32] Another consideration for the gene delivery method was the low yield process of clonal line generation in principal. With the known poor survival of single cell isolation, it was vital that the method chosen did not add additional cytotoxic burden and further decrease the chances of expansion of modified cells. In the interest of maintaining consistency with the published method and the described constraints, electroporation was chosen as the method of transient transfection.

Three separate plasmids were used to deliver the required components. The modified Cas9 vector, pCMV-BE3, was a gift from David Liu (Addgene plasmid # 73021). The targeting sequences for the modification sites were cloned separately into the backbone pFYF1320, a gift from Keith Joung (Addgene plasmid # 47511). The base editor and guide RNA plasmids were delivered at a mass ratio of 3:1 in favor of the base editor in all experiments. The plasmid containing the fluorescent reporter GFP, pEF-GFP, was a gift from Connie Cepko (Addgene plasmid # 11154). This plasmid was delivered at as low a concentration as possible in order to increase the likelihood that

positive GFP expression was correlated with transfection and expression of the other two plasmids.

The evaluation methods for the HEK293 verification were used as the basis for all evaluation. Verification of survival and transfection was first assessed qualitatively via fluorescent microscopy. Each sample was photographed at twenty-four hour intervals following electroporation up to seventy-two hours using brightfield and fluorescent channels (EVOS® FL). Brightfield imaging indicated the relative viability of the cells post-transfection. Positive GFP expression in the fluorescent channel was used as indication of successful transfection. At seventy-two hours post-transfection, this relative viability and GFP expression were quantified using flow cytometry. Genetic modification was assessed through sequencing. Sanger sequencing was performed on PCR products generated from bulk samples. Sanger sequencing and next-generation Illumina sequencing were also performed on PCR products from samples enriched for GFP positive cells through fluorescence-activated cell sorting (FACS). The next-generation sequencing method provided quantification of the modification efficiency.

The first HEK293 verification experiment employed increasing plasmid concentrations to determine the concentration required to effect the desired mutation at a detectable scale. HEK293 cells were harvested from 10cm culture plates at ~80% confluency using Accutase. The cells were pelleted and washed with phosphate-buffered saline (PBS), then pelleted and resuspended in Neon® Resuspension buffer at approximately $5*10^6$ cells per aliquot. The three-plasmid system was added to each aliquot in a gradation of plasmid concentration from $0.12\mu g/1*10^6$cells to $1.6\mu g/1*10^6$cells with a constant concentration of $20ng/1*10^6$cells of pEF-GFP and a 3:1 ratio of pCMV-BE3:APOE- pFYF1320. Each sample was then electroporated using a

100µL Neon® Tip with settings of 1100V pulse amplitude, 20ms pulse length, and 2 pulses. (Neon® Transfection System). Samples were immediately deposited drop-wise into a 10cm culture plate containing antibiotic-free media, DMEM + 10% FBS, and 5µM Rho kinase inhibitor, and returned to the incubator. Microscopy images were recorded every twenty-four hours followed by harvest and flow cytometry analysis after seventy-two hours total. Genomic DNA (gDNA) was extracted from the harvested cells. (QIAGEN® DNeasy® Blood & Tissue Kits) PCR amplification of the APOE gene was performed using the Phusion® High-Fidelity PCR Kit. The resulting PCR products from the heterogeneous population were submitted for Sanger sequencing to assess modification on a bulk scale.

Following this coarse assessment of modification, the transfection procedure described above was repeated with a set plasmid concentration of 1.6µg/1*10^6cells. At seventy-two hours post-transfection, the transfected cells were sorted via FACS based on GFP expression. gDNA was harvested from the GFP positive, GFP negative, and unsorted populations, followed by PCR amplification of the APOE gene. These resulting PCR products were submitted for both Sanger sequencing and next-generation sequencing.

3.2. Translation to hiPSCs

The generation of the hiPSC lines from patients was performed by a third party. Peripheral blood mononuclear cells (PBMCs) were harvested and redirected to pluripotency by exposure to the KMOS transcription factors, KLF4, c-MYC, OCT4, and SOX2, transiently expressed by Sendei virus transduction. Two lines of interest were used specifically in this project, one from an Alzheimer's disease patient, iPS-160, and one from a non-demented control patient, iPS-384. The genotyping for each line was performed using Sanger sequencing for both the PBMCs and the reprogrammed hiPSCs.

The iPS-160 and iPS-384 cell lines were used to transition from the model HEK293 cell line to the patient-derived hiPSCs. Simple carryover of the electroporation technique used to transfect the HEK293 cell line revealed that more optimization was required to yield living, transfected cells. The electroporation settings were revisited in a multi-parameter assessment of cell survival and GFP expression with respect to changes in voltage, pulse length, and pulse number. Voltage was either 1100V and 1200V, pulse length was 10ms, 20ms, or 30ms, and pulse number was either 1 or 2 pulses. All samples were plated in a single well of a 24-well culture plate and assessed for survival and GFP expression via flow cytometry after seventy-two hours. The optimal method resulting from this analysis of 1200V, 10ms pulse length, and 2 pulses was then used as the baseline comparison for evaluation of a method proposed by a publication describing Cas9-mediated modification in hiPSCs. This alternate approach employed a lesser cell electroporation density of $5*10^5$ cells, varied electroporation settings of 1100V, 30ms pulse length, and 1pulse, and the use of a smaller 10µL electroporation cuvette tip. [33] For this experiment, the variables modulated included plasmid concentration, tip size and electroporation cell density, and the set of three electroporation parameters. Plating density was held constant at $1*10^6$ cells per well of a 6-well culture plate and cells were assessed via flow cytometry after seventy-two hours.

Given the optimal electroporation protocol, the remaining variable to potentially improve transfection efficiency was plasmid concentration. Various series of testing were performed with increasing concentrations of the pCMV-BE3:APOE and pFYF1320 plasmids and the same low concentration of pEF-GFP. The range of concentrations spanned 0.8µg/$1*10^6$cells to 4.0µg/$1*10^6$cells. As in previous experiments, electroporated cells were cultured for seventy-two hours prior to analysis via flow

cytometry for survival and GFP expression. The maximum plasmid concentration was also repeated at a larger scale in order to assess modification on a bulk scale via Sanger sequencing following FACS-enrichment based on GFP expression.

Alternate gene delivery approaches were also explored, specifically lipopolymer packaging methods. The products TransIT®-LT1 and TransfeX™ were both evaluated as transfection reagents. (TransfeX™ Transfection Reagent, ATCC® ACS-4005™) The efficacy of TransIT®-LT1 was tested in the iPS-384 cell line. Cells were plated at standard passaging density and allowed to expand for forty-eight hours. The transfection solution was prepared by incubating 2µg total plasmid with varying ratios of TransIT®-LT1 ranging from a µL TransIT®-LT1: µg plasmid ratio of 2:1 to 6:1. The established ratio of 3:1 for pCMV-BE3:APOE- pFYF1320 was used for the plasmid distribution with 100ng pEF-GFP. Following the incubation period, the transfection solution was added drop-wise to the wells. The cells were immediately returned to the incubator and assessed for survival and GFP expression after twenty-four hours. gDNA was also harvested at this point for Sanger sequencing of the APOE gene. A larger scale transfection was repeated using the TransIT®-LT1. After twenty-four hours, these samples were sorted via FACS for GFP expression. The gDNA associated with the enriched and depleted populations was submitted for analysis via next-generation Illumina sequencing.

The efficacy of TransfeX was next to be evaluated. For this experiment, iPS-160s were plated in a 6-well at standard density and allowed to expand for twenty-four hours. The transfection solution was prepared by incubating 4µL TransfeX with 2µg total plasmid in serum-free media at room temperature. The established ratio of 3:1 for pCMV-BE3:APOE- pFYF1320 was used with 100ng pEF-GFP. Following the incubation

period, the transfection solution was added drop-wise to the wells. The iPS-160s were immediately returned to the incubator. Microscopy and flow cytometry analysis occurred twenty-four hours after transfection.

3.3. Preparation for clonal line generation

In parallel with the efforts to optimize the delivery of the base modification system, parameters for the clonal line isolation and characterization were explored. Given the emphasis on a homogenous population, it was preferable to minimize the number of cells from which the clonal line would be expanded. FACS was chosen as the methodology for its ability to isolate single cells. In preparation for sorting modified cells, untransfected cells were sorted to predict a survival rate of hiPSCs and provide guidance as to the cell numbers required. Untransfected iPS-384 cells were treated with 5μM RockI for 24 hours, then enzymatically harvested using Accuctase. Viable cells, distinguished by relatively low side scatter and relatively high forward scatter, were sorted into 96-well plates containing E8 media and 10μM RockI. 1-, 2-, 10- and 100-event sorts per well were performed in replicate. Plated cells were cultured for 14 days and assessed for cell survival and proliferation via microscopy.

In addition to testing of the isolation method, a procedure for assessing off-target effects of the base modification system was designed. Five potential off-target sites were generated based on the sgRNA sequences for each modification site via CCTop, a prediction algorithm published by the Heidelberg Centre for Organismal Studies. [34] Appropriate primer pairs were designed for each site. The baseline pre-modification sequence for each of the ten sites was determined for the iPS-160 cell line via PCR and Sanger sequencing.

4. FINDINGS

4.1. Direct Base Conversion in HEK293 Cells

The direct base conversion using the model HEK293 cell line verified the transfection method, the targeting system for the ApoE ε3 to ApoE ε2 conversion, and the colocalization of the plasmids that allows for selection via GFP expression. Fluorescent microscopy following transfection of the HEK293 cells clearly indicated successful transfection and expression of the GFP cassette. (Figure 1). Survival post-transfection was also supported by microscopy as cells were seen to expand in culture, recovering from the stress inflicted by electroporation. The quantification of these two results, survival and GFP expression, indicated a convergence between surviving cells and surviving cells that were GFP positive as plasmid concentration increased, suggesting both increased cell death and increased transfection efficiency with high plasmid concentrations. (Figure 2).

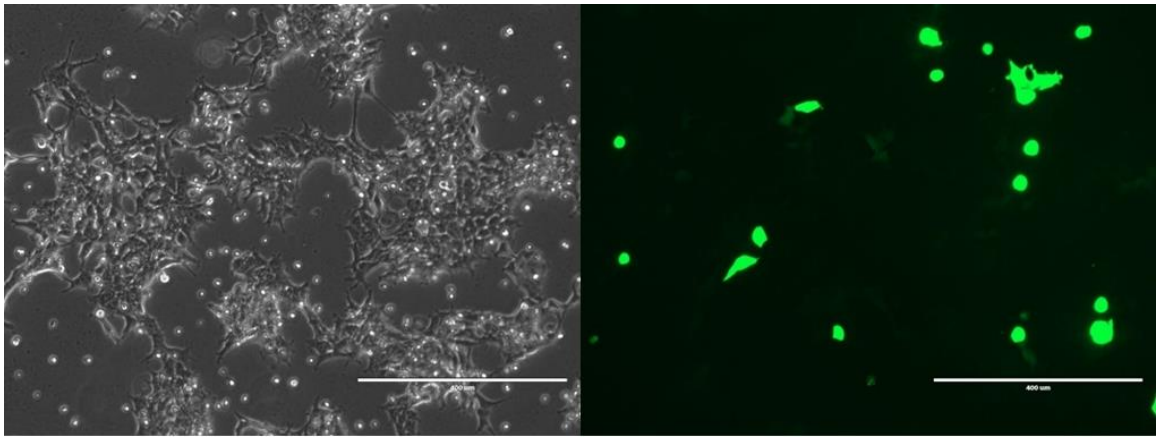

Figure 1. Microscopy images of HEK293 cells 72hrs post-transfection. Brightfield image (left) with the corresponding fluorescent image (right). Electroporation plasmid concentration of 1.6μg/1*10^6cells. 10x magnification.

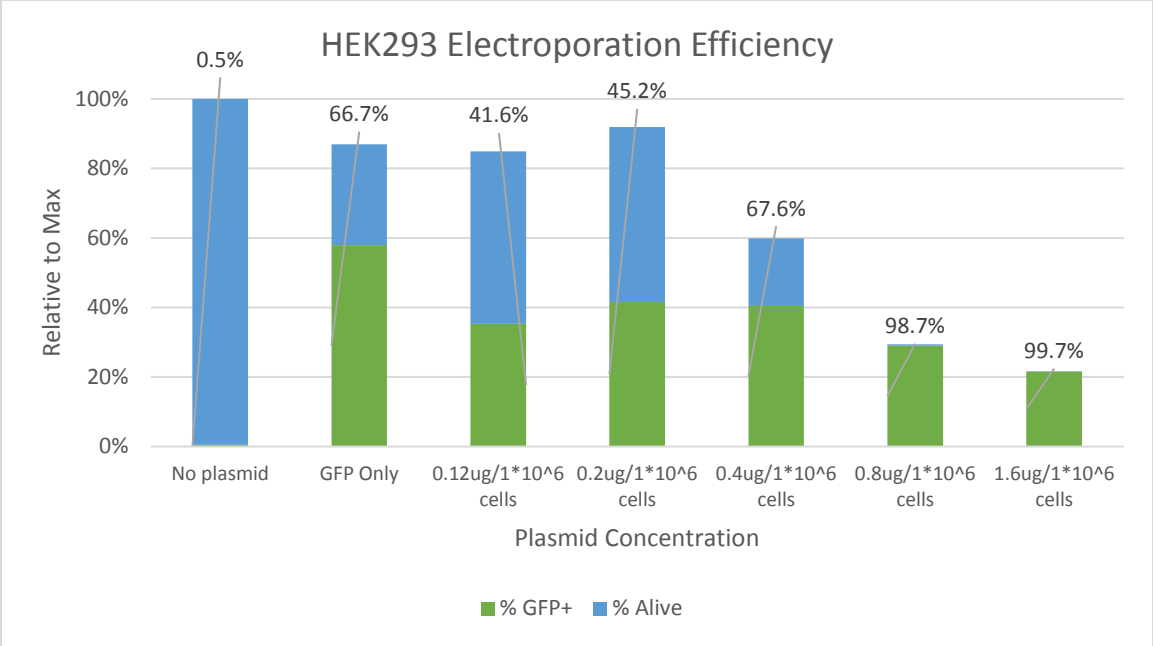

Figure 2. Flow cytometry results characterizing yield of live and GFP-positive HEK293 cells 72hrs post-transfection. Total yield was normalized to the control sample. Data labels indicate the percent of live cells that were found to be GFP-positive for each sample.

The sequencing results for the varying levels of concentrations did not show a significant cytosine to thymine conversion, however, the 1.6μg/1*10^6 cells plasmid concentration sample showed a clear thymine signal at the three bases expected to be converted by successful base editing. (Figure 3). This signal indicates that the higher plasmid concentration is required to see base conversion in a bulk analysis. When the same concentration was repeated followed by enrichment of the GFP positive cells, the thymine signal is much more pronounced in the GFP positive population than in the unsorted population. No thymine signal above background noise is observed for the GFP negative population, all results indicating that there is some colocalization of expression of the three plasmids. (Figure 4).

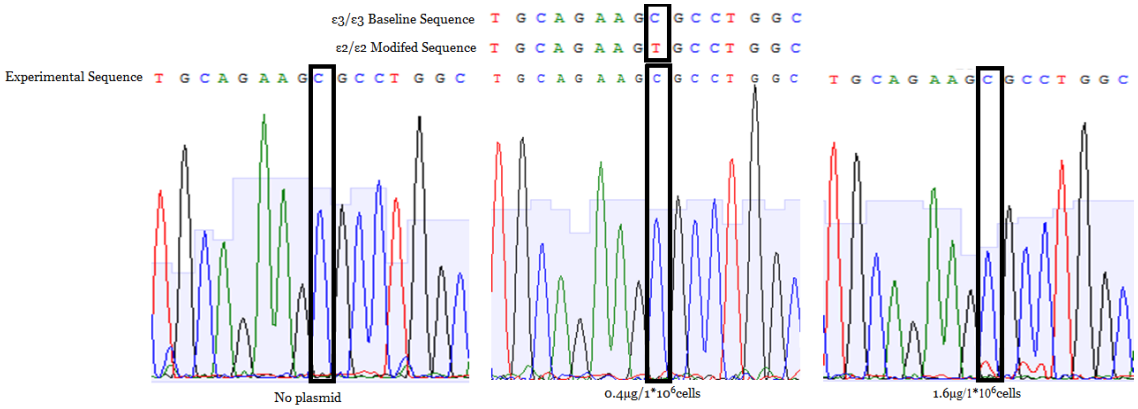

Figure 3. Sanger sequencing results for bulk sequencing HEK293 samples at the 158 amino acid site. Results not shown for electroporation plasmid concentrations 0.12µg/1*10⁶ cells, 0.2µg/1*10⁶ cells as results are represented by the 0.4µg/1*10⁶ cells sample. Black boxes indicate the nucleotide targeted for mutation from cytosine to thymine.

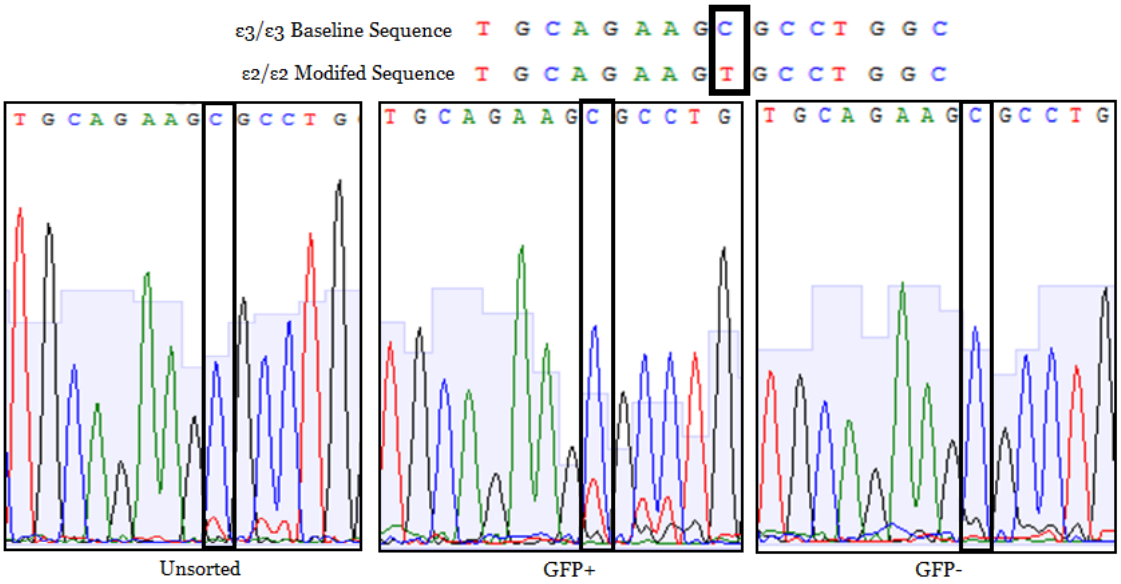

Figure 4. Sanger sequencing results for FACS-enriched HEK293 samples at the 158 amino acid site. Black boxes indicate the nucleotide targeted for mutation from cytosine to thymine.

The use of next-generation sequencing methods to achieve a quantitative measure of this thymine conversion revealed that approximately 14% of the HEK293 cells were successfully modified from an *APOE* ε3/ε3 genotype to having at least one *APOE* ε2 allele. Interestingly, the GFP positive population had a lesser incidence of indels than the GFP negative population. (Figure 5).

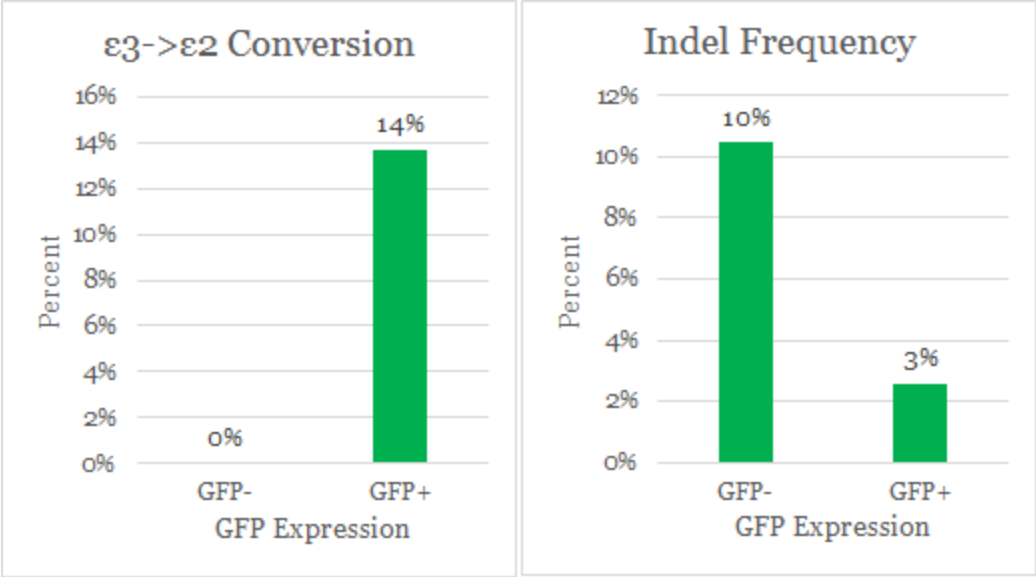

Figure 5. Next-generation Illumina sequencing results for FACS-enriched HEK293 samples indicating genotype based on the sequence at the 158 amino acid site.

4.2. Generation of hiPSCs

The generation of hiPSCs required verification of a stable genome over the course of reprogramming. Sanger sequencing confirmed that the iPS-160 and the iPS-384 cell lines were genotyped as *APOE* ε4/ε4 before and after the reprogramming occurred. (Figure 6).

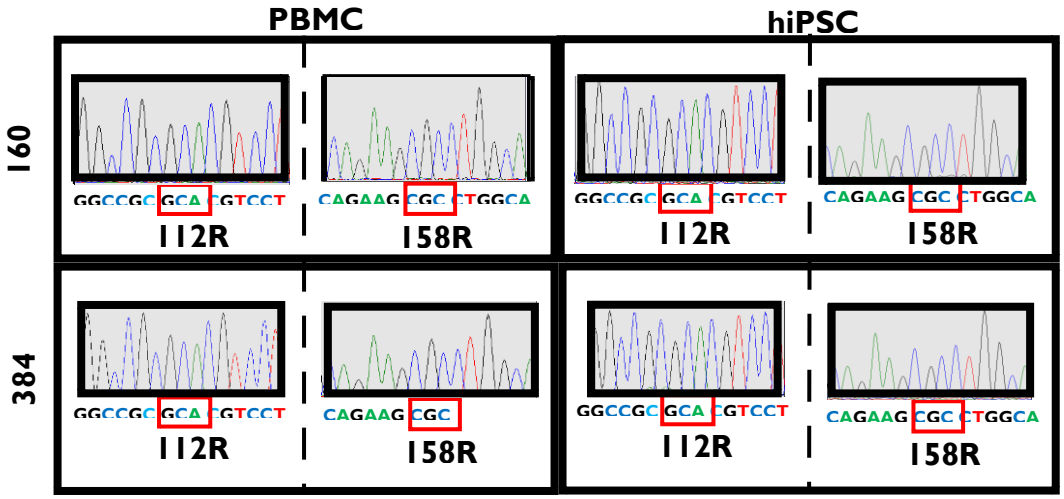

Figure 6. Genotyping of hiPSC lines via Sanger sequencing. Sequencing was performed before and after reprogramming. The red boxes indicate the codons whose polymorphisms define the allele present.

4.3. Electroporation Optimization in hiPSCs

Given the lack of success generating viable cells using the exact protocol used to effect transfection in the HEK293 cells in addition to a general lack of literature citing protocols using the Neon Transfection System to modify hiPSC lines, it was necessary to take such an involved approach to determining the optimal electroporation settings. Evaluation of different combinations of voltage, pulse length, and number of pulses suggested that a combination of 1200V, 10ms or 20ms, and 1 or 2 pulses generated consistent results with respect to survival and GFP expression as well as limited variability between the different combinations. 1200V, 10ms, and 2 pulses was selected as the protocol with the average results for those groups. (Figure 7).

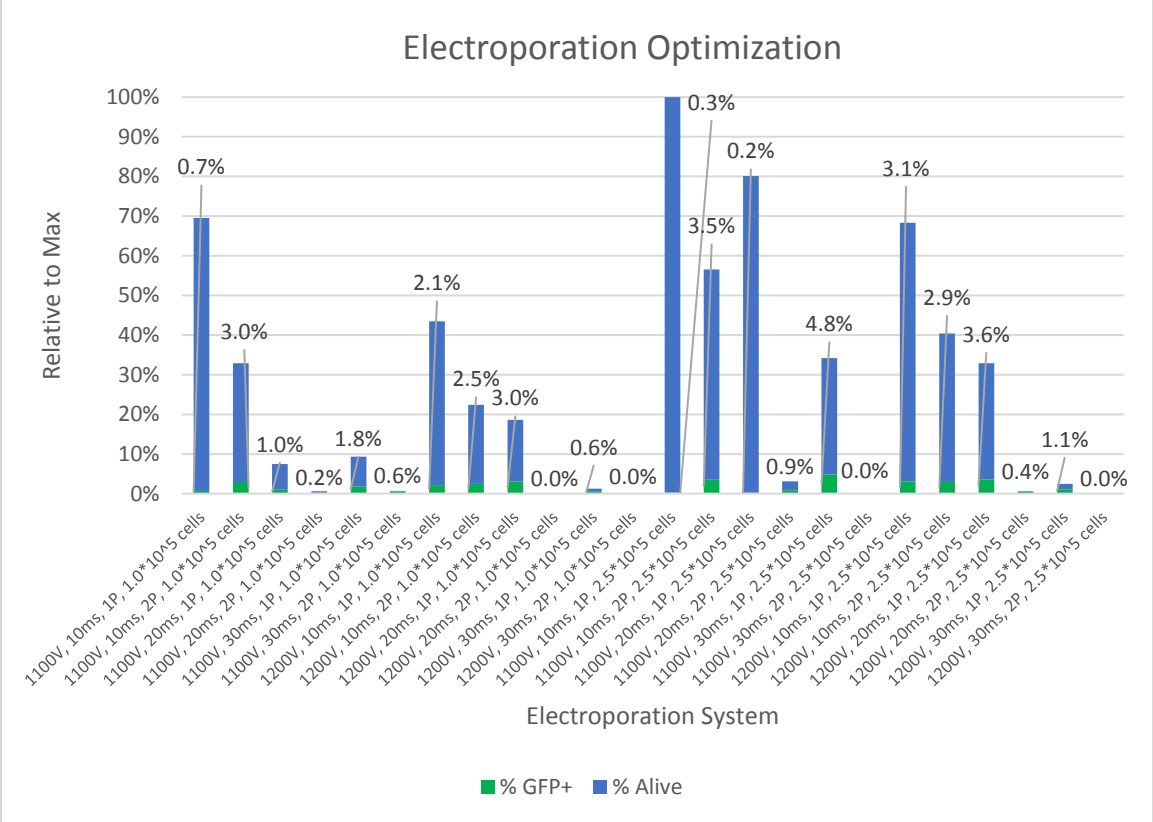

Figure 7. Flow cytometry results from 24-well optimization of the electroporation settings used characterizing yield of live and GFP-positive iPS-160 cells 72hrs post-transfection.

Discovery of literature that did use the Neon Transfection System in pluripotent stem cells offered an additional parameter of the electroporation tip size and the resulting possible electroporation cell density. Comparison of the previously established optimal method and this method did not offer evidence for using the 10µL tip procedure. Modification of the electroporation tip from 100µL to 10µL did not improve the viability or transfection efficiency of the electroporation method. (Figure 8).

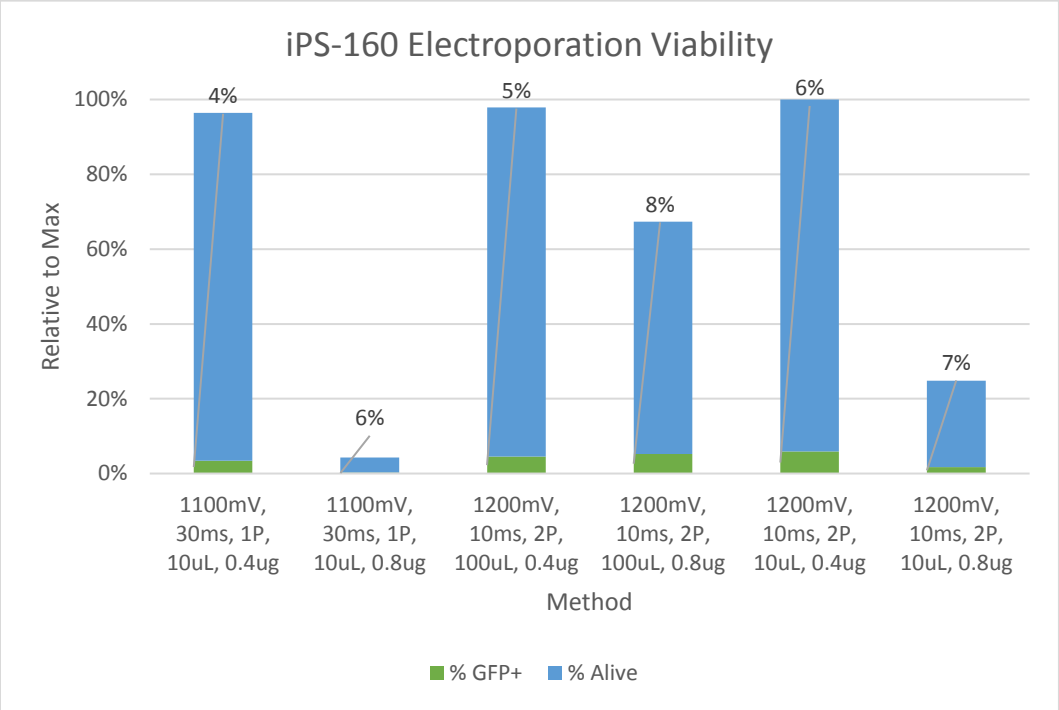

Figure 8. Flow cytometry results characterizing yield of live and GFP-positive iPS-160 cells 72hrs post-transfection. Total yield was normalized to the maximum value. Data labels indicate the percent of live cells that were found to be GFP-positive for each sample.

Plasmid concentration remained as the final major variable. Achieving convergence between survival and GFP expression was desired as this convergence in the HEK293 cell verification model corresponded with successful base editing. Such a convergence would also increase the chances of any one live cell expressing the genetic modification given the evidence supporting coexpression of all three plasmids.

Increasing the base editing plasmid concentration never achieved such a convergence, suggesting that the optimal maximum plasmid concentration has not yet been tested. (Figure 9). This is further supported by the lack of successful cytosine to thymine conversion at both conversion sites observed even after FACS-mediated enrichment for GFP expression. (Figure 10). The seemingly heterogenous representation of the 112-site sequence in both the GFP positive and the GFP negative populations presents an immediate complication for evaluating base editing at that site, particularly as the earliest passages of these cell lines did not display such heterogeneity. Future work assessing the genomic stability of these lines will be required, particularly as the 112-site is that to be targeted in achieving the APOE ε4 to APOE ε3 conversion for generating the first new clonal lines.

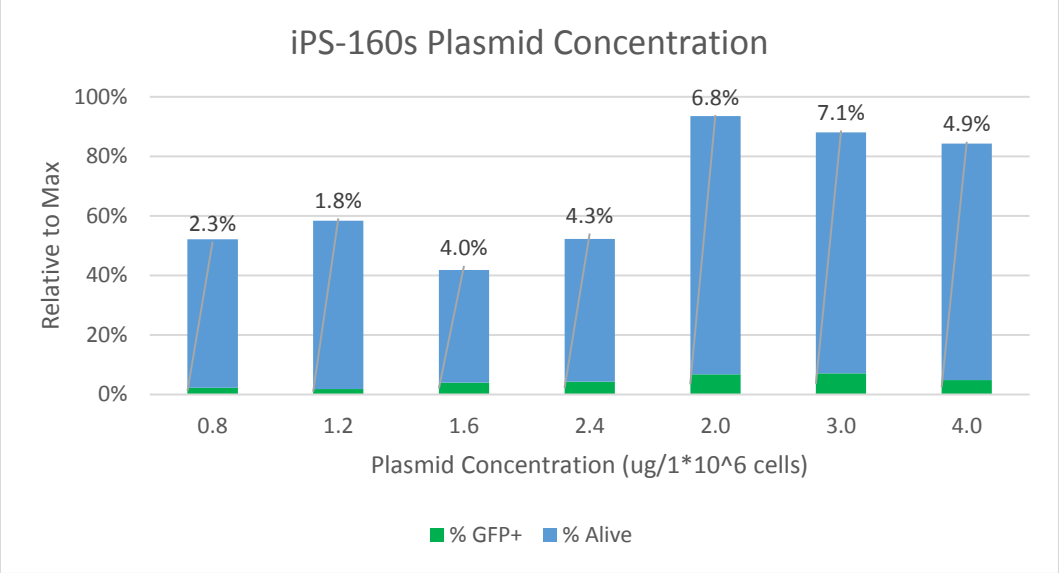

Figure 9. Flow cytometry results from plasmid concentration optimization characterizing yield of live and GFP-positive iPS-160 cells 72hrs post-transfection. Data labels indicate the percent of live cells that were found to be GFP-positive for each sample.

Target Site	# Live Cells	# GFP+ Cells	% GFP+
112	941281	101744	10.8%
158	711000	103000	14.5%

Table 2. FACS results from large-scale electroporation of iPS-160 cells 72hrs post-transfection, including yield and relative GFP expression for both samples.

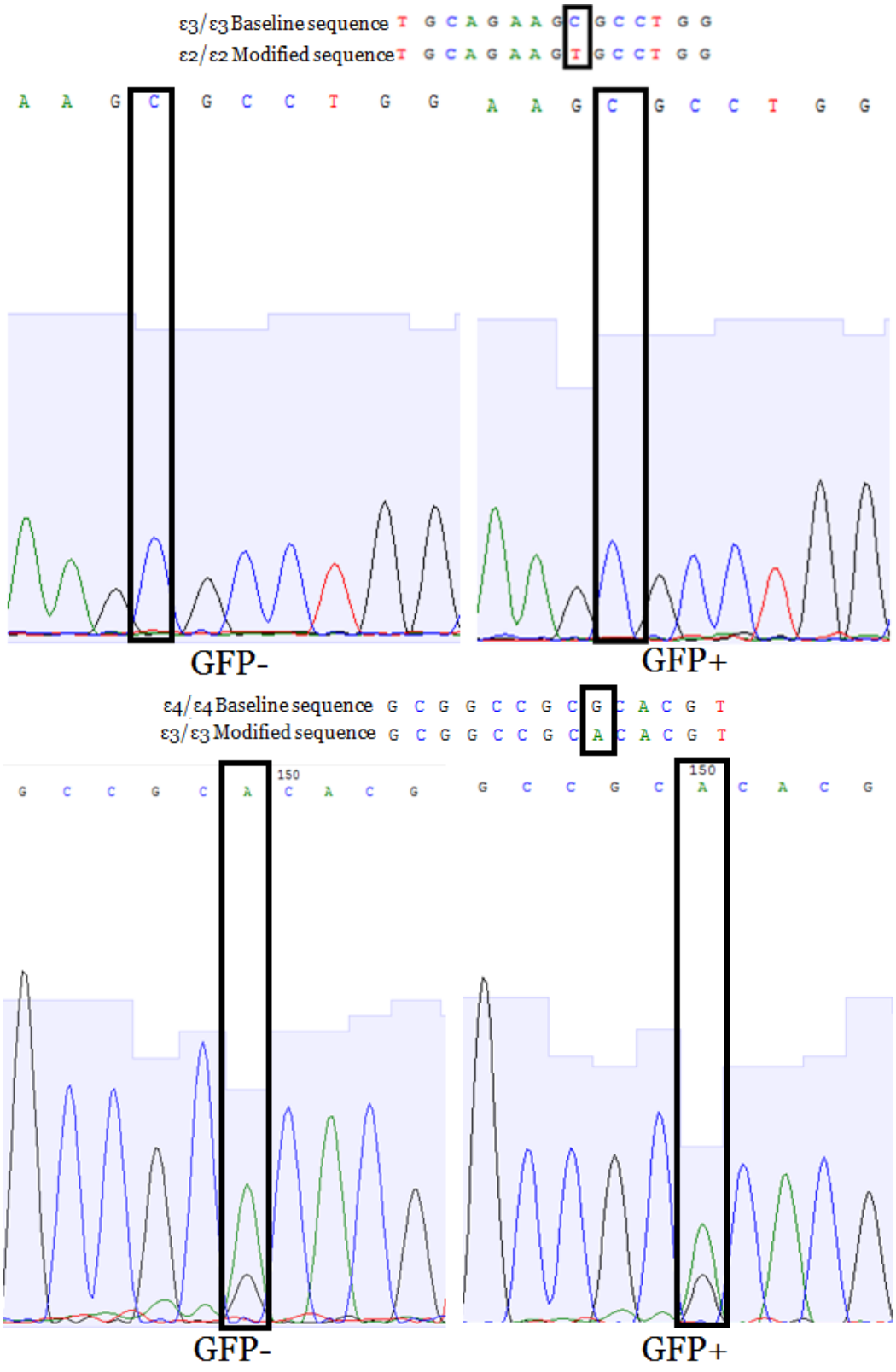
ε3/ε3 Baseline sequence T G C A G A A G C G C C T G G
ε2/ε2 Modified sequence T G C A G A A G T G C C T G G
A A G C G C C T G G
A A G C G C C T G G
GFP-
GFP+
ε4/ε4 Baseline sequence G C G G C C G C G C A C G T
ε3/ε3 Modified sequence G C G G C C G C A C A C G T
150
G C C G C A C A C G
150
G C C G C A C A C G
GFP-
GFP+

Figure 10. Sanger sequencing results from the bulk enrichment of the large-scale electroporation of iPS-160 cells. Black boxes indicate the nucleotide targeted for mutation from cytosine to thymine for each conversion site.

4.4. Alternate Transfection Methods in hiPSCs

Use of the TransIT®-LT1 reagent did not result in a greater incidence of GFP expression in live cells than that observed in electroporation. (Figure 11). However, the sequencing results clearly indicated a conversion signal above background at the 158 site for the APOE ε3 to APOE ε2 conversion. This conversion was most clearly observed in the sample with a 4:1 ratio of µL TransIT-LT1: µg total plasmid. (Figure 12). The next-generation sequencing results confirmed and quantified the base conversion for a replicate sample, revealing an editing efficiency of 4.19% and indel frequency of 1.35%. (Figure 13). Given that this sequencing was performed on a bulk population, it is probable that the editing efficiency would be even greater if quantified from an enriched population based on positive GFP expression.

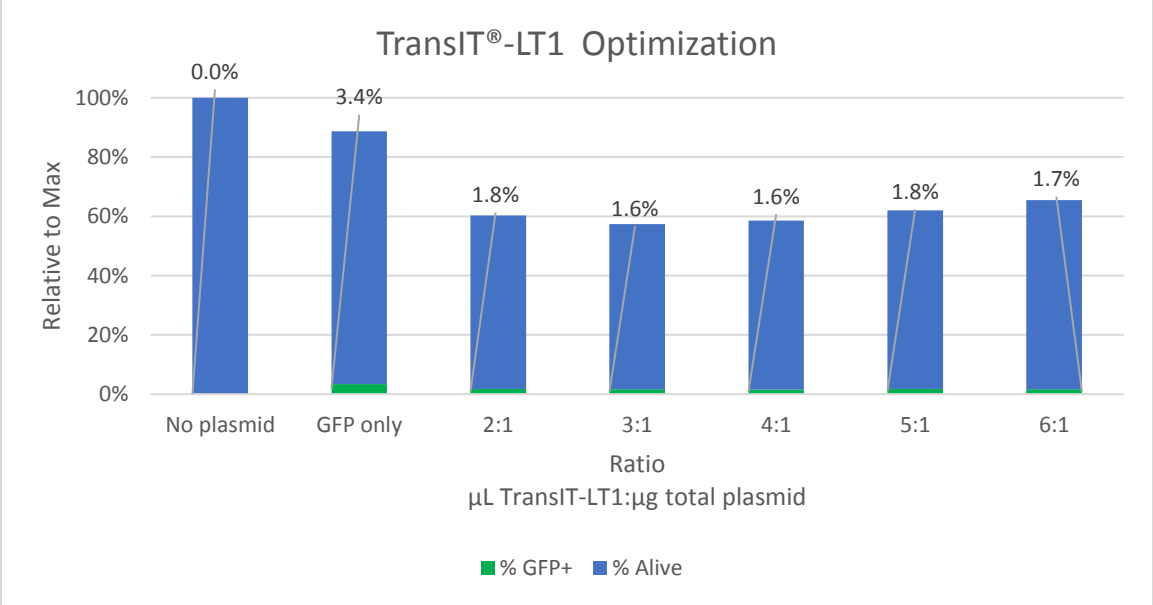

Figure 11. Flow cytometry results from transfection reagent ratio optimization for TransIT®-LT1 characterizing yield of live and GFP-positive iPS-384 cells 24hrs post-transfection. Data labels indicate the percent of live cells that were found to be GFP-positive for each sample.

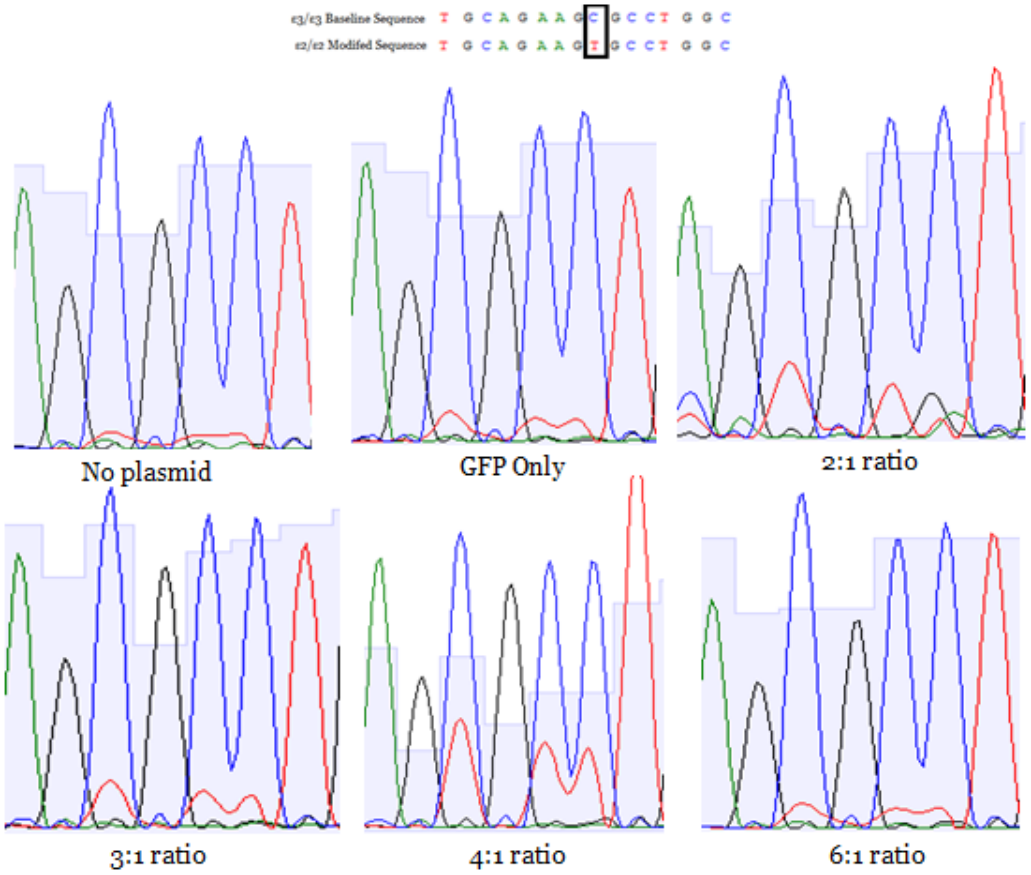

Figure 12. Sanger sequencing results from the transfection reagent ratio optimization for TransIT®-LT1.

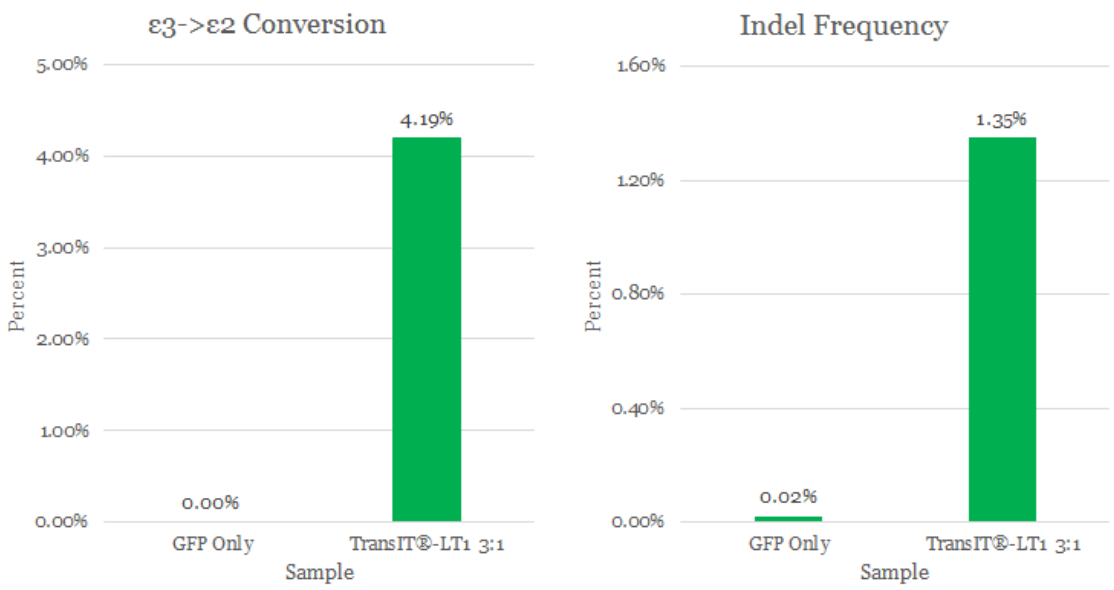

Figure 13. Next-generation Illumina sequencing results for TransIT®-LT samples indicating genotype based on the sequence at the 158 amino acid site. The GFP-only sample is presented here as a control.

The alternate transfection reagent Transfex did result in an improved incidence of GFP expression in live cells 24-hours post-transfection than that observed via electroporation 72-hours post-transfection. Such increased efficiency is readily apparent in the microscopy images captured at this point in time. (Figure 14) This suggests that use of this reagent may result in an even greater successful base editing efficiency than the TransIT®-LT1 reagent. Such prediction is further supported by the flow cytometry analysis of the transfected cells in which it was observed that use of the Transfex reagent appears to be closer to achieving the previously mentioned convergence between live cells and GFP expression. (Figure 15)

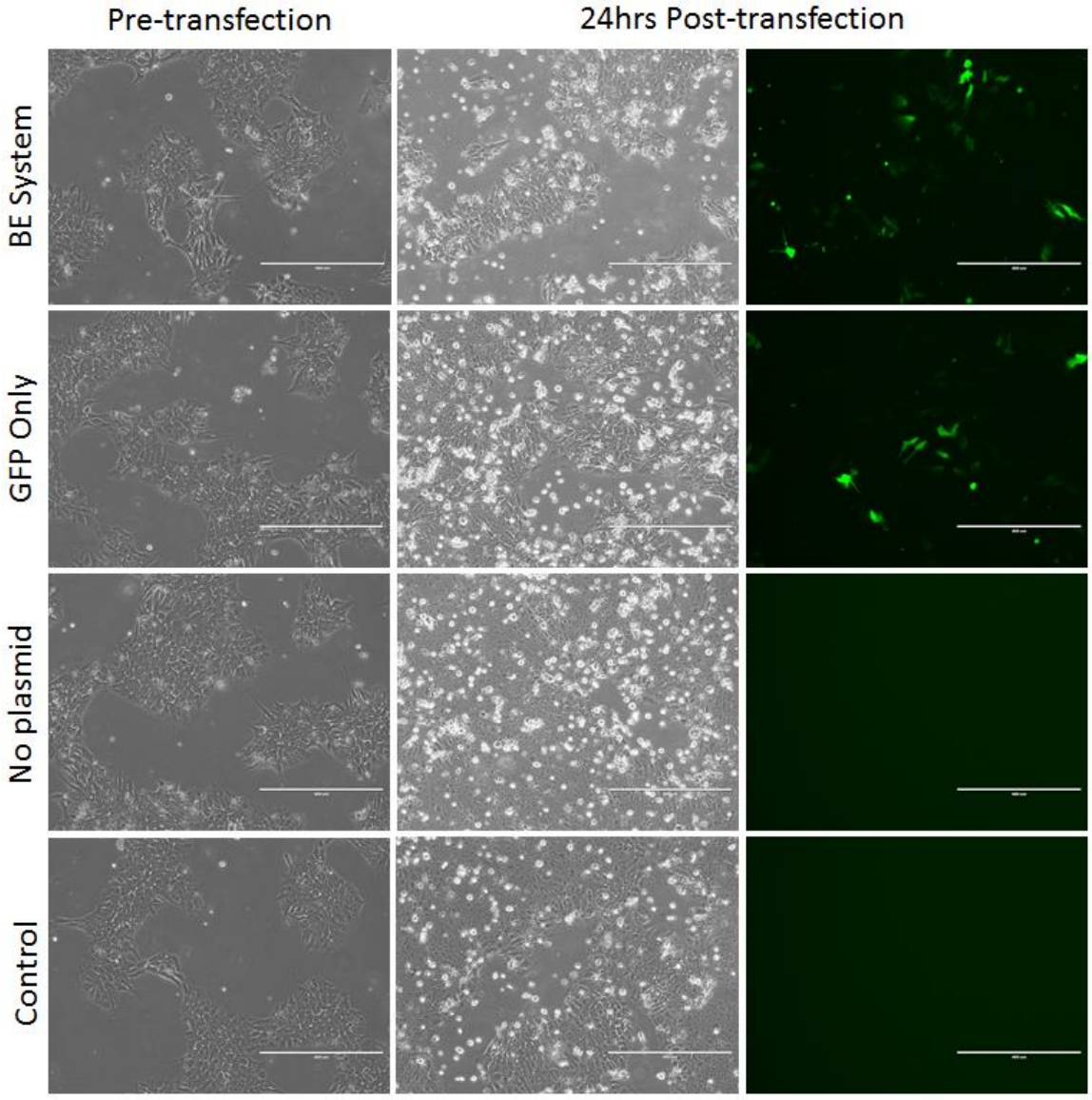

Figure 14. Microscopy images of iPS-160 cells immediately preceding and 24hrs following transfection with Transfex reagent. 10x magnification.

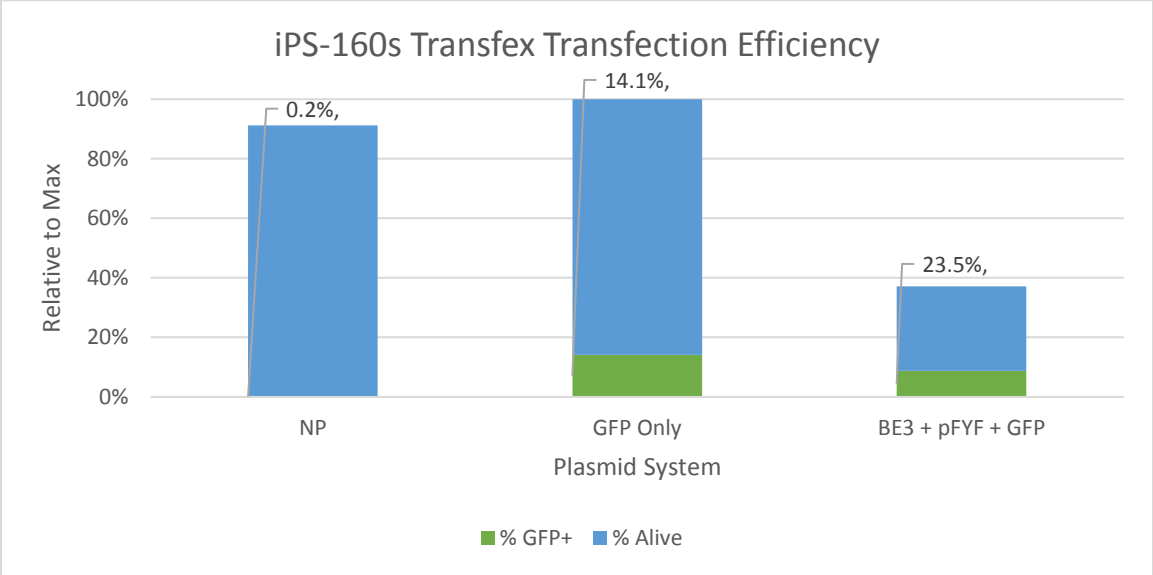

Figure 15. Flow cytometry results characterizing yield of live and GFP-positive iPS-160 cells 24hrs post-transfection. Total yield was normalized to the maximum yield observed. Data labels indicate the percent of live cells that were found to be GFP-positive for each sample.

4.5. Preparation for Clonal Line Generation

The control iPS-384 cells isolated through FACS failed to proliferate after incubating over 14 days. Potential avenues for improving future yield include increased cell numbers sorted into each well, modulated concentrations of RockI, and the use of mouse embryonic fibroblasts to encourage proliferation. Alternate isolation methods could also be explored, such as mechanical separation of colonies.

Sanger sequencing for the ten predicted off-target sites resulted in confirmed baseline sequencing for six of the ten sites. Three of the five sites identified for the 112 off-target analysis did not return sequencing data that matched the expected sequence or even the intended chromosome, indicating that the primers used require redesign. One of the five sites identified for the 158 off-target analysis simply did not yield sufficient product for a confident analysis.

Off-Target Site	Location	Mismatches	Expected Sequence	Actual Sequence
112-1	chr4:24913056-24913078	2	GAGG[GGCGCGGCCGCC]	Not found
112-2	chr11:46380711-46380729	1	GACGAGCGCGGCCGCC	GACGAGCGCGGCCGCC
112-3	chr16:67155478-67155500	2	GGCG[AGCGCGGCCGCC]	Not found
112-4	chr14:102517000-102517018	1	GACG[TCCGCGGCCGCC]	Not found
112-5	chr1:23743300-23743318	2	GGAG[TGCGCGGCCGCC]	GGAGTGCGCGGCCGCC
158-1	chr11:944367	3	CCTGcTACAaTGCCAGGCaCTT	CCTGCTACAATGCCAGGCACTT
158-2	chr19:42356330	3	CCTaGgACACTGCCAGcCGCTT	CCTAGGACACTGCCAGCCGCTT
158-3	chr19:42401325-42401343	1	CTCC[TGGCAGTGTACC]	CTCCTGGCAGTGTACC
158-4	chr1:148987471-148987489	1	CGCA[TGGCAGTGTACC]	Not found
158-5	chr18:54823613-54823631	2	TACC[TGGCAGTGTACC]	TACCTGGCAGTGTACC

Table 3 Results from Sanger sequence analysis determining baseline sequence of potential off-target sites in the iPS-160 cell line.

5. SUMMARY & DISCUSSION

The use of the HEK293 cell line provided verification of the delivery method, CRISPR/Cas9 system, and targeting sequence with respect to direct base conversion at the 158 amino acid site in that cell line. This verification demonstrated the functionality of the modified base editor in effecting an ε3 to ε2 conversion at the 158 site as reported by Komor et al. (2016). The same electroporation methodology did not effect the base conversion in hiPSC lines, however, use of a lipofection technique with comparable plasmid concentration did result in confirmed base editing at the 158 site in the iPS-160

cell line. Successful base editing at the 112 site was not able to be observed given a sequencing artifact at the targeted nucleotide.

With the promising establishment of the base modification and clonal line isolation methods, future work will first and foremost involve the implementation of these established methods to generate the isogenic ε3/3 clonal lines from the current ε4/4 lines. The prevalence of off-target effects will be assessed for the 112-targeting modification system. The methodology would then be repeated to generate the corresponding ε2/2 clonal lines with off-target effect assessment performed for the 158-targeting modification system. Differentiation from pluripotency to neuronal cell types, specifically the astrocytes responsible for ApoE production in vivo, would provide opportunity for a variety of analyses to assess differences in cell function that could only be attributed to the differences in the APOE variant. These potential analyses include assays for cholesterol transport levels, extracellular Aβ accumulation, expression of phosphorylated tau, and cell survival. Any differences observed would help to clarify the role of ApoE and the effects of the isoforms on the pathogenesis of Alzheimer's disease. Such an outcome would be a new model for in vitro study of Alzheimer's disease and could provide the missing human component for drug screening currently thwarted by rodent models. Given success of this system to model APOE variants, it is also possible that this genetic modification model in hiPSC lines could be applied to other genetic targets identified as being correlated with varied risk of developing AD or other diseases.

REFERENCES

[1] "2017 Alzheimer’s Disease Facts and Figures," *Alzheimer's and Dementia*, no. 13, pp. 325-373, 2017.

[2] "Changing the Trajectory of Alzheimer's Disease," Alzheimer's Association, 2015.

[3] "Estimates of Funding for Various Research, Condition, and Disease Categories (RCDC)," U.S. Department of Health & Human Services, 10 February 2016. [Online]. [Accessed 2017].

[4] "Health, United States, 2015: With Special Feature on Racial and Ethnic Health Disparities," National Center for Health Statistics, Hyattsville, MD, 2016.

[5] M. M. ,. B. D. S. Eric Karran, "The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics," *Nature Reviews Drug Discovery,* vol. 10, no. 9, pp. 698-712, September 2011.

[6] S. L. C. a. R. Vassar, "The Alzheimer's disease β-secretase enzyme, BACE1," *Molecular Neurodegeneration,* p. 22, 15 November 2007.

[7] P. D. X. J. G. Q. Shirwany NA, "The amyloid beta ion channel hypothesis of Alzheimer's disease," *Neurophysiatric Disease and Treatment,* vol. 4, no. 5, pp. 597-612, October 2007.

[8] H. U. W. S. M. J. T. J. B. M. M. M. R. F. D. A. S. L. W. D. K. M. M. U. Hick M, "Acute function of secreted amyloid precursor protein fragment APPsα in synaptic plasticity," *Acta Neuropathologica,* vol. 129, no. 1, pp. 21-37, January 2015.

[9] J. L. B. P. C. G. L. H. Y. W.-Y. H. M. J. A. A. P. W. R. J. B. C. A. H. K. A. W. N. L. E. A. H. C. M. C. a. A. D. R. M. A. Pericak-Vance, "Linkage studies in familial Alzheimer disease: Evidence for chromosome 19 linkage," *American Journal of Human Genetics,* vol. 48, no. 6, pp. 1034-1050, June 1991.

[10] C. F. a. K. Garai, "Concerning the structure of apoE," *Protein Science,* 1 Oct 2013.

[11] C. A. P.-L. K. H. W. Danny M. Hatters, "Apolipoprotein E structure: insights into function," *Trends in Biomedical Sciences,* vol. 31, no. 8, pp. 445-454, August 2006.

[12] A. M. S. W. J. S. B. J. C. C. M. H. S. H. J. M. A. P.-V. D. G. a. A. D. R. D E Schmechel, "Increased amyloid beta-peptide deposition in cerebral cortex as a consequence of apolipoprotein E genotype in late-onset Alzheimer disease," *Proceedings of the National Academy of Sciences,* vol. 90, no. 20, 15 October 1993.

[13] J. M. C. K. G. Y. W. H. J. A. S. G. B. C. F. a. D. M. H. Philip B. Verghese, "ApoE influences amyloid-β (Aβ) clearance despite minimal apoE/Aβ association in physiological conditions," *Proceedings of the National Academy of Sciences,* vol. 110, no. 19, 25 April 2013.

[14] C. K. T. X. H. &. B. G. Liu, "Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy," *Nature Reviews Neurology,* vol. 9, no. 4, pp. 184-184, 2013.

[15] P. Walter A. Kukull, P. Roger Higdon, M. James D. Bowen and e. al, "Dementia and Alzheimer Disease Incidence: A Prospective Cohort Study," *Archives of Neurology,* vol. 59, no. 11, pp. 1737-1746, 2002.

[16] R. F. d. B. a. M. A. Ikram, "Cardiovascular risk factors and future risk of Alzheimer's disease," *BMC Medicine,* vol. 12, 11 November 2014.

[17] V. E. L. K. B. D. C. K. N. T. K. M. M. C. Yaffe K, "Posttraumatic Stress Disorder and Risk of Dementia Among US Veterans," *Archives General Psychiatry,* vol. 67, no. 6, pp. 608-613, 2010.

[18] Y. L. X. L. S. Z. J. Z. X. Z. G. T. Yanjun Li, "Head Injury as a Risk Factor for Dementia and Alzheimer's Disease: A Systematic Review and Meta-Analysis of 32 Observational Studies," *PLoS One,* vol. 12, no. 1, 9 January 2017.

[19] H. A. B. A. A. G. A. T. A. B. Z. a. Y. W. May A BeydounEmail author, "Epidemiologic studies of modifiable factors associated with cognition and dementia: systematic review and meta-analysis," *BMC Public Health,* vol. 14, p. 643, 24 June 2014.

[20] S. P. Valenzuela MJ, "Brain reserve and dementia: a systematic review," *Psycological Medicine,* vol. 36, no. 4, pp. 441-454, April 2006.

[21] "Animal Models of Alzheimer Disease," *Cold Spring Harbor Perspectives in Medicine,* November 2012.

[22] M. C.-T. a. A. García-Osta, "Current Animal Models of Alzheimer's Disease: Challenges in Translational Research," *Frontiers in Neurology,* no. 5, p. 182, 24 September 2014.

[23] F. M. L. a. K. N. Green, "Animal Models of Alzheimer's Disease," *Cold Spring Harbor: Perspectives in Medicine,* vol. 2, no. 11, 1 November 2012.

[24] R. F. a. A. Cedazo-Minguez, "Successful therapies for Alzheimer's disease: why so many in animal models and none in humans?," *Frontiers in Pharmacology,* vol. 5, p. 146, 25 June 2014.

[25] Y. S. Takahashi K, "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors," *Cell,* vol. 126, no. 4, pp. 663-76, 25 August 2006.

[26] N. M. a. M. S. Rao, "A Review of the Methods for Human iPSC Derivation," *Methods in Molecular Biology,* vol. 997, pp. 23-33, 2013.

[27] M. N. a. M. Otsu, "Development of Sendai Virus Vectors and their Potential Applications in Gene Therapy and Regenerative Medicine," *Current Gene Therapy,* vol. 12, no. 5, pp. 410-416, October 2012.

[28] S. C. C. W. v. d. B. M. H. C. A. R. C. D. D. S. D. W.-v. O. A. A. W. C. R. B. A. O. V. C. F. C. L. M. Richard P. Davis, "Cardiomyocytes Derived From Pluripotent Stem Cells Recapitulate Electrophysiological Characteristics of an Overlap Syndrome of Cardiac Sodium Channel Disease," *Circulation,* vol. 125, no. 25, pp. 3079-3091, 25 June 2012.

[29] E. P. P. H. K. S. M. C. M. J. T. C. A. F. Y. M. G. J. M. F. S. A. V. V. T. M. S. &. L. S. Gabsang Lee, "Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs," *Nature,* vol. 461, pp. 402-406, 17 September 2009.

[30] C. R. M. H. C. R. P. S. D. G. C. T. S. L. N. P. B. D. M. W. D. J. S. T. L. Young-Pearse, "The familial Alzheimer's disease APPV717I mutation alters APP processing and Tau expression in iPSC-derived neurons," *Human Molecular Genetics,* vol. 23, no. 13, pp. 3523-3536, 12 February 2014.

[31] Y. B. K. M. S. P. J. A. Z. &. D. R. L. Alexis C. Komor, "Programmable editing of a target base in genomic," *Nature,* vol. 533, pp. 420-242, 2016.

[32] F. Z. Yi-ye Zhou, "Integration-free Methods for Generating Induced Pluripotent Stem Cells," *Genomics, Proteomics & Bioinformatics,* vol. 11, no. 5, pp. 284-287, October 2013.

[33] B. M. A. G. J. V. S. P. M. H. L. A. G. E. a. E. G. S. Sara E. Howden, "A Cas9 Variant for Efficient Generation of Indel-Free Knockin or Gene-," *Stem Cell Reports,* vol. 7, no. 3, pp. 508-517, 13 September 2016.

[34] T. T. ,. M. d. S. K. J. W. ,. J. L. M. Manuel Stemmer, "CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool," *PLOS One,* 24 April 2015.