Generation of Knockout Human iPSCs to Investigate Genes Associated with Telomere Length

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Noor Farhan Shamkhi BSc, MSc

Department of Cardiovascular Sciences

University of Leicester

2017

Generation of knockout human iPSCs to investigate genes associated with telomere length

Noor Farhan Shamkhi

Abstract

Telomeres are repetitive sequences located at the ends of human chromosomes. During DNA replication, DNA polymerase is unable to fully replicate telomeric DNA causing a progressive reduction in telomere length with each cell division. Consequently, telomeres shorten with age and when telomere length reaches a critical length the cell becomes dysfunctional or senescent. Increasing numbers of senescent cells results in reduced organ function and telomere length has been associated with age-associated diseases including cardiovascular disease and cancer. Telomere length is highly heritable and genome-wide association studies have identified several loci that associate with telomere length. Interestingly, two of these loci do not contain genes with known roles in telomere maintenance.

In this study, CRISPR/Cas9 genome editing was used to knockout telomere length associated genes in human induced pluripotent stem cells (hiPSCs) with the ultimate aim of definitively linking a novel telomere length associated gene to telomere maintenance. CRISPR/Cas9 was used to create hiPSCs with no telomerase activity by knocking out *TERT*, which encodes the catalytic subunit of telomerase. Telomere length was maintained in control iPSCs, but reduced by approximately 1% per day during extended culture of *TERT* knockout hiPSCs. Flow cytometric analysis of mutant hiPSCs revealed that they continued to express markers of pluripotency but had increased expression of a differentiation marker. Next, CRISPR/Cas9 was used to generate mutations in the candidate telomere maintenance genes *ACYP2* and *TSPYL6*, however, no difference in telomere length was observed after extended culture of hiPSCs carrying mutations in either gene.

In conclusion, CRISPR/Cas9 genome editing was successfully used to generate mutant hiPSCs for *TERT*, which resulted in telomere shortening and for two candidate telomere maintenance genes, which had no effect on telomere length. Further analysis will be required to determine which gene mediates the association with telomere length at the *ACYP2* locus.

ACKNOWLEDGEMENTS

I would first like to thank my thesis advisors Dr. Tom Webb, Dr. Veryan Codd and Professor Sir Nilesh Samani. The door to Dr. Webb's office was always open whenever I ran into a trouble spot or had a question about my research or writing. He consistently allowed this paper to be my own work, but steered me in the right direction whenever he thought I needed it. Very special gratitude goes out to all at the Higher Committee for Education Development in Iraq (HCEDiraq) for helping and providing the funding for this Ph.D. study.

I would like to acknowledge all the members of the Department of Cardiovascular Sciences who have helped and guided me throughout my PhD, and have made the department a nice place to be. In particular, I am grateful to Dr. Pete Jones, Mr. Mike Kaiser, Dr. Peng Gong and Maryam Ghaderi Najafabadi for help with experiments and to Dr. Sameer Kurmani and Dr. Rob Turnbull for all their help with the flow cytometer. A huge thank for all my friends specially Dr. Wadhah Mahbuba for their advice.

Finally, I must express my very profound gratitude to my parents and to my sisters, Asmaa, Dianaa, Asraa and Haneen for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Noor Shamkhi

Table of Contents

List of Tables

List of Figures

The list of abbreviations

36B4: acidic ribosomal phosphoprotein PO

A: Adenine

A: Alanine

ACYP2: Acyl phosphatase-2

ALS: amyotrophic lateral sclerosis

ALT: Alternative lengthening of telomere

Antp^{+}: Antennapedia

ASB3: ankyrin repeat and SOCS box containing 3

BSA: bovine serum albumin

C: Cysteine

C: Cytosine

C2: complement C2

CAD: Coronary artery disease

CD9: cell surface antigen

cDNA: complementary Deoxyribonucleic acid

CHAC2: ChaC cation transport regulator homolog 2

CHD: Coronary heart disease

CHS: Cardiovascular health study

CGH: genomic hybridization array

***c-Myc*:** myelocytomatosis oncogene

CNVs: copy number variations

CRISPR: Clustered regularly interspaced short palindromic repeat

crRNA: CRISPR-derived RNAs

CST: CTC1, STN1 and **TEN1** complex

CTC1: Conserved telomere maintenance componenet1

CTE: C-terminal domain or C-terminal extension domain

CVD: Cardiovascular disease

D: Aspartic acid

DAPI: 4′, 6-diamidino-2-phenylindole

DDR: DNA damage response

DKC1: Gene encodes the dyskerin protein

DMEM: Dulbecco's Modified Engle's Medium

DNA: Deoxyribonucleic acid

Dnmt3a: DNA methyltransferase 3 alpha

Dnmt3b: DNA methyltransferase 3 beta

DSBs: Double strand breaks

E: Glutamic acid

ECM: extracellular matrix

EDTA: Ethylenediaminetetraacetic acid

EML6: echinoderm microtubule associated protein like 6

ERLEC1: endoplasmic reticulum lectin 1

ESCs: Embryonic stem cells

F: Phenylalanine

FBS: Fetal bovine serum

FCS: Foetal calf serum

FISH: fluorescent *in situ* hybridization

FITC: fluorescent isothiocyanate

FS: forward scatter

G: Glycine

G: Guanine

GAR1: GAR1 ribonucleoprotein

GDF3: growth differentiation factor 3

gDNA: genomic Deoxyribonucleic acid

GFP: green fluorescent protein

GTEx: Genotype-Tissue Expression project

GWAS: Genome-wide association studies

H: Histidine

***HBB*:** hemoglobin subunit beta

HDR: homology-directed repair

HEK293: human embryonic kidney 293 cell line

hESCs: Human embryonic stem cells

hiPSCs: Human induced pluripotent stem cells

HLA: surface protein antigen

hPSCs: Human pluripotent stem cells

HR: Homologous recombination

I: Isoleucine

IDLVs: integrase-defective lentiviral vectors

iPSCs: Induced pluripotent stem cells

K: Lysine

K562: human immortalised myelogenous leukemia line

Kbp: Kilobase pairs

Klf4: Kruppel-like factor 4

L: Leucine

M: Methionine

MAD2: mitotic arrest deficient-like 2

MAD2L2: mitotic arrest deficient 2 like 2

MEF: mouse embryonic fibroblast

MEF-CM: mouse embryonic fibroblast serum-conditioned media

MiPSCs: mouse induced pluripotent stem cells

N: Asparagine

NAF1: nuclear assembly factor 1 ribonucleoprotein

NANOG: Nanog homeobox

NAP: Nucleosome assembly protein

NCBI: National Centre for Biotechnology Information

NHEJ: Error-prone non homologous end joining

NHP2: NHP2 ribonucleoprotein

NOP10: NOP10 ribonucleoprotein

NRT: No reverse transcriptase

NTC: No template control

OBFC1: oligonucleotide/oligosaccharide binding fold containing 1

***Oct3/4*:** organic cation/carnitine transporter 3/4

P: Proline

PAM: Protospacer adjacent motif

PBS: Phosphate buffer saline

PCR: Polymerase chain reaction

PE: phycoerythrin

PFA: Paraformaldehyde

PMSB4: proteasome subunit beta 4

POT1: Protein protection of telomeres 1

POU5F1: POU class 5 homeobox 1

PSCs: Pluripotent stem cells

PSME4: proteasome activator subunit 4

Q: Glutamine

qPCR: quantities Polymerase chain reaction

qRT-PCR: quantities real time-Polymerase chain reaction

R: Arginine

Rap1: Repressor activator protein 1

RAP1: Repressor/activator site-binding protein

REX1: RNA exonuclease 1 homolog

***RPL23AP32*:** ribosomal protein L23a pseudogene 32

RT: reverse transcriptase domain

RTEL1: Regulator of telomere elongation helicase 1

RT-PCR: Real-time Polymerase chain reaction

RVDs: Repeat variable diresidues

S: Serine

SCNT: Somatic cell nuclear transfer

SET: protein domain

sgRNA or gRNA: Single-guide-RNA

siRNA: Small interfering RNA

SNPs: Single-nucleotide polymorphisms

Sox2: SRY-box 2

SPTBN1: spectrin beta, non-erythrocytic 1

SS: side scatter

SSEA: stage-specific embryonic antigen

STN1: Suppressor of cdc thirteen homolog

Suv39h2: suppressor of variegation 3-9 homolog 2

T: Threonine

T: Thymine

TALENs: Transcription Activator- like Effector Nucleases

TCAB1/ WRAP53: WD repeat containing antisense to TP53

TEN: telomere essential N-terminal domain

TEN1: Telomere length regulation protein TEN1 homolog

TERC: Telomerase RNA Component

TERT: Telomerase reverse transcriptase

Thy1(CD90): Thy-1 cell surface antigen

TIN2: TRF1-interacting nuclear factor 2

TPP1: telomerase processivity factor

TRA: Trafalgar antigen

tracrRNA: Trans-activating crRNA

TRBD: telomerase RNA binding domain

TRF1: Telomeric repeat binding factor 1

TRF2: Telomeric repeat binding factor 2

TSCE: telomere sister chromatid exchange

TSPYL6: TSPY like 2

TTSN: *TSPY/TSPYL/SET/NAP-1*

UCSC: University of California Santa Cruz genome browser

V: Valine

***VPS34/ PIK3C*:** phosphatidylinositol 3-kinase catalytic subunit type 3

W: Tryptophan

WGS: whole genome sequencing

WOSCOP: West of Scotland Primary Prevention Study

WT: Wild type

WTCCC: Welcome Trust Case-Control Consortium

Y: Tyrosine

ZFNs: Zing-finger nucleases

ZNF208: zinc finger protein 208

Zscan4: zinc finger and SCAN domain containing 4, pseudoge

Chapter 1

Introduction

1.1 Introduction

Telomeres are nucleoprotein structures located at the end of chromosomes that function to maintain the integrity and stability of genomic DNA. These structures naturally shorten with each cell division causing telomere length to reduce with age. Telomere length varies between individuals, is highly heritable and is associated with age-related disease such as coronary artery disease (CAD) and cancer. Recent genome-wide association studies (GWAS) have identified several loci that associate with telomere length and implicated genes with no known role in telomere maintenance. This study aims to identify the genes involved and their function in telomere biology.

1.1.1 Telomeres

Telomeres are specialized structures located at the extreme end of the eukaryotic chromosomes and composed of repetitive DNA and various interacting proteins. The telomeric nucleotide sequence comprises noncoding double-stranded Guanine (G) rich repeats which extend several thousand base pairs (bp) and end in a 3'single stranded overhang termed a G-tail (Serrano and Andres 2004) **(Figure 1.1)**. The repeat sequence varies in length from a few hundred base pairs in yeast to several kilo base pairs (kbp) in vertebrates (Oeseburg et al. 2010). The mammalian telomeric repeat sequence is TTAGGG and extends to 10-15kbp of double-stranded DNA with a G-tail of 150-200bp (Meyne et al. 1989, Wright et al. 1997, Opresko and Shay 2016). The telomere is an important structure because it preserves genomic stability and integrity by preventing chromosomal fusion at the ends of the chromosomes. A protein complex called shelterin masks the ends of telomeres and prevents them being recognized as double strand breaks (DSBs) by the DNA repair machinery (di Fagagna et al. 2003, Szilard and Durocher 2006).

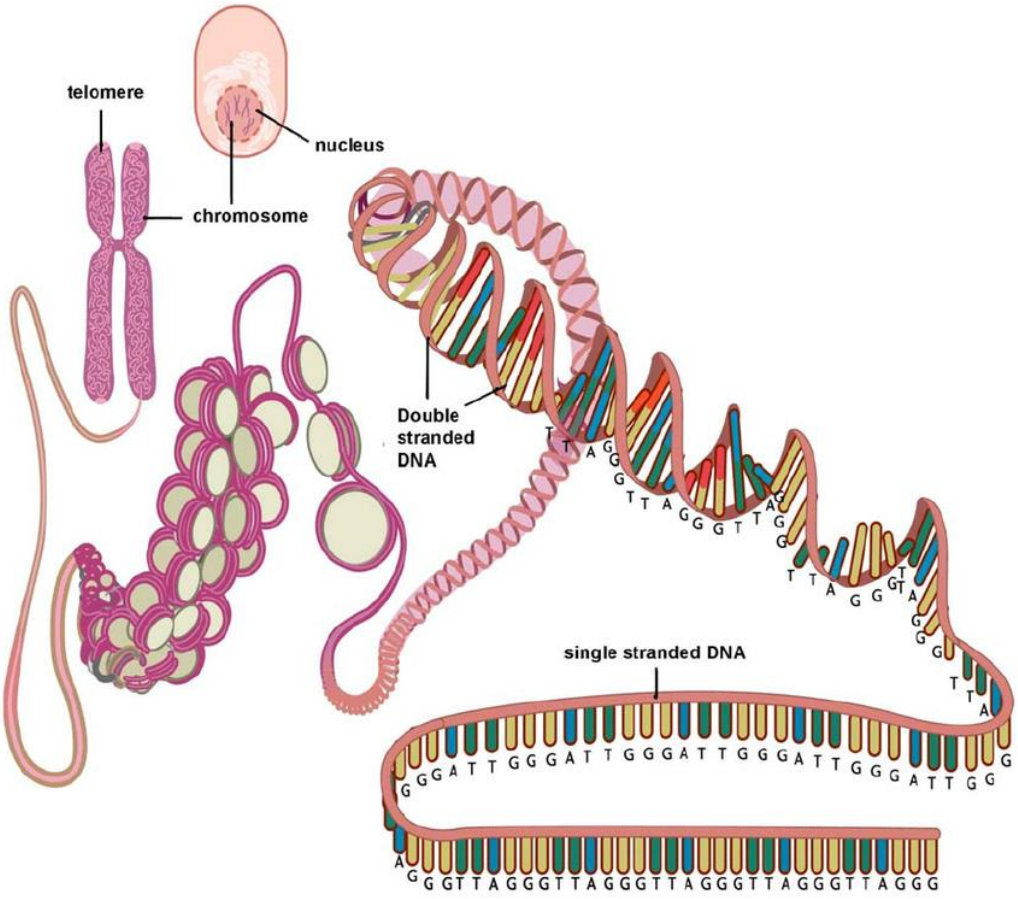


Figure 1.1: Telomere Structure. Telomere is positioned at the end of chromosomes and consists of TTAGGG repeats. Taken from Huzen *et al*. (2008).

During DNA replication, DNA polymerase is not able to replicate the 3' end of the lagging strand (Greider and Blackburn 1987), this results in a shortening in telomere length with each mitotic cycle and is termed "the end replication problem" (Olovnikov 1971). When telomere length reaches a critical limit, the cell is no longer able to divide (Hayflick and Moorhead 1961) and the cells undergo replicative senescence (Dimri et al. 1995, Oeseburg et al. 2009, Shelton et al. 1999). The number of times a cell is able to divide is therefore determined by its telomere length; a phenomenon called the "Hayflick limit". It is thought that the accumulation of senescent cells leads to age-related tissue deterioration and disease features (Campisi 2001, Campisi et al. 2001, Stewart and Weinberg 2006). For this reason, telomeres serve as a "molecular clock of biological aging" (Shay and Wright 1996).

1.1.1.2 Telomere associated proteins

Telomerase is a protein/RNA complex that maintains telomere length, however, it is not expressed in all cell types. During embryonic development telomerase is highly expressed, particularly at the blastocyst stage of embryogenesis, while it is suppressed after birth in most somatic cells (Wright et al. 1996). The expression of telomerase is high in proliferative cells such as stem cells, progenitor cells, lymphocytes and skin keratinocytes (Flores et al. 2006) as well as male germ cells (Kim et al. 1994, Shay and Bacchetti 1997, Wright et al. 1996).

Telomerase has two primary components, the non-coding RNA *TERC* and the reverse transcriptase TERT (Harrington et al. 1997, Kilian et al. 1997, Lingner et al. 1997, Meyerson et al. 1997) (**Figure 1.2**). *TERC* is a 451 nucleotide non-coding RNA, which contains the 11bp sequence CAAUCCCAAUC that acts as the template sequence for the addition of TTAGGG DNA repeats to the G-Tail (Morin 1989). The TERT protein contains four main functional domains, the telomere essential N-terminal domain (TEN), the reverse transcriptase domain (RT), the telomerase RNA binding domain (TRBD) and the C-terminal domain or C-terminal extension domain (CTE) (Zvereva et al. 2010, Podlevsky and Chen 2012, Sandin and Rhodes 2014). The TEN domain binds RNA and telomeric DNA and is responsible for addition of telomere repeats (Jacobs et al. 2006). The RT domain is the central catalytic domain and consists of seven motifs, five of which

form the catalytic site while the other two bind incoming nucleotides (Lue 2004). The CTE is essential for telomerase activity (Banik et al. 2002). Other components of telomerase include Dyskerin, TCAB1, NOP10, NHP2, pontin/reptin and GAR1 proteins, which are involved in telomerase biogenesis, stability and localisation (Fu and Collines 2007) **(Figure 1.2)**.

Telomere lengthening is a step-wise process **(Figure 1.3)**. During elongation, the 11bp RNA template hybridizes with the extreme 3' telomeric repeat and nucleotides are then incorporated to complete the gap at the end of the RNA template. Consequently, a complete telomeric repeat is synthesized and assembled within the catalytic domain of TERT. Finally, the newly synthesized strand is translocated in a 3' to 5' direction producing a new gap and allowing the cycle of synthesis to be repeated (Gomez et al. 2012).

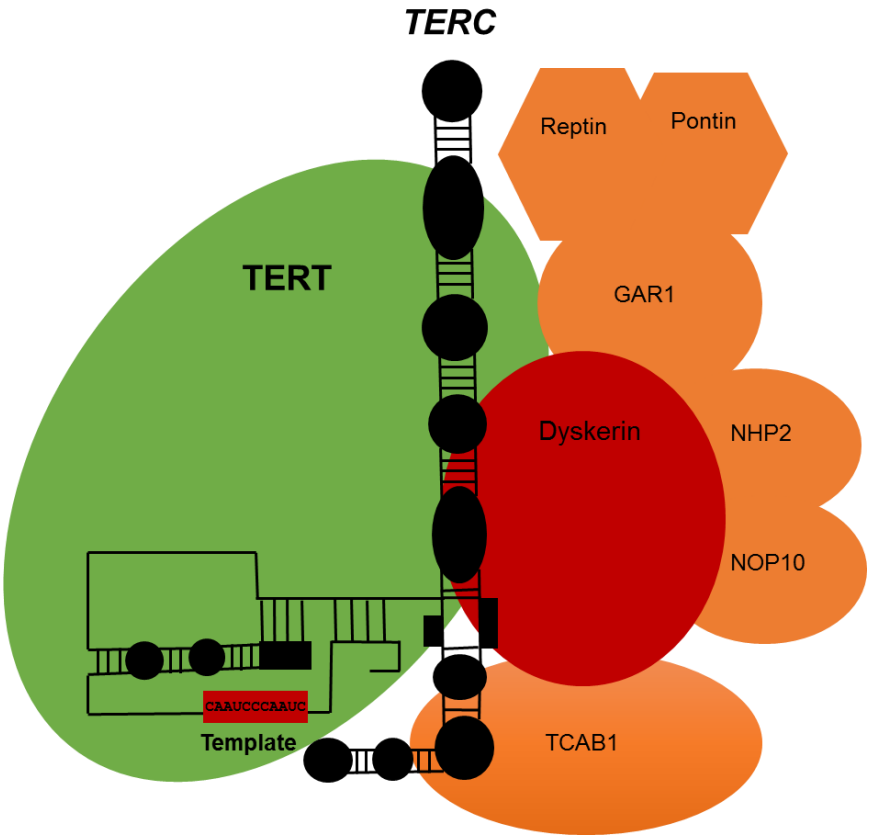


Figure 1.2: Schematic illustration of telomerase and its associated proteins.
Adapted from Gomez *et al*. (2012).

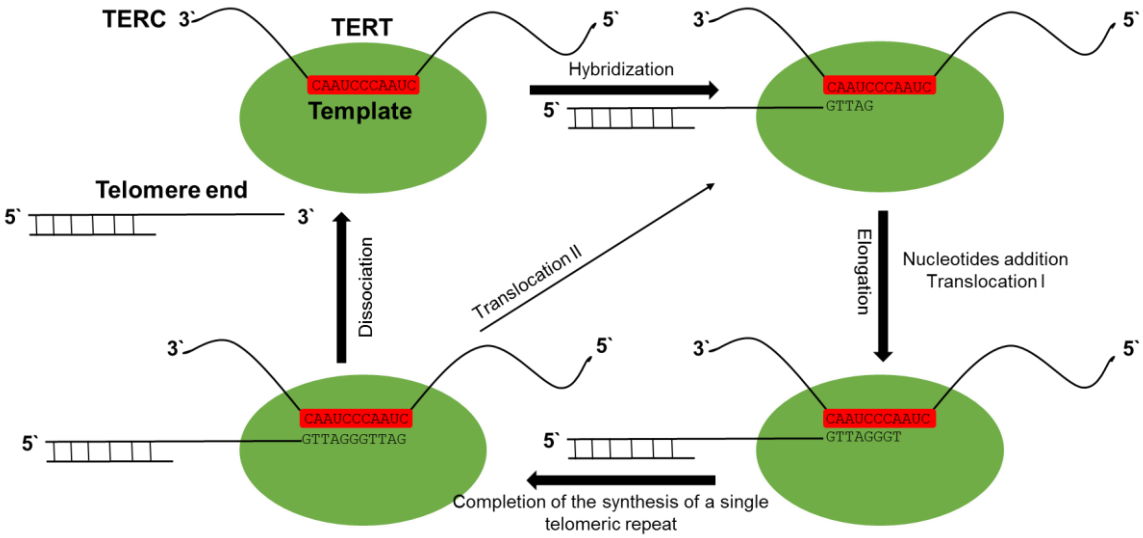


Figure 1.3: Schematic representation of telomere lengthening by telomerase. The 3' end of the telomeric DNA hybridizes with the *TERC* template sequence. The gap at the end of the template is completed by the addition of nucleotides producing a complete telomeric repeat. The newly synthesized repeat is released in a 3' to 5' direction and the cycle is either repeated or telomerase dissociates from the telomere. Adapted from Gomez *et al*. (2012).

Mutations in the *TERT* promoter reactivate telomerase expression in several types of human cancer (Vinagre et al. 2013). Expression of telomerase leads to the immortalization of cells, allowing them to escape replicative senescence (Bodnar et al. 1998, Ouellette et al. 2000, Ramirez et al. 2001, Yang et al. 1999). In addition, in the absence of telomerase proliferative human cell lines and some tumors elongate their telomeres using a mechanism called alternative lengthening of telomeres (ALT) (Bryan et al. 1995, Bryan et al. 1997). This process involves homologous recombination (HR) with DNA sequences being copied from telomere to telomere (Dunham et al. 2000) .

The shelterin protein complex consists of six core proteins and is required for the protection, stabilization and regulation of telomeres **(Figure 1.4)**. The shelterin subunits Telomeric repeat binding factor 1 and 2 (TRF1, TRF2) directly bind TTAGGG repeats and Protein protection of telomeres 1 (POT1) recognizes binds the G-tail **(Figure 1.4 A)**. TRF1 and TRF2 interact with POT1 via TRF1-interacting nuclear factor 2 (TIN2), telomerase processivity factor (TPP1) and Repressor/activator site-binding protein (RAP1) form the shelterin complex (O'Connor et al. 2006, Opresko and Shay 2016). Together with the T-loop structure, the shelterin complex prevents the G-tail being recognized as a double-strand break (DSB) and triggering the DNA damage response (DDR) (di Fagagna et al. 2003, Opresko and Shay 2016). Shelterin causes the telomere to fold in on itself to form a large loop structure, termed a telomeric loop (T-loop) **(Figure 1.4 B)** (Griffith et al. 1999, Doksani et al. 2013, Gomez et al. 2012). The G-tail inserting is inserted between double-stranded telomeric DNA and disrupts the double–helix to form the displacement loop (D-loop).

The shelterin complex is also involved in the regulation of telomere length. When telomeres are too long, more shelterin complexes bind and act to inactivate telomerase (de Lange 2005). When telomeres are short the shelterin complex becomes more relaxed allowing telomerase to maintain the length of telomeres.

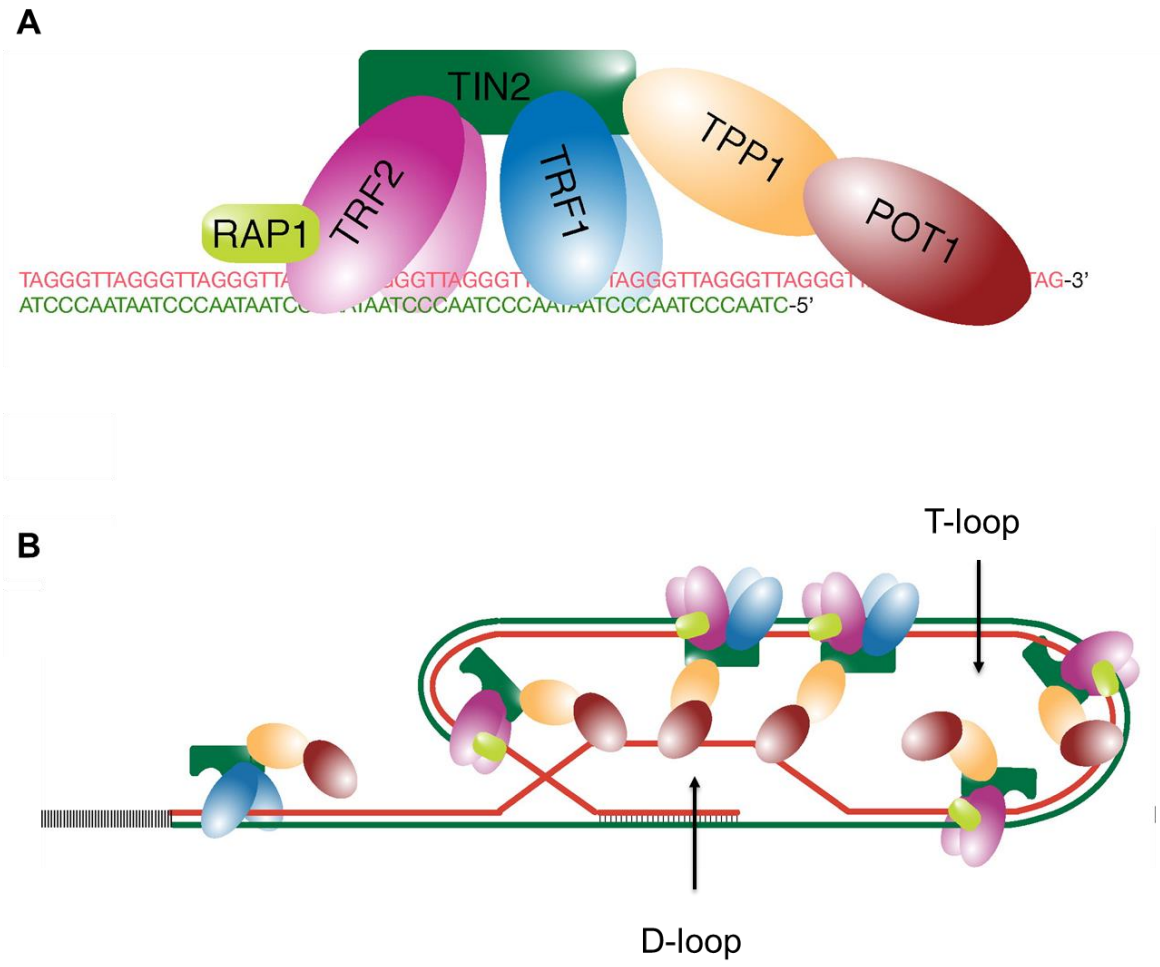


Figure 1.4: Schematic representation of the telomere shelterin complex. Proteins: TRF1, TRF2, POT1, Rap1, TIN2, and TPP1 are associated with telomere that leads to stability of chromosome. This figure is taken from Denchi (2009).

Another complex called CST, which consists of three proteins, Conserved telomere maintenance componenet1 (CTC1), Suppressor of cdc thirteen homolog (STN1) and Telomere length regulation protein TEN1 homolog (TEN1), is involved in telomere protection and DNA metabolism (Miyake et al. 2009, Casteel et al. 2009, Surovtseva et al. 2009, Nakaoka et al. 2012). CST inhibits telomerase activity by binding to POT1-TPP1, stabilizing the shelterin complex and preventing binding of telomerase. (Wang et al. 2007, Latrick, Cech 2010, Chen et al. 2012).

1.1.1.3 Variation in telomere length

Telomere length varies between individuals and declines with age. Telomere length also differs between different organs of the same individual. Telomere length is similar between organs in a newborn but decreases at different rates due to variation in tissue renewal times leading to variable rates of telomere length attrition (Okuda et al 2002). For instance, the renewal time of gastrointestinal tissue is rapid, while renewal of hepatic and renal tissue is slow. Although the rate of attrition is different between the various tissues, there is a correlation within an individual meaning that an individual with shorter blood telomere length would also have shorter telomeres in other organs (Takubo et al 2002, Wilson et al. 2008, Daniali et al. 2013) **(Figure 1.5)**. Telomere length from easily accessible cell-types can therefore be used as a marker for the telomere length of an individual.

A number of different factors affect telomere length, but a significant proportion is due to genetic factors. Heritability of telomere length has been reported to be between 34 and 82% (Slagboom et al. 1994, Bischoff et al. 2005, Vasa-Nicotera et al. 2005). The largest study to date involved a meta-analysis of almost 20,000 individuals and estimated telomere length heritability at 70% with strong agreement between the contributing studies (Broer et al 2013). The same study also identified a greater maternal than paternal influence on telomere length and confirmed a positive association with paternal age (Unryn et al. 2005, De Meyer et al. 2007, Kimura et al. 2008). The cause of the maternal effect is not known, but might include X-linked effects, mitochondrial influences or imprinting (Broer et al 2013).

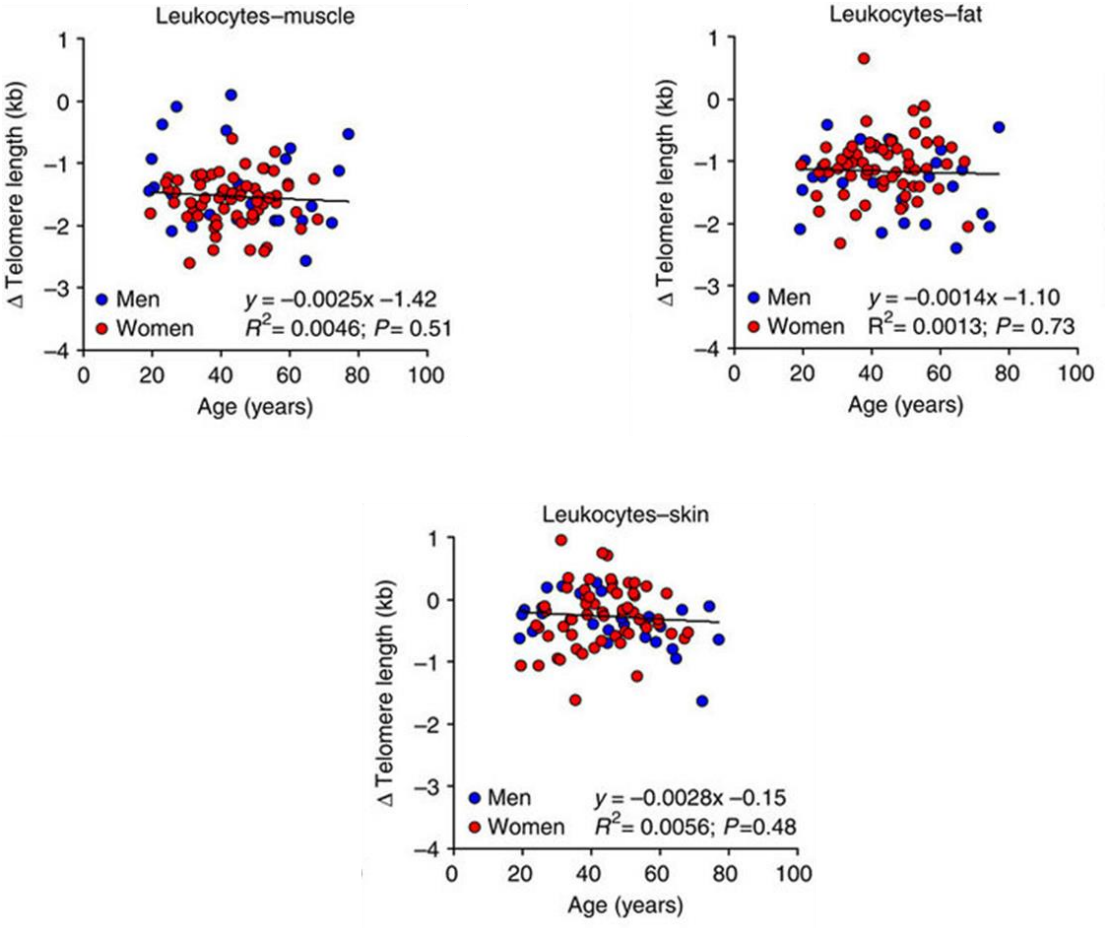


Figure 1.5: Telomere length attrition in different tissues within the same individuals during human lifetime. This figure is taken from Daniali *et al.* (2013).

Telomere length also varies between males and females, with telomere length longer in women than in men, possibly a result of stimulation of telomerase by the hormone estrogen (Okuda et al. 2002, Kyo et al. 1999). Moreover, Hunt *et al*. (2008) revealed variation in telomere length between different ethnicities with African Americans having longer telomeres than Caucasians.

Various lifestyle and environmental factors including, stress, obesity, smoking and diet have been reported to affect telomere length. Previous studies have shown that shorter telomere length is associated with risk factors such as stress, obesity, smoking and diet (Epel et al. 2004, Song et al. 2010, McGrath et al. 2007, Morla et al. 2006, Cassidy et al 2010). However, the majority of these had relatively small sample size and require investigation in larger datasets.

1.1.1.4 Telomere length and disease

Various monogenic and complex age-related diseases demonstrate the importance for telomere length maintenance in health and disease. Importantly, the risk factors that affect telomere length are the same as those that contribute to common age-related diseases, suggesting a potential causal role for telomere length in these disorders.

1.1.1.4.1 Dyskeratosis Congenita

Dyskeratosis congenita is a severe inherited disorder caused by an increased rate in telomere shortening. The clinical abnormalities of this disease include mucosal leukoplakia, abnormal skin pigmentation, and nail dystrophy (Dokal 2000). In addition, symptoms indicative of premature aging such as, pulmonary disease, dental abnormalities, alopecia and esophagostenosis are also commonly present (Carroll, Ly 2009). The disease is caused by inheritance of many mutations in genes involving in telomere length such as X-linked recessive mutation in *DKC1* or autosomal dominant mutation in *TERT, TERC, RTEl1* and *TINF2*, and even the inheritance of the autosomal recessive mutations in *TERT, RTEL1*, *CTC1*, *NOP10*, *NHP2*, or *WRAP53* (Walne et al.

2007, Vulliamy et al. 2008, Keller et al. 2012, Nelson and Bertuch 2012, Walne et al. 2013a, Walne et al. 2013b, Ballew et al. 2013, Le Guen et al. 2013).

1.1.1.4.2 Idiopathic pulmonary fibrosis

Idiopathic pulmonary fibrosis is a severe and progressive disease that eventually leads to pulmonary failure. Around 8-15% of the idiopathic pulmonary fibrosis individuals have a documented family history with mutations in telomerase components *TERT* and *TERC* (Armanios et al. 2007, Tsakiri et al. 2007). A recent GWAS for pulmonary fibrosis identified seven new loci including the telomere maintenance gene *OBFC* (Fingerlin et al. 2013).

1.1.1.4.3 Cancer

Telomere shortening plays an important role in mediating tumor suppression by causing cells that have been through many cell divisions and have critically short telomeres to enter replicative senescence (Shay 2010). In the absence of appropriate check-points cancer cells bypass replicative senescence and can proliferate indefinitely. A high percentage of cells do this by extending their telomere length through the reactivation of telomerase or through the ALT pathway (Günes and Rudolph 2013, Shay, Bacchetti 1997). Both longer and shorter telomere length has been found to correlate with different cancers. Longer telomere length is associated with cancers including breast, lung cancer, hepatocellular carcinoma and renal cell carcinoma (Han et al. 2008, Svenson et al. 2008, Gramatges et al. 2010, Shen et al. 2011, Liu et al. 2011, Lan et al. 2009, Svenson et al. 2009, Rode et al. 2016). Cells with longer telomeres have the capacity to go through more cell divisions than cells with short telomeres meaning they have more time to accumulate potential cancer causing mutations. Other studies found that the shorter telomere length has been linked to cancers of kidney, lung, prostate, bladder, breast, neck and head (Wu et al. 2003, Shen et al. 2009, Ma et al. 2011).

1.1.1.4.4 Telomere length and cardiovascular disease

Coronary artery disease (CAD), the leading cause of infirmity and death worldwide (Wang et al. 2016), is an age-related disease caused by atherosclerosis. Multiple studies have shown that aging is a key risk factor of atherosclerosis, the underlying cause of CAD (Grundy et al. 1999, Kiechl, Willeit 1999). In addition, telomere length was shorter in atherosclerosis lesions of the human abdominal aorta than in non-atherosclerotic portions (Chang and Harley 1995, Aviv et al. 2001, Okuda et al. 2000). Samani *et al*. (2001) reported the first CAD case-control study to investigate telomere length, found a significant association between shorter telomeres and CAD risk. In the prospective WOSCOPS study, Brouilette *et al* (2008) found that individuals in the tertile with the shortest telomere length had a 44% increased risk of CAD over a 5 and half-year follow-up period compared to those in the tertile with the longest telomere length. More recently, Codd *et al* (2013) found evidence of a causal link between shorter telomere length and increased risk of CAD using a genetic risk score incorporating common genetic variants that associate with telomere length.

1.1.1.5 Identification of novel genetic loci controlling telomere length

1.1.1.5.1 Genome-wide association studies

Genome-wide association studies (GWAS) are designed to discover common variants related to traits or diseases **(Figure 1.6)** (Musunuru, Kathiresan 2010, Smith et al. 2009, Hindorff et al. 2009, Barrett et al. 2008). To identify the genetic variants that contribute to a disease or trait GWAS employ either a case-control or population-based study design and genotype thousands of single-nucleotide polymorphisms (SNPs) in many individuals. After genotyping, a linear regression analysis is performed for a quantitative trait (such as telomere length) and a logistic regression or chi squared analysis is carried out for a discrete trait (such as CAD) to test the association of each SNP. Because of the large number of tests performed, genome-wide significance is set to the Bonferroni corrected threshold of $P<5x10^{-8}$ (Bush and Moore 2012). The first large-scale GWAS was published in 2007 by the Wellcome Trust Case Control Consortium (WTCCC) (Zeggini et al. 2007) and involved the comparison of 2,000 individuals for seven different complex

diseases with a shared set of 3,000 controls. In the last decade hundreds of additional GWAS have been performed to identify the genetic contribution to a variety of other diseases and traits.

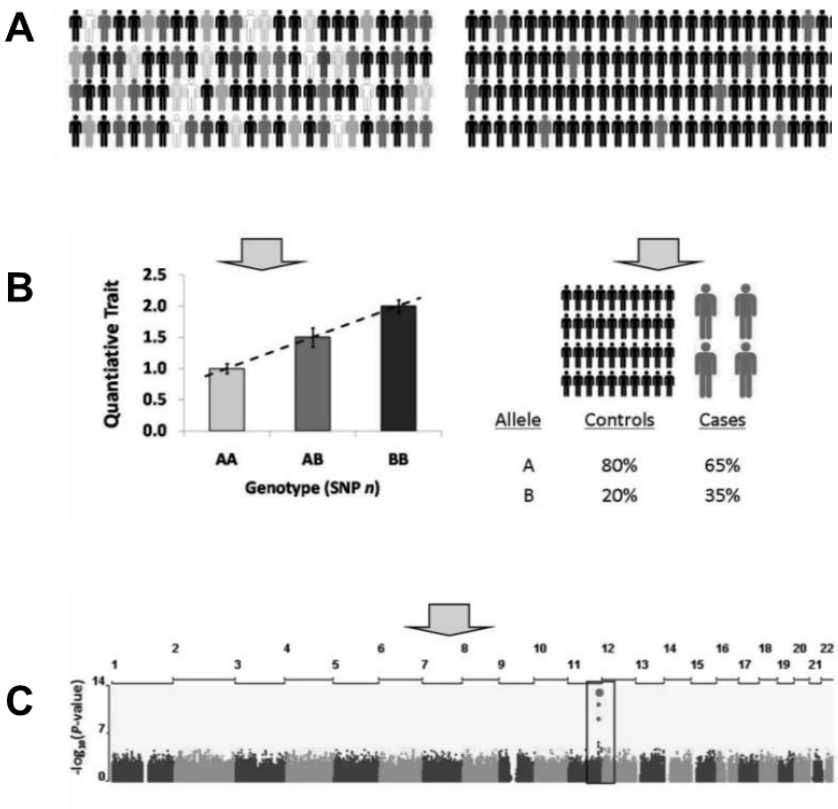


Figure 1.6: The genome-wide association study. (A) A population-based study design for a quantitative trait (left) and a case-control study design for a disease discrete trait (right). SNPs from each individual are genotyped using a high-density microarray and **(B)** association between SNPs and phenotype of interest is tested. **(C)** A representative Manhattan plot showing a single association signal on chromosome 11 reaching genome wide significance ($P<5x10^{-8}$). Taken from Hegele (2010).

1.1.1.5.2 Genome-wide association studies of telomere length

The first GWAS to identify a genome-wide significant association with telomere length was reported by Codd *et al*. (2010). The study involved the measurement of mean leukocyte telomere length in 2,917 individuals with follow-up replication in 9,492 subjects and identified a locus on chromosome 3 close to *TERC*. The effect of each allele corresponds to a reduction in telomere length about 75bp, equivalent to around 3 years of age-related telomere-length attrition. Subsequently, Levy *et al*. (2010) reported a locus on chromosome 10 containing *OBFC1*, which encodes a component of the CST complex and is involved in the replication and capping of telomeres (Palm and de Lange 2008, Miyake et al. 2009). In the largest study to date, Codd *et al*. (2013) identified seven loci associated with telomere length **(Figure 1.7)**.

Five of the loci identified by Codd *et al* contain strong candidate genes involved in telomere biology (*TERC, TERT, OBFC1, RTEL1, and NAF1*). As described previously, TERT and *TERC* are the main components of telomerase and OBFC1 is a member of the CST complex (Miyake et al. 2009, Palm and de Lange 2008). *RTEL1* encodes a DNA helicase that contributes to the elongation and maintenance of telomeres (Wang and Meier 2004, Egan and Collins 2012, Miyake et al. 2009, Ding et al. 2004, Barber et al. 2008). NAF1 is a protein required for maturation of H/ACA small nuclear RNAs, which includes *TERC* (Egan and Collins 2012). Interestingly, the two other loci found to associate with telomere length, *ACYP2* and *ZNF208*, do not contain genes with known role in telomere biology.

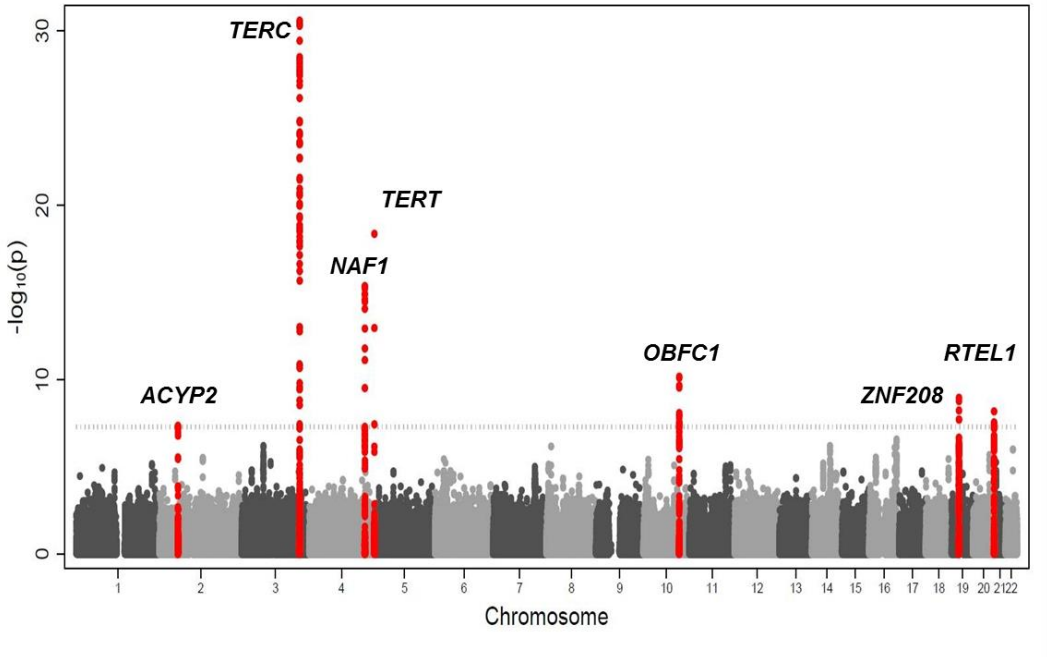


Figure 1.7: Manhattan plot of the seven loci associated with telomere length. The X axis shows chromosome location for the 2,362,330 SNPs that were tested. Y axis shows significance (-log10(P value)) of association. The dotted line indicates a genome-wide level of significance of $P = 5\times10^{-8}$. The red plots refer to associated loci. Taken from Codd *et al*. (2013)

The *ZNF208* locus is located on chromosome 19 and contains a cluster of genes which encode zinc finger proteins involved in various cellular and developmental differentiation processes (Pieler and Bellefroid 1994, Eichler et al. 1998). The *ACYP2* gene is located on chromosome 2 and encodes the enzyme Acylphosphatase-2 that stimulates the hydrolysis reaction of carboxyl-phosphate bond of acyl phosphates (Modesti et al., 1993). Within intron 3 of *ACYP2* there is an additional gene called Testis Specific Y like 6 (*TSPYL6*), a member of the *TSPY* like gene superfamily. There have been no studies into TSPYL6 protein function, but the TSPYL6 protein contains a nucleosome assembly protein (NAP) domain. Proteins that contain NAP domains are histone chaperones and can function as regulators of gene expression, cell division and chromatin remodelling (Rodriguez et al., 1997, Carlson et al., 1998).

As the *ACYP2* loci and *ZNF208* telomere length loci do not contain genes known to be involved in telomere length. Therefore, a better understanding of the genes involved in telomere length maintenance can be achieved by investigating these loci. As GWAS do not identify the causal gene we decided to develop a functional screen that would allow the identification of genes required for telomere length maintenance. The screen involves the knockout of candidate genes in human induced pluripotent stem cells (hiPSCs), which express telomerase and maintain telomere length using CRISPR/Cas9 mediated genome editing. Subsequent measurement of telomere length in knockout cell lines would demonstrate whether a gene is required for telomere length maintenance.

1.1.2 Genome editing

Genome editing is a type of genetic engineering in which a specific target site in the DNA sequence of a cell line or organism is altered by either deletion, insertion or replacement using an engineered nuclease. In the past few years, there has been enormous interest in the newly emerging genome editing technologies that allow the manipulation of specific genes in many organisms, including humans, and has a range of applications in industry, agriculture and therapeutics. In addition, genome editing can be applied as a research tool to target specific genes or introduce single disease or trait associated sequence variants.

1.1.2.1 Genome editing with engineered nucleases.

Genome editing relies on the targeting of engineered nucleases to specific DNA sequences where they induce a double strand break (DSB) (**Figure 1.8**) (Cathomen and Joung 2008, Carroll 2008, Urnov et al. 2010). Cells employ two mechanisms to repair DSBs; high-precision homologous recombination (HR) and error-prone non-homologous end joining (NHEJ) (Lieber 1999, Huertas 2010, Carbery et al. 2010). Genome editing technologies exploit these mechanisms to allow the introduction of genetic alterations at specific sites in the genome (Sun et al. 2012, Perez-Pinera et al. 2012). HR, through homology directed repair (HDR), can mediate substitution between the endogenous targeted genomic region and an exogenous DNA fragment, which carries site-specific sequence differences resulting in precise genetic alterations. NHEJ produces imprecise changes such as small deletions or insertions, which can result in gene disruption by introducing frame-shift mutations (Sun and Zhao 2013) (**Figure 1.8**).

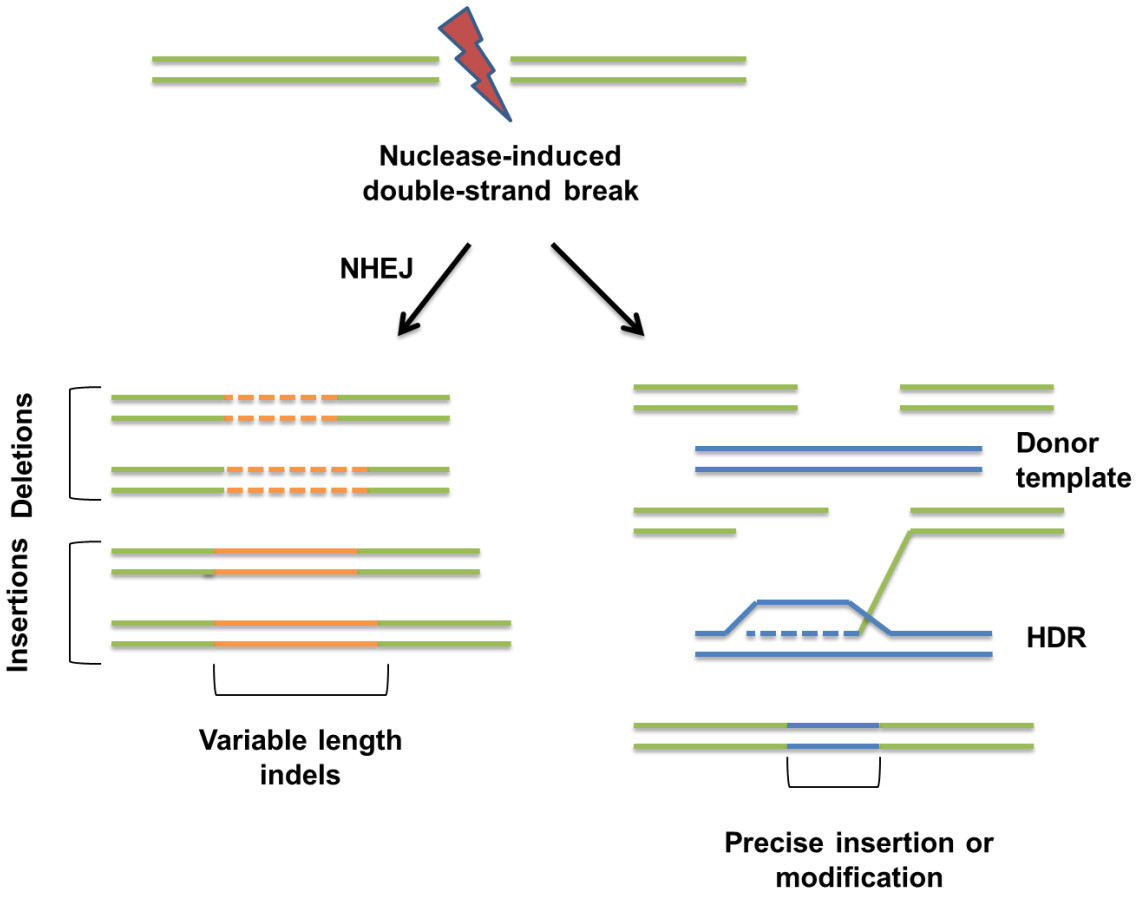


Figure 1.8: Nuclease-induced genome editing. Genome editing relies on the site-specific induction of double strand breaks (DSBs) induced by an engineered nuclease. A DSB can be repaired by either of two pathways: non-homologous end joining (NHEJ) or homology-directed repair (HDR). NHEJ can result in insertions or deletion mutation while the inclusion of a donor template with homology to the DSB site can introduce precise insertion by HDR. This figure is updated from Sander and Joung (2014).

There are three main types of engineered nuclease; Zinc finger nucleases (ZFNs), which first emerged in the late 1990s, Transcription activator-like effector nucleases (TALENs), which appeared in 2011 and the most recent, and perhaps most interesting discovery the Clustered regularly interspaced short palindromic repeat (CRISPR)/Cas9 system.

1.1.2.1.1: Zing-Finger nucleases

ZFNs are artificial enzymes which consist of a non-specific type II endonuclease FokI DNA cleavage domain fused with a series of zinc finger polypeptides (Bibikova et al. 2003, Miller et al.1985, Porteus and Baltimore 2003). The zinc-finger domains supply the ZFN with the ability to bind a specific DNA sequences. Each zinc-finger domain, which are approximately 30 amino acids in length and contain a tandem array of Cys2-His2 fingers, is able to recognize 3 DNA base pairs (Beerli and Barbas 2002, Miller et al. 1985, Wolfe et al. 2000). Typically, arrays of 3-6 zinc-finger domains are linked together to produce a DNA-binding domain with specificity to 9-18 base pairs (Carlson et al. 2012). To produce DSBs, two endonucleases Fok I domains are required, therefore the ZFNs are utilized in pairs and are designed to bind to each side of the specific target DNA sequence (Händel et al. 2009, Urnov et al. 2010) (**Figure 1.9**).

ZFNs allowed several different species that had not previously been amenable to genetic manipulation to be targeted and provoked interest in utilizing these tools for gene therapy (Zou et al. 2009, Hockemeyer et al. 2011, Soldner et al. 2011). As well as in human cells (Moehle et al. 2007, Zou et al. 2009), ZFNs have been successful utilized in Zebrafish (Doyon et al. 2008, Meng et al. 2008), mice (Carbery et al. 2010), rats (Geurts et al. 2009, Mashimo et al. 2010), rabbits (Flisikowska et al. 2011) and pigs (Whyte et al. 2011, Hauschild et al. 2011). However, ZFNs are both expensive and technically demanding to design and produce. Ramirez *et al*. (2008) reported difficulties in designing ZFNs that target desired sites due to the limitations in the modular structure of ZF domains. Moreover, undesired mutations and chromosome aberrations can be caused by off-target DSB induction by ZFNs (Pattanayak et al. 2011, Radecke et al. 2010).

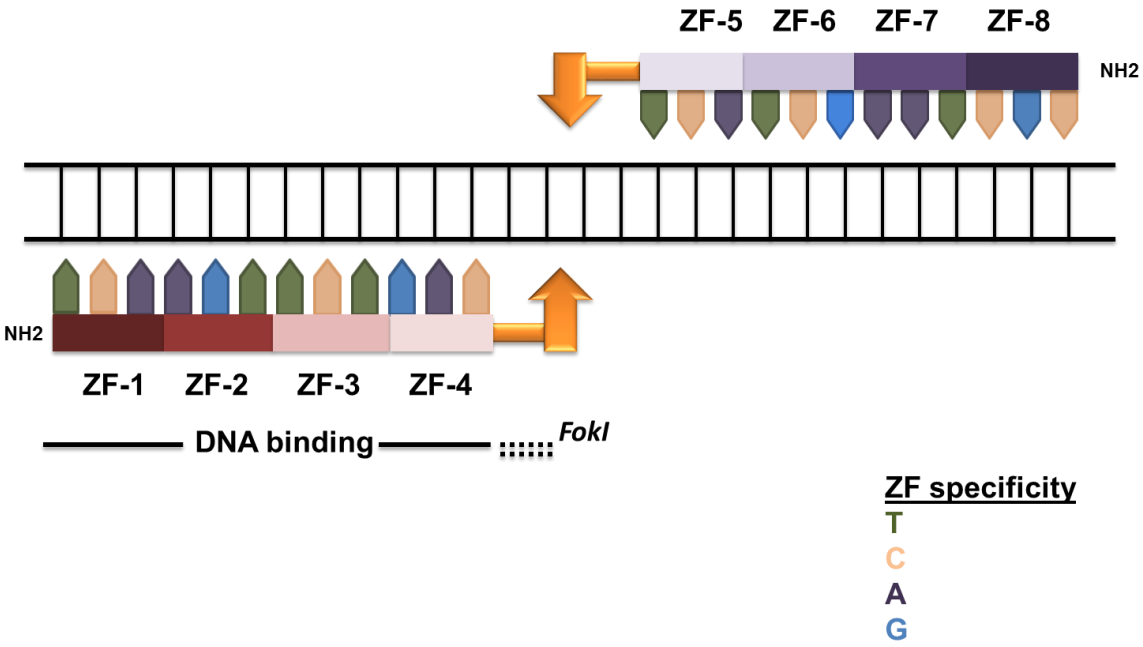


Figure 1.9: Structure and design of zinc-finger nucleases. Each zinc-finger nuclease (ZFN) consists of a single FokI nuclease cleavage domain (orange arrows) and a DNA binding domain designed from ZF modules. Each module recognizes 3bp DNA sequence. DSB induction requires two ZFNs. This figure is adapted from Carlson *et al*. (2012).

1.1.2.1.2 Transcription Activator-Like Effector Nucleases

TALENs, which are of bacterial origin, emerged after ZFNs and are an alternative method for producing targeted DSBs (Carlson et al. 2012). Like ZFNs, TALENs incorporate the Fok I nuclease as the DNA cleavage domain with a targetable DNA binding domain (Sun and Zhao 2013). The genes encoding Transcription Activator- like Effector (TALE) proteins are found in the plant pathogen *Xanthomonas*, which releases TALEs into plant cells in order to prepare a suitable environment for the invading pathogen (Boch and Bonas 2010). TALEs consist of a series of 34 amino acid tandem repeats. Two amino acids termed repeat variable disresidues (RVDs) at the center of each repeat are variable and this variability controls base pair recognition. For example, NI (asparagine, isoleucine) recognizes adenine, HD (histidine, aspartic acid) recognizes cytosine, NG (asparagine, glycine) recognizes thymine, and NN (asparagine, asparagine) recognizes guanine and adenine (Boch et al. 2009, Moscou and Bogdanove 2009) (**Figure 1.10**). The final repeat contains only 20 amino acids and is called a "half- repeat" (Sun and Zhao 2013, Benetos et al. 2004).

Unlike ZFNs, TALENs can be easily and rapidly designed to target almost any DNA sequence. In addition, they cause fewer off-target effects and have reduced cytotoxicity compared with ZFNs (Ding et al. 2013, Mussolino et al. 2011).

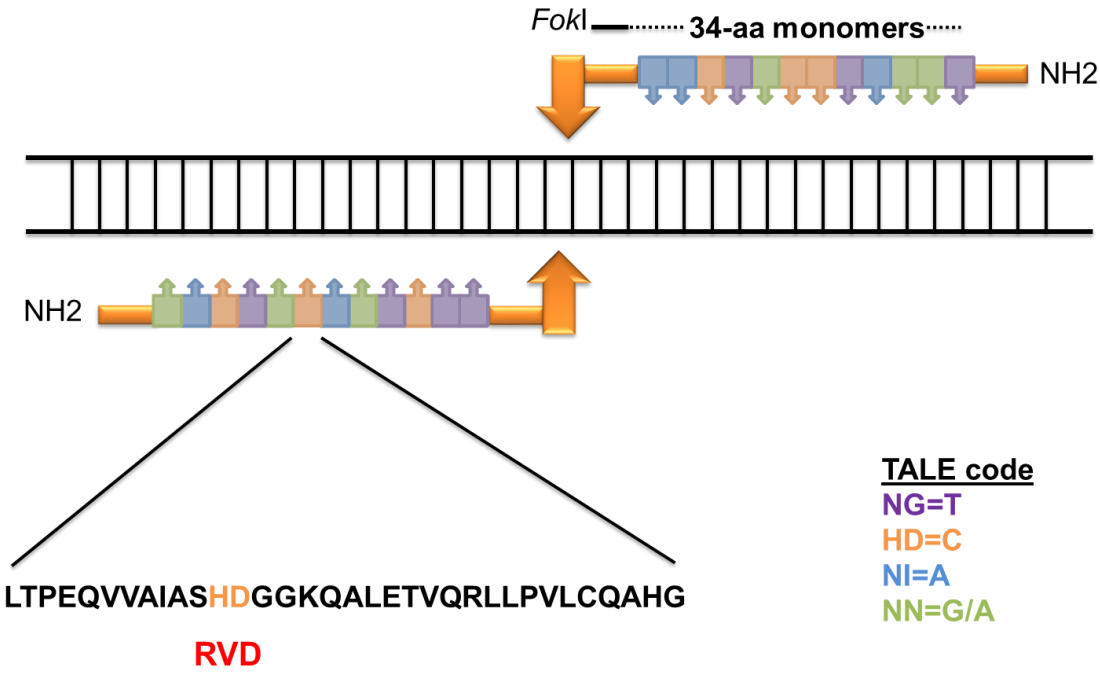


Figure 1.10: Transcription activator like effector nucleases. TALENs consist of two parts; a non-specific cleavage Fok I cleavage domain (orange arrow) and a DNA binding domain consisting of a series of 34 amino acid repeats. Repeat variable di-residues (RVDs) (shown in red) is variable and responsible for specific recognition of the target DNA base-pair. This figure is adapted from Carlson *et al*. (2012).

1.1.2.1.3 The CRISPR/Cas9 system

Shortly before the start of this project a CRISPR/Cas9 system was reported as a more affordable and easier to use genome editing tools (Mali et al. 2013, Cong et al. 2013). There has been an enormous amount of scientific interest in this system over the past 4 years and CRISPR/Cas9 targeting has been performed in wide range of different species.

CRISPRs were originally identified in *Escherichia coli* in 1987 and characterized as a rare sequence element consisting of a sequence of 29 nucleotide repeats separated by 32 nucleotides 'spacer' sequences, however, the functional relevance of these repeats remained obscure for many years (Ishino et al. 1987). Over a period of 24 years a combination of, bioinformatic, genetic and functional studies revealed that CRISPR is part of a prokaryotic adaptive immune system, which utilizes RNA to target and damage parasites (Al-Attar et al. 2011, Deveau et al. 2010, Makarova et al. 2011). Previously, it was thought that adaptive immunity only occurred in eukaryotes (Wiedenheft et al. 2012). The mechanism of CRISPR adaptive immunity in bacteria involves the integration of a short sequence of foreign nucleic acid (e.g viral genomic DNA) into the host genome at one end of a CRISPR repeat **(Figure 1.11)** (Andersson and Banfield 2008, Barrangou et al. 2007, Garneau et al. 2010). These foreign protospacer sequences now serve as memory for the same invaders during the next infection.

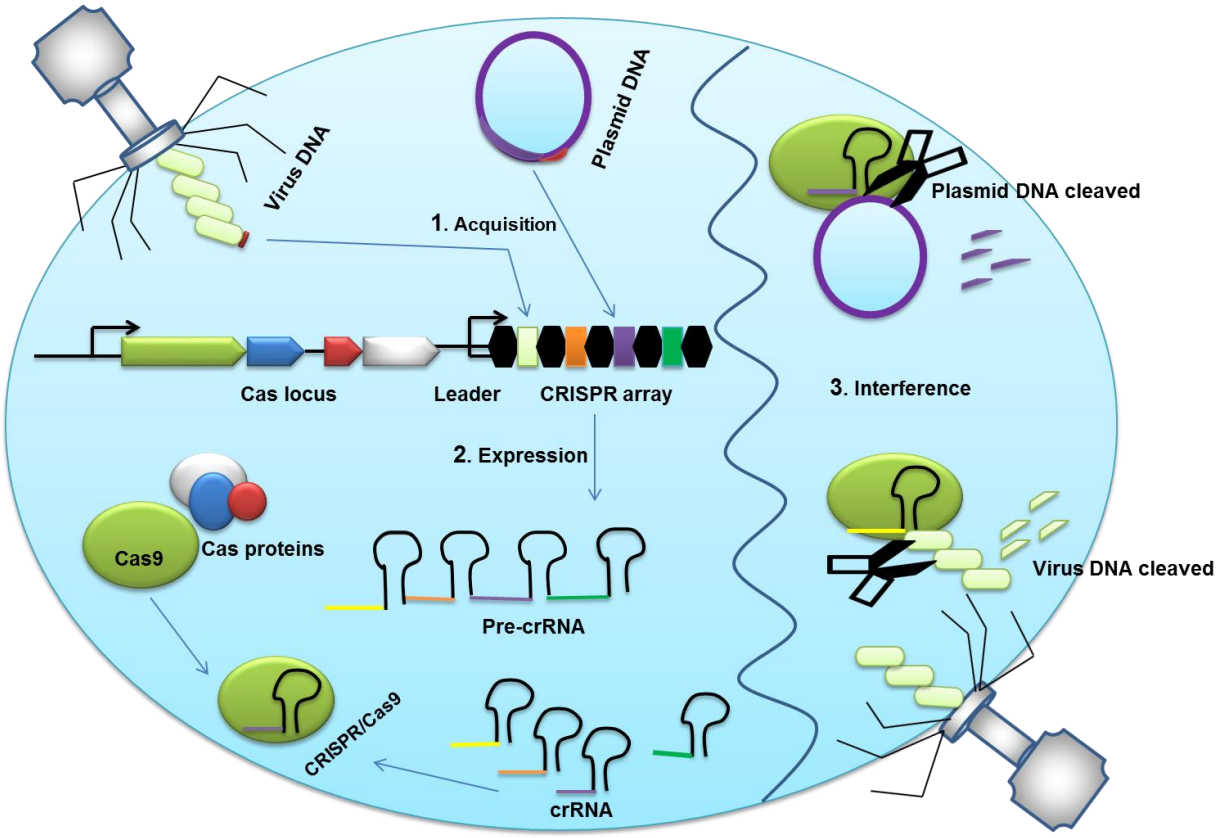


Figure 1.11: The CRISPR/Cas system in the bacterial adaptive immune system. 1. On the left of the figure; foreign DNA from invading pathogens is incorporated into the CRISPR locus. 2. The CRISPR array is expressed as pre-crRNA, processed into crRNA before complexing with Cas9. 3. On the right of the figure; the CRISPR/Cas9 complex cleaves the DNA of invading pathogens. The figure is adapted from Wiedenheft *et al*. (2012).

The CRISPR repeat is transcribed to pre-crRNA before being processed into short CRISPR-derived RNAs (crRNA) (Wiedenheft et al. 2012) containing 20 nucleotide sequence complementary to the foreign DNA immediately followed by a trinucleotide sequence called the protospacer adjacent motif (PAM), such as 5-NGG-3 (where N is any nucleotide) (Ran et al. 2013b). Each crRNA complexes with the Cas9 endonuclease, which cleaves the foreign DNA (Hwang et al. 2013).

In the bacterial system, the Cas9 protein complexes with both the crRNA and a trans-activating crRNA (tracrRNA) (Jinek et al. 2012) (**Figure 1.12, A**). Jinek *et al* (2012) combined these RNAs to produce a single-chain chimeric RNA called a single-guide-RNA (sgRNA) and programmable sgRNA/Cas9 complex (**Figure 1.12, B**). This approach with a sgRNA and Cas9 protein has been used in subsequent genome editing applications.

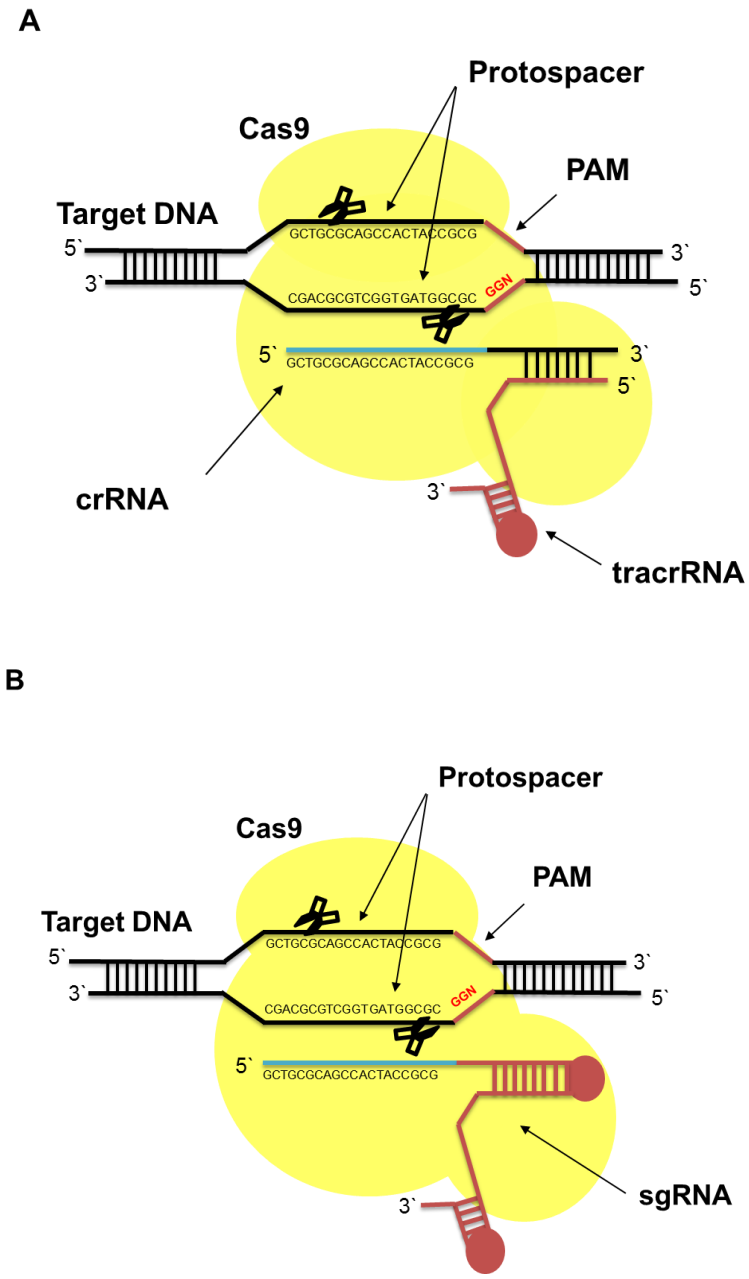


Figure 1.12: Natural and engineered CRISPR-Cas9 systems. (**A**) Natural CRISPR/Cas9 system involves the integration of "protospacer" foreign DNA sequence within the CRISPR repeats producing crRNA containing the "protospacer". A trans activating CRISPR RNA (tracRNA) hybridizes with the crRNA forming an RNA complex that can associate with Cas9 nuclease and cleave the target DNA. (**B**) The engineered CRISPR/Cas9 system involves using crRNA, consisting of a sequence designed to target a specific DNA sequence, fused with part of the tracrRNA to produce a single-guide RNA (sgRNA). The sgRNA associates with Cas9 and cleaves the target DNA. This figure is updated from Jinek *et al.* (2012).

After the first publication showing that Cas9 can be directed to cleave DNA in a site-specific manner (Jinek et al. 2012) there have been many papers reporting efficient CRISPR/Cas9 mediated genome editing in a variety of cells and organisms. The first studies demonstrated that Cas9 could be directed to target endogenous genes in bacteria (Jiang et al. 2013), human cancer cell lines and human pluripotent stem cells (Mali et al. 2013, Cong et al. 2013, Jinek et al. 2012, Cho et al. 2013) as well as in whole organisms such as zebrafish (Hwang et al. 2013). Within one year the CRISPR/Cas9 system had been utilized to alter the genomes of yeast (DiCarlo et al. 2013), tobacco (Li et al. 2013, Nekrasov et al. 2013), thale cress (Li et al. 2013), rice (Shan et al. 2013, Xie, Yang 2013), wheat (Shan et al. 2013), sorghum (Jiang et al. 2013b), mice (Wang et al. 2013, Shen et al. 2013), rats (Li et al. 2013), rabbits (Yang et al. 2014), frogs (Nakayama et al. 2013), fruit flies (Yu et al. 2013, Bassett et al. 2013), and roundworms (Friedland et al. 2013).

CRISPR/Cas9 mediated genome editing has been reported to have many advantages over ZFNs and TALENs with the main benefit being ease of design. The Cas9 nuclease can be directed to any target site consisting of 20 nucleotides followed by a PAM simply by changing the first 20 nucleotides of the sgRNA to match the target DNA sequence, and similar or higher targeting efficiency and the capacity to facilitate diverse types of genome editing (Cong et al. 2013, Mali et al. 2013, Hou et al. 2013, Jinek et al. 2013, Ran et al. 2013b, Sander and Joung 2014). CRISPR/Cas9 had been used to stimulate the NHEJ-mediated introduction of insertion and mutation mutations as well as to induce HR with both single-stranded oligonucleotide donor templates and double-stranded plasmid donors. This CRISPR/Cas9 system also enabled the parallel induction of DSBs at multiple target sites; a unique advantage of this platform relative to ZFNs and TALENs (Cong et al. 2013, Xiao et al. 2013, Horii et al. 2013).

Regarding improvements in the technology, there has also been a lot of effort to improve the efficiency of this technology. One method has involved the identification of Cas9 alternatives, such as homologues from other bacterial species and the directed mutagenesis of Cas9 itself. An example Cpf1 (CRISPR from *Prevotella* and *Fracisella1*) is an RNA-guided nuclease characterized by Zetsche *et al* (2015), mediates efficient DNA cleavage with distinct advantages over Cas9. Cpf1 does not require tracrRNA and causes a 4-5 nucleotide 5' overhang, which contributes to improve HR efficiency. In addition,

the sgRNA uses a T-rich rather than G-rich PAM, therefore, expanding the number of targetable sites in a genome (Zetsche et al. 2015).

1.1.2.1.3.1 CRISPR/Cas9 off-target effects

A major concern with the CRISPR/Cas9 system is the potential for off-target effects (Fu et al. 2013, Cho et al. 2014, Kim et al. 2015) which can disrupt gene function and cause genome instability and therefore influence research and therapeutic applications. There are several different methods that can be used for detection. The most basic method is to use Sanger sequencing to screen computationally predicted off-target sites (i.e. the regions of the genome that share some homology to the sgRNA target sequence), but this has the potential to be time-consuming and miss unexpected off-target mutations. This can be overcome using deep sequencing techniques such as whole-exome sequencing, which will detect the off-target effects most likely to produce a phenotype and whole genome sequencing (WGS), which can detect all induced changes (Fu et al. 2014, Smith et al. 2014, Veres et al. 2014, Kim et al. 2015). There are also examples of targeted deep sequencing methods which rely on tagging CRISPR/Cas9 induced DSBs. For example, Wang *et al*. (2015) used integrase-defective lentiviral vectors (IDLVs) to detect off-target DSBs. In this technique, IDLVs integrate at CRISPR/Cas9 induced DSBs. The integration sites can be amplified by PCR and integration sites mapped by deep sequencing.

Because of the potential functional consequences of off-target effects there have been major efforts to develop methods to minimising off-target CRISPR/Cas9 activity. The primary method of reducing off-target effects is through better sgRNA design as a result of an improved understanding of CRISPR targeting. Many sgRNA design tools (see below) predict possible off-target effects and as most CRISPR/Cas9 applications allow flexibility in sgRNA guide selection, therefore, the sgRNAs with the lowest off-target predictions can be selected for experiments. Good experimental practice can also help control for off-target effects. By using multiple targeted clones, which are unlikely to all contain the same off-target mutations, or multiple sgRNAs, which are unlikely to induce the same off-target changes, the functional consequence of any effects can be controlled. Screening of targeted lines also allows clones carrying unwanted changes to be identified

and an informed decision made as to whether each clone is suitable for down-stream investigation. Methods for reducing the frequency of off-target induction include using a Cas9 "nickase", which has a mutation that results in single-strand rather than DSBs being induced. This allows a strategy where two offset sgRNAs are used and therefore increases specificity by reducing off-target homology (Ran et al. 2013a). Adding two extra guanine (G) nucleotides at the 5` end of the sgRNA has also been reported to reduce off-target effects (Cho et al. 2014) and truncated sgRNAs have also been shown to decrease the number off-targets (Fu et al. 2014). Furthermore, a catalytically dead Cas9 (dcas9)-FokI fusion, which works as a dimer requiring two sgRNA was shown to dramatically increase targeting specificity (Gullinger et al. 2014, Tsai et al. 2014). In addition, transfection of purified recombinant Cas9 protein and sgRNA ribonucleoproteins (RGENs) induced site-specific mutations at frequencies up to 79% with reduced off-target induction compared to plasmid transfection (Kim et al. 2014). Finally, Kleinstiver *et al*. (2016) recently created a high fidelity *Streptococcus Pyogenes* Cas9 (SpCas9-HF1) by mutagenesis.

1.1.2.1.3.1.1 sgRNA design

There are several non-commercial CRISPR sgRNA design resources available **(Table 1.1)**. These web tools can be used to predict potential CRISPR target and design sgRNAs in a user-given sequence or in reference genomes. These tools also provide off-target and activity prediction to allow the best sgRNAs to be selected. In addition, numerous biotechnology companies also host sgRNA design tools or have libraries of pre-designed sgRNAs. It should be noted that the sgRNAs used in this study (described in **Chapter 4** and **5**) were designed and generated by Horizon Discovery Ltd (Cambridge, UK) using an in-house design method prior to their launch of CRISPR/Cas9 products. To the best our knowledge, the only design tools available at the time were the MIT and Broad hosted tools.

Design tool	Web adress
Deskgen	https://www.deskgen.com/landing/
MIT	http://crispr.mit.edu:8079
Broad	https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design
GeneArt	http://www.thermofisher.com/uk/en/home/life-science/genome-editing/geneart-crispr/geneart-crispr-search-and-design-tool.html
ChopChop	http://chopchop.cbu.uib.no/
CCTop	http://crispr.cos.uni-heidelberg.de
OFF-Spotter	https://cm.jefferson.edu/Off-Spotter/
GT-Scan	http://gt-scan.csiro.au
CRISPRseek	http://www.bioconductor.org/packages/release/bioc/html/CRISPRseek.html
crispr grna	https://www.atum.bio/eCommerce/cas9/input
CRISPR RNA Configurator	http://dharmacon.gelifesciences.com/gene-editing/crispr-rna-configurator/
Cas-OFFinder	http://www.rgenome.net/cas-offinder/
Benchling	https://benchling.com/crispr
CRISPRdirect	http://crispr.dbcls.jp
Breaking Cas	http://bioinfogp.cnb.csic.es/tools/breakingcas/
SYNTHEGO	https://design.synthego.com/

Table 1.1: sgRNA design web tools.

1.1.2.1.3.2 Molecular assessment of telomeres

1.1.2.1.3.2.1 Telomere length measurement

There are various methods that have been employed to measure the telomere length and these can be classified into three categories; 1) terminal restriction fragment (TRF), whichto measures the average length and length distribution of telomeres in a sample, 2) fluorescent *in situ* hybridization (FISH) techniques such as Q-FISH and Flow-FISH, which are utilized to examine variation in telomere length between chromosomes and cells and 3) quantitative polymerase chain reaction (qPCR)-based techniques such as monochrome multiplex quantitative polymerase chain reaction (MMqPCR) and absolute telomere length quantitative polymerase chain reaction (aTLqPCR) that measures telomeric DNA relative to a single copy gene DNA. In general, these methods use primers or nucleic acid probes to probe the telomeric repeat sequence and each technique has certain advantages and disadvantages properties **(Table 1.2)**.

Technique	Methodology	Advantages	Disadvantages	References
TRF	Exploits specific restriction enzymes that digest genomic DNA but cannot digest telomeric repeats. This treatment results in long undigested telomeres and shorter genomic fragments. The products can be resolved depending on size by the agarose gel electrophoresis and the telomere fragments identified using labelled telomere oligonucleotide probes in gel-hybridization or southern blotting.	A “Gold standard” technique used as a reference when comparing the advantages and drawbacks of alternative telomere length assays. Does not require specialized equipment.	Requires large amount of DNA. Subtelomeric polymorphisms can impact the data. Measures mean telomere length, but does not provide readings for individual telomeres, short telomeres or ends lacking a telomere. Labour intensive	(Allshire et al. 1989, Harley et al. 1990, Kimura et al. 2010, Aubert et al. 2012)
Q-FISH	Measures telomere length in fixed cells in metaphase or interphase using a synthetic peptide nucleic acid (PNA) probe $(CCCAT)_3$ hybridized with nonspecific DNA stain such as 4′, 6-diamidino-2-phenylindole (DAPI) or propidium idodide for visualization.	Provides a high resolution measurement of a single telomere. Measures the telomere lengths in individual cells. Can identify variation in individual telomeres, identify signal free ends, end-to-end telomeres and provide a mean telomere length measurement.	Measures mean telomere length in fluorescent relative units compared to a standard (centromeric) value. Requires dividing cells for metaphase chromosomes. Highly technical and labour intensive. Requires a speciialised fluorescent microscope.	(Landsdorp et al. 1996, Krejci and Koch 1998, Aubert et al. 2012, Vera and Blasco 2012, Artandi et al. 2000, Goldman et al. 2008, Canela et al. 2007)

F-FISH	An adaptation of the Q-FISH technique which assesses the telomere lengths of cells in suspension by flow cytometry technology to quantify mean fluorescence.	Measures the mean telomere lengths of specific cells. Can provide cell type specific information if combined with other labels. Potential for automation.	Measures mean telomere length in fluorescent relative units. Does not provide reading of individual short telomeres or ends lacking a telomere. Highly technical and labour intensive. Requires flow cytometer.	(Hultdin et al. 1998, Alter et al. 2007, Samassekou et al. 2010, Aubert et al. 2012)
qPCR (MMqPCR, aTLqPCR)	Telomeric repeats are amplified over 20-40 cycles and quantified using a fluorophore such as SYBR green., Telomere length is then determined by dividing the quantity of telomeric repeats (T) to the quantity of a single copy gene (S) to give the T/S ratio. By using a standard curve of known telomere lengths, it can provide a measurement of absolute telomere length rather than relative telomere length.	Only requires a small amount of DNA. Less labour intensive. Relatively low cost.	Measures mean telomere length and does not provide a reading of individual telomeres, short telomeres or ends lacking a telomere. Some variation among replicate estimates. Does not provide an absolute kilobase length estimate unless coupled with a standard. Requires qPCR equipment.	(Ding and Cantor 2004, Cawthon et al. 2009, Cawthon 2009, O'Callaghan and Fenech 2011, Cunningham et al. 2013, Koppelstaetter et al. 2005, Aviv et al. 2011)

Table 1:2 Advantages and disadvantages of the different methods used to measure telomere length. The table is adapted from Montpetit *et al*. (2014)

1.1.2.1.3.2.2 Telomerase activity

As described earlier, telomere length is largely regulated by telomerase and there are several ways in which telomerase activity can be assessed. The most basic method to assess telomerase is to analyse the expression of the core telomerase components *TERC* and *TERT* at the RNA or protein level. Other techniques, which directly examine telomerase activity have been used since the first description of telomerase. In 1985 Carol Greider and Elizabeth Blackburn demonstrated the presence of a "telomere terminal transferase" in protozoan extracts using a telomere elongation activity assay (Greider and Blackburn 1985). Four years later, using a modified version of the same assay, Gregg Morin proved thattelomerase is present in human cells (Morin 1989). In this assay, telomeric sequences are elongated *in vitro* by telomerase reverse transcription present in cell lysates using a telomeric repeat oligonucleotide primer and radiolabelled nucleotides. The products of this reaction are then separated by polyacrylamide gel electrophoresis and compared with positive bands by autoradiography. Although this method is reliable for detection of telomerase activity, it is time consuming, requires a large amount of sample and is technically complex. Kim *et al*. (1994) developed this technique to establish the telomeric repeat amplification protocol (TRAP) in which the telomeric repeat is added by telomerase to a non-telomeric substrate primer followed by amplification by PCR and this technique is now widely used to assess telomerase activity with different modifications (Zhou and Xing 2012, Gobourdes et al. 2004, Kong et al. 2007, Xiao et al. 2010, Wang et al. 2017, Zhang et al. 2017).

1.1.3 Human pluripotent stem cells

Pluripotent cells are able to differentiate into any of the cell-types within a body. There are two types of pluripotent stem cells (PSCs); embryonic stem cells (ESCs), which are isolated from embryos (Thomson et al. 1998, Reubinoff et al. 2000, Adaikalakoteswari et al. 2005) and induced pluripotent stem cells (iPSCs), which are produced by reprogramming somatic cells using transcription factors (Takahashi and Yamanaka 2006, Takahashi et al. 2007).

Telomeres are elongated during reprogramming of mouse and human somatic cells into iPSCs due to the upregulation in levels of telomerase and both ESCs and iPSCs maintain their telomere length over extended culture (Marion et al. 2009, Yehezkel et al. 2011, Mathew et al. 2010, Suhr et al. 2009). Therefore, telomere length can be considered a marker of pluripotency. Because of these reasons, human induced pluripotent stem cells (hiPSCs) make an ideal cell model for the investigation of genes associated with telomere length. Moreover, the emergence of genome editing technologies, which allow the manipulation of genes in hiPSCs, provides the experimental tools to investigate the requirement of specific genes on telomere maintenance and hiPSC phenotype.

1.1.3.1 Discovery of iPSCs

The discovery of iPSCs was the culmination of numerous findings over 60 years of research. In 1962, Dr Gurdon produced normal tadpoles by delivering nuclei from differentiated intestinal epithelium cells of other tadpoles into unfertilized eggs and therefore demonstrating that the intestinal cell nuclei contain all of the genetic information necessary for all somatic cells (Gurdon, 1962). More than 30 years later, the first animal cloned from an adult cell, Dolly the sheep, was generated by somatic cloning and showed that the oocyte contains factors that can induce somatic nuclei to control the development of an entire organism (Wilmut et al. 1997). Five years later, Tada and colleagues developed an experimental system for the nuclear reprogramming of somatic cells *in vitro* by fusing adult thymocytes with ESCs, demonstrating that ESCs have the capacity to reprogram somatic cells (Tada et al. 2001). Each of the studies outlined above achieved reprogramming via nuclear transfer. Over the same time-period other groups were investigating transcription factors as mediators of reprogramming. In 1987, by overexpressing the transcription factor $Antp^{+}$, Schneuwly and colleagues were able to cause *Drosophila* to develop legs in place of antennae (Schneuwly et al. 1987). In the same year, Davis and his group transfected mouse fibroblasts with myoblast cDNA causing conversion of the fibroblasts into stable myoblasts (Davis at el. 1987). These results led to the concept of “master regulators”, which are transcription factors with the ability to induce a particular lineage (Yamanaka 2012). The convergence of these strands of research with mouse ESCs and human ESC culture (Evan and Kaufman 1981, Martin 1981, Thomson et al 1998) led to the discovery of iPSCs (Takahashi and Yamanaka 2006,

Takahashi et al. 2007). In 2006, Takahashi and Yamanaka, reported the generation of iPSCs from mouse embryonic and adult fibroblasts by introducing transcription factors such as Oct3/4, Sox2, c-Myc, and Klf4 involved in the maintenance of ESC pluripotency under ESC culture conditions (Takahashi and Yamanaka 2006). The resulting cells showed the same growth properties and morphology of ESCs, expressed ESC marker genes and had the ability to form all three germ layers. The following year the same group reported the reprograming of human dermal fibroblasts into hiPSCs (Takahashi et al. 2007).

1.1.3.2 Generation of iPSCs

The first iPSCs were generated by viral reprogramming and used retrovirus (Takahashi and Yamanaka 2006) or lentivirus (Sommer et al. 2009) to introduce the four transcription factors. However, a major disadvantage of these viral delivery methods is that they might insert randomly in the genome with unpredictable consequences. Moreover, it has been found the use of exogenous oncogenes for reprogramming, such as c-Myc, may increase tumorigenic risk and alternative approaches might be need to be use for clinical applications (Okita et al. 2007, Tong et al. 2011). Because of this, a number of technical modifications have been investigated to generate iPSCs using non-integrative methods including episomal vectors (Yu et al. 2009), piggy Bac transposons (Woltjen et al. 2009), sendai viral vectors (Fusaki et al. 2009), plasmid vectors (Okita et al. 2010) and a minicircle vector system (Jia et al. 2010) **(Table 1.3)**. However, further optimization of these previous approaches is needed in order to improve reprogramming efficiency and ensure safety for clinical applications.

Programming method	Example	Advantages	Disadvantages	Programming efficiency %	References
Viral methods	Lentivirus	Transduction of both dividing and nondividing cells	Genomic integration (risk of insertional mutagenesis). Lack of silencing in pluripotent state	0.1-1.5	(Somers et al. 2010, Takahashi and Yamanaka 2006, Brambrink et al. 2008, Blelloch et al. 2007, Yu et al. 2007)
	Retrovirus	Silenced in pluripotent cells	Genomic integration (risk of insertional mutagenesis). Transduction of only dividing cells		
	Adenovirus	No genomic integration	Low efficiency. Requires further validation.	0.0001-0.001	(Stadtfeld et al. 2008, Zhou and Freed 2009)
	Sendai virus		Need culture high temperature and long time to remove the virus from reprogramed cells	0.1	(Fusaki et al. 2009, Ban et al. 2011, Seki et al. 2010)
Non-viral methods	mRNA transfection	Footprint-free. Technically straightforward	Labour intensive. Requires validation in other cell types.	1.4-4.4	(Warren et al. 2010)
	miRNA infection/transfection		Low efficiency. Need more studies	0.002	(Miyoshi et al. 2011)

	Piggy Bac		Low efficiency. Requires further validation.	0.02-0.05	(Kaji et al. 2009, Woltjen et al. 2009, Mali et al. 2010)
	Minicircle vector		Low efficiency. Only performed in neonatal fibroblasts. Requires further validation.	0.005	(Narsinh et al. 2011)
	Episomal plasmid		Multiple transfection	0.0003-0.1	(Yu et al. 2007, Hu et al. 2011, Chen et al. 2011)
	Protein transduction	Footprint-free iPSCs Direct delivery of transcription factors Avoid complication of nucleic-acid-based delivery	Low efficiency. Can be technically challenging	0.0001-0.0006	(Zhou et al 2009, Kim et al. 2009)
	Stimulus-trigged acquisition of pluripotency (STAP)	No introduction of transcription factors or nuclear transfer	Required further validation Limited capacity of self-renewal when compared to ESCs		(Obokata et al. 2014)

Table 1.3: iPSC reprogramming methods. The table is updated from Malik and Mahendra (2013), Sachamitr *et al*. (2014).

As both mouse and human adult fibroblasts had previously been demonstrated to be amenable to reprograming via nuclear transfer (Wakayama et al. 1998, Cowan et al. 2005, Tada et al. 2001), these cells were chosen in the first attempts at inducing reprogramming of somatic cells into pluripotent cells (Takahashi and Yamanaka 2006, Takahashi et al. 2007). In addition, Fibroblasts can be cultured in the same culture conditions as ESCs and can also serve as a feeder layer for ESCs, making them a suitable candidate for the first attempts at reprogramming (Maherali and Hochedilinger 2008). Later, there has been many examples of generating iPSCs from different somatic cells such as liver cells (Aoi et al. 2008), neural progenitor cells (Emini et al. 2008), pancreatic cells (Stadtfeld et al. 2008), lymphocytes (Hanna et al. 2008) and human keratocytes (Aasen et al. 2008). These studies reported different reprogramming efficiencies and using different reprograming factors depending on the cell type. Therefore, factors that should be taken into consideration when choosing the type of donor cell include the age of the donor as cells from older individuals are more likely to have genetic alternations than younger cells and the accessibility of the donor cell type.

1.1.3.3 Characterisation of iPSCs

Regardless of the source of somatic cells or the reprogramming method used in their generation hiPSCs have certain characteristics, which they share with ESCs. Morphologically, they are small, with a large nucleus to cytoplasmic ratio, prominent nucleoli and they form compact colonies with a clear edge (Amit and Itskovitz-Eldor 2012, Bosnakovski et al. 2012, Brouwer et al. 2016). Important growth properties are self-renewal and high-levels of telomerase activity. As with ESCs, they express specific cell surface antigens such as the glycolipid antigens SSEA-3 and SSEA-4, the keratin sulfate antigens TRA-1-60 and TRA-1-81, and express transcription factors including NANOG, OCT4, GDF3 and REX1 (Chan et al. 2009, Boulting et al. 2011, Singh et al. 2012). Cell surface antigen expression is routinely used for hiPSC characterization with the glycolipid antigen SSEA-1, a marker of differentiation, also normally included as a negative control (Adewumi et al. 2007, Martí et al. 2013). The defining feature of hiPSCs is their ability to differentiate into any cell type. Pluripotency of newly generated hiPSCs is functionally tested using the *in vivo* teratoma assay or *in vitro* embryoid body formation

assay, both of which evaluate the differentiation capacity of hiPSCs into each of the three germ layers (Prokhorova et al. 2009, Gutierrez-Aranda et al. 2010, Martí et al 2013). There is some debate regarding the necessity of the teratoma assay for hiPSC characterization, in many studies the *in vitro* assays are seen as sufficient evidence for pluripotency, however, other groups argue that the behavior of cells in tissue might be completely different from those in tissue culture and suggest that i*n vivo* testing is therefore necessary (Marti et al. 2013).

1.1.3.4 iPSCs as an experimental model

Over the past decade, an enormous number of studies have utilised hiPSCs to investigate everything from basic biology to regenerative medicine. Basic research is focussed on understanding the mechanisms and signalling pathways controlling reprogramming and improving the methods used to generate hiPSCs as well as the optimisation and standardisation of their differentiation into different cell and tissue types.

The ability of iPSCs to differentiate into any tissue type has made iPSCs a powerful model for the study of disease and as a way to develop treatments. An early example of hiPSCs as a disease model involved the generation of hiPSCs from fibroblasts of amyotrophic lateral sclerosis patients with subsequent differentiation into motor neuron cells (Dimos at el. 2008). Subsequently, hiPSCs have been generated from patients with a variety of genetic diseases with either Mendelian or complex inheritance to provide a model for disease investigation and drug development, such as using hiPSCs to test toxicity and the dosage of prior to human studies (Guo at el. 2011, Park at el. 2008). For polygenic disease, several studies have investigated patient-specific hiPSCs for Alzheimer disease (Israel et al. 2012, Yagi et al. 2011, Yahata et al. 2011), Parkinson's disease (Devine et al. 2011), and schizophrenia (Brennand et al. 2011). In mouse studies, a number of miPSCs were successfully conducted the transplantation an attempt to treat diseases. For example, Hanna *et al.* (2007) have reported that humanized sickle cell anaemia mouse model was rescued after transplantation of corrected hematopoietic progenitors produced from miPSCs. Recently, murine model of myocardial infraction has

appeared improvement in left ventricular function by injected iPSC-derived cardiomyocytes (Rojas et al. 2017).

For clinical application, hiPSCs have promised solving the problem of the immune rejection and ethics that associated with clinical application of the human ESCs. The first clinical trial was in human in 2014 when the pigment epithelial cells differentiated from human fibroblast iPSCs were implanted in Japanese woman eye to treat the age-related macular degeneration (AMD) (Cyranoski 2014). Therefore, hiPSCs have been expected to treat many incurable diseases such as neurodegenerative disease (Ross and Akimov 2014), spinal cord injury (Nakamura and Okano 2013), diabetes (Holditch et al. 2014), heart failure (Hsiao et al. 2013) and retinal diseases (Ramsden et al. 2013). However, further investigation in term of safety and efficiency in generation of hiPSCs is still required as mentioned above. More recently, hiPSC banks, consisting of hundreds of lines derived from both patients and controls have become available as a resource for the wider scientific community.

1.1.3.5 Genetic instability of iPSCs

Several studies have reported examples of genetic instability in hiPSCs. including aneuploidy (an abnormal number of chromosomes), copy number variation (CNV) (deletions and duplications), inversions and single nucleotide variants (Chin et al. 2009, Mayshar et al. 2010, Ruiz et al. 2013). While the occurrence of genetic alterations in iPSCs has the potential to cause confounding effects in the interpretation of in vitro iPSC data they are most important in regenerative medicine where their presence could have serious safety implications There are three different causes of these genetic alterations; 1) they were already present in the parental somatic cells and became fixed during clonal selection in the course of iPSC generation, 2) they were induced as part of the reprogramming process or 3) they arose spontaneously during prolonged culture (Yoshihara et al. 2017). Regardless of the procedures used for generation of iPSCs, the central question that has been asked whether the mutations in these cells were mainly arisen from spontaneous mutation events which occur in their original cells or whether they were arisen during reprogramming process. More recently, it has been argued that about 75% of the mutation happened during the reprogramming special at initiation steps

(Ji et al. 2012, Sugiura et al. 2014). Sugiura and colleagues have also demonstrated that the rate of the genetic alterations were much higher in ESCs than in the iPSCs (Sugiura et al. 2014).

The presence of genetic abnormalities can be detected on a genome-wide scale by a number of approaches. One of the most widely use and conventional methods in genetic evaluation laboratories and researches for routine cell line quality control is the Giemsa (G)-banding technique that can identify changes in chromosome content and chromosome structure (inversions and translocations) as well as larger duplications and deletions (Yunis 1979, Meisner and Johnson 2008). Large chromosome abnormalities (e.g. aneuploidy) can be detected using fluorescent *in situ* hybridization (FISH), which involves hybridizing fluorescently labelled DNA probes to complementary target sequences on chromosomes (Bauman et al. 1980). Chromosome painting, a type of FISH, involves the use of probes against the whole genome resulting in fluorescent labelling of the entire length of each chromosome (Pinkel et al. 1998). As well as allowing chromosome content to be determined chromosome painting also allows the detection of chromosomal rearrangements including translocations, inversions or the presence of extra-chromosomal fragments of DNA. (Pinkel et al. 1998). G-banding, FISH and chromosome painting are only effective at identifying large changes in chromosome structure and do not detect smaller CNVs or single base changes. DNA microarrays including comparative genomic hybridization arrays (CGH) (Kallioniemi et al. 1992) and single nucleotide polymorphism arrays (Wang et al. 1998) allow high resolution recognition of CNVs on a genome-wide scale. More recently, WGS has been utilised to detect genetic variation with single nucleotide resolution across entire genome and detect single base changes as well as small CNVs and inversions. Such high-throughput technology has detected various mutations in hiPSCs, but most of these are likely to be benign (Bhutani et al. 2016).

1.1.4 Project hypothesis

GWAS has recently identified associations between mean leukocyte telomere length and loci containing genes with no known role in telomere maintenance. This study will use CRISPR/Cas9 mediated genome editing of hiPSCs to investigate the hypothesis that;

> *Knockout of genes associated with telomere length in hiPSCs will result in loss of telomere maintenance and confirm the causal gene at the novel loci*

1.1.5 Project Aims

The aims of this study are;

- To establish and optimize hiPSC culture and CRISPR/Cas9 mediated genome editing.
- To determine the effects of telomerase deficiency on telomere maintenance and hiPSC phenotype
- To determine the identity of the causally associated gene at the novel telomere length associated *ACYP2* locus.

Chapter 2

Materials and Methods

2.1 Materials and methods

All experiments were conducted at room temperature unless stated otherwise. Required materials are listed before each method.

2.1.1 Cell culture

2.1.1.1 Cell lines

- Human iPSCs (GM23720*B) were purchased from the NIGMS collection at the Coriell institute for medical research/America.
- K562 cell line was purchased from ECACC, 89121407, Sigma Aldrich.

2.1.1.2 Human iPSC culture

2.1.1.2.1 Preparation of Matrigel™-coated plates

Required materials:

- Dulbecco's Modified Engle's Medium, DMEM (Sigma)
- BD Matrigel™ hESC-qualified Matrix, Lot. 325691 (BD Biosciences)
- 6-well plates (Cell star®, greiner bio-one)
- Ice box
- Parafilm (Demis)

Protocol:

11µl of Matrigel™, BD Biosciences, was added to 1ml of cold DMEM medium and mixed. This mixture should be prepared on ice.

1ml/well Matrigel™/DMEM solution was immediately used to coat 6-well plates. The plates were swirled gently to spread the BD Matrigel™ solution evenly across the plates and used straight away or sealed with parafilm for storage at 2-8°C for up to 7 days. For stored plates, the parafilm was removed and the plates incubated at 37°C for 30 mins

before use. Finally, BD Matrigel™ solution was removed by aspiration before plating cells. **Table 2.1** lists the dilution of Matrigel™ for each type of vessel used in this study.

Type of vessel	The Volume of Matrigel™/ DMEM medium for each well	Overall volume of the medium required for whole plate
6-well plate	11µl/1ml	6ml
96-well plate	0.363µl/33µl	3.3ml
24-well plate	2.53µl/230µl	5.6ml
90 mm petri-dish	55µl/5ml	8ml

Table 2.1: The dilution of the Matrigel™ for each type of plate used in the study.

2.1.1.2.2 Reviving human iPSCs

Required materials:

- mTesR™1 medium, Lot. 14B52766D (STEMCELL™) with 5x supplement, Lot. 13L52098D, (STEMCELL™)
- ROCK inhibitor Y-27632, Lot. Doo155678, (STEMCELL™)
- BD Matrigel™ -coated 6-well plates prepared as in the **section 2.1.1.2.1**
- incubator at 37°C with 5% CO_2

Protocol:

The GM23720*B iPSC line (Coriell institute, America) was removed from liquid nitrogen and thawed quickly with constant, moderate agitation in a 37°C water bath until thawed. The cells were transferred to 15ml sterile conical tubes containing 1mL warm

mTesR™1 medium supplemented with a final concentration of 10µM ROCK inhibitor Y-27632. Following this step, cells were centrifuged at 300xg for 2 min at room temperature. The medium was then aspirated and the cells resuspended in 2ml warm mTesR™1 including the 5x supplement provided with the media. The iPSCs were then plated on a BD Matrigel™-coated 6-well plate and incubated at 37°C with 5% CO_2.

2.1.1.2.3 Enzymatic passaging of iPSCs

Required materials:

- DMEM (Sigma)
- mTesR™1 medium, Lot. 14B52766D (STEMCELL™) with 5x supplement, Lot. 13L52098D, (STEMCELL™)
- Dispase, CAT# 07923, (STEMCELL™)
- Scraper, NY1483, (Costar® corning)
- BD Matrigel™-coated 6-well plates as prepared in the **section 2.1.1.2.1**
- CKX41 microscope (Olympus)
- Incubator at 37°C with 5% CO_2

Protocol:

iPSCs were cultured at 37°C with daily media change and were passaged when they reached 70-80% confluence. The media was removed from the plate by aspiration and the cells washed twice with 1ml DMEM. 1ml/well of 1mg/ml Dispase II was added and the plates incubated at 37°C with frequent observation at 10x magnification with an Olympus CKX41 microscope. When the edges of colonies appeared to be slightly folded back, but the colonies were still attached to the plate (approximately 7 mins) the Dispase II was aspirated. The cells were washed twice with 2ml DMEM and 2ml/well mTesR™1 added. iPSC colonies were then carefully detached using a cell scraper. The cells were transferred by pipetting to a 15ml conical tube and pelleted by centrifugation at 300xg for 5 mins at room temperature. The supernatant was removed and cells were gently resuspended in 5-6 ml of fresh mTesR™1. 1ml/well of cells were replated on Matrigel™-coated 6-well plates and the volume supplemented with an additional 1ml mTesR™1.

2.1.1.2.4 Non-enzymatic passaging of iPSCs

Required materials:

- ReLeSR™ enzyme-free reagent, Cat# 05872, (STEMCELL™)
- DMEM (Sigma)
- mTesR™1 medium, Lot. 14B52766D (STEMCELL™) with 5x supplement, Lot. 13L52098D, (STEMCELL™)
- BD Matrigel™-coated 6-well plates as prepared in the **section 2.1.1.2.1**
- Incubator at 37°C with 5% CO_2

Protocol:

ReLeSR™ enzyme-free reagent was used to select and passage the iPSCs. The iPSCs were cultured at 37°C and were passaged when they reached 70-80% confluence. The media was removed by aspiration and the cells washed twice with 1ml DMEM and then 1ml/well ReLeSR™ added. After 1 min, the ReLeSR™ reagent was aspirated, leaving a thin film of reagent over the cells, which were then incubated for 5 mins at 37°C, followed by addition of 1ml/well mTeSR™1 medium. The plate was then tapped gently on the side for 30-60 sec. The resuspended cell suspension was transferred to a 15ml conical tube. The cell aggregate should be around 50-200µm for plating and this was achieved by pipetting the suspension up and down using a 2ml serological pipette. 5-6ml mTeSR™ 1 medium was added to the cell suspension and 1ml/well was replated on 6-well Matrigel™-coated plates. An additional 1ml of mTeSR™1 was added to each well to achieve a final volume of 2ml/well.

2.1.1.3 K562 culture

Required materials:

- RPMI 1640 medium (Sigma-Aldrich)
- L-Glutamine (Hyclone)
- Foetal calf serum, FCS (Gibco)
- Culturing flask (CELLSTAR®-greiner bio-one)

Protocol:

K562 is an immortalized myelogenous leukaemia cell line and was used as a calibrator for telomere length measurement. Briefly, cells were cultured in suspension in RPMI 1640 medium supplemented with 2mM L-glutamine and 10% FCS and maintained between $1x10^5$ to $1x10^6$ cells/ml at 37˚C with 5% CO_2.

2.1.1.4 Cryopreservation of iPSCs

Required materials:

- mTesR™1 medium, Lot. 14B52766D (STEMCELL™) with 5x supplement, Lot. 13L52098D, (STEMCELL™)
- ROCK-inhibitor Y-27632, Lot. Doo155678, (STEMCELL™)
- DMSO, Dimethyl sulfoxide, Lot. SZBD287OV, (Sigma)
- Knock out serum replacement, Lot. 1392634, (Gibco)
- Cryovials (Sarstedt)
- isopropanol freezing container (Nalgene®, Thermo Scientific)
- Ice box

Protocol:

iPSCs were collected as described in the **section 2.1.1.2.4**, but cell pellets were resuspended in cooled cryoprotectant solution (50% mTeSR™1, 40% Knock out Serum Replacement, 10% DMSO, 1µl/ml of 10mM of Y-27632 ROCK-inhibitor and transferred to cryovials on ice. The cryovials were stored in isopropanol freezing container at -80°C overnight and subsequently transferred to liquid nitrogen for long-term storage.

2.1.1.5 Cell counting

Required materials:

- Dulbeco's phosphate buffer saline, PBS
- 1xTrypsin/EDTA (Lonza) was prepared by dissolving 5g/l trypsin and 2g/l EDTA to give 10x working dilution
- DMEM (Sigma)
- Foetal calf serum, FCS, (GIBCO)
- Incubator at 37°C with 5% CO_2
- 4% Trypan blue solution (Hyclone)
- Haemocytometer slide with cover (Hawksley)
- CKX41 microscope (Olympus)

Protocol:

An exact single cell suspension per ml was required in particular experiments such as clonal isolation, flow cytometry and karyotyping. iPSCs cultured in 6-well plates were washed twice with PBS and trypsinized using 1ml/well 1x trypsin/EDTA at 37°C for 10-12 mins, then neutralizied by adding DMEM medium supplemented with 10% FCS, transferred to a 15ml conical tube and centrifuged at 300xg for 5 mins at room temperature. The pellet was resuspended with either 1 or 2ml mTesR™1 medium and 10µl of the suspension mixed with 10µl of 4% trypan blue solution. 10µl of this was added to the haemocytometer (cleaned with 70% IMS and dried thoroughly). The average of the number of the cells in two or four outer squares observed at 10x magnification with CKX41 microscope was multiplied by $2x1x10^4$ to achieve the number of the cells per ml

of the suspension. This value was used to calculate the concentration of the cells required for the experiments.

2.1.1.6 Optimization of iPSC clonal dilution

Required materials:

- Matrigel™-coated 90mm petri-dishes as described in the **section 2.1.1.2.1**
- mTesR™1 medium, Lot. 14B52766D (STEMCELL™) with 5x supplement, Lot. 13L52098D, (STEMCELL™)
- Conditioned mTesR™1 medium (preparation described below)
- Dulbecco's PBS
- 1x Trypsin/EDTA (Lonza)
- FCS (GIBCO)
- DMEM (Sigma)
- 0.22µm membrane (Millex®GP)
- Syringe (BD Plastipak™)
- 4% Trypan blue solution (Hyclone)
- Haemocytometer (Hawksley)
- Inverted microscope (Evos®)
- Incubator at 37°C with 5% CO_2

Protocol:

Three different concentrations of iPSCs single cell dilutions (10^3, 10^4 and 10^5 cells/ml) were tested in order to isolate clones from individual cells. Dilutions were cultured in triplicate in Matrigel™-coated 90mm Petri dishes containing 8ml of fresh mTesR™1 medium and incubated at 37°C with 5% CO_2. The medium of the 10^3 and 10^4 cells/ml dilutions was replaced the following day with mTesR™1 medium consisting of 50% fresh mTesR™1 medium and 50% conditioned mTesR™1 (prepared from mTesR™1 collected from iPSCs at 40-50% confluence and filtered through a 0.22µm membrane filter and stored at -20°C freezer for up to 5 months). On day 3, the media of the 10^3 and 10^4 cells/ml

dilutions was replaced with 75% fresh and 25% mTesR™1 medium. From day 3, media was replaced with 100% fresh mTesR™1 medium. The media of the 10^5 cells/ml dilution was replaced every two days with fresh mTesR™1 medium. The iPSCs were monitored daily until the colonies reached approximately 0.5 to 1mm in diameter. The media change was then replaced daily with fresh mTesR™1 medium. 10^3 cells/ml was identified as the optimal seeding density as there was sufficient space between colonies once they had reached the appropriate size for picking.

2.1.1.7 Optimized clonal isolation

Required materials:

- Matrigel™-coated 90mm petri-dishes, described in **the section 2.1.1.2.1**
- Matrigel™-coated 96-well plates (described in **the section 2.1.1.2.1**)
- mTesR™1 medium, Lot. 14B52766D (STEMCELL™) with 5x supplement, Lot. 13L52098D, (STEMCELL™)
- Dulbecco's PBS
- Conditioned mTesR™1 medium (as prepared in **the section 2.1.1.6)**
- 1x Trypsin/EDTA (Lonza) Foetal calf serum, FCS (GIBCO)
- DMEM (Sigma)
- Inverted microscope (Evos®)
- 4% Trypan blue solution (Hyclone)
- Haemocytometer (Hawksley)
- Incubator at 37˚C with 5% CO_2

Protocol:

After electroporation iPSCs were cultured in Matrigel™-coated 6-well plates as described in **section 2.1.1.2.4**. When the cells reached 70-80% confluence, they were trypsinized using 1x trypsin/EDTA for 10-12 mins and cell-count determined as described in **section 2.1.1.5**. 10^3 cells/ml were seeded in Matrigel™-coated 90mm Petri-dishes in triplicate as described in **section 2.1.1.6**. When the colonies reached 2-5mm in diameter, the colonies were visualized under the 10x objective of an Evos® inverted

microscope and picked using a 200µl pipette adjusted to 45µl and transferred to Matrigel™-coated 96-well plates contained 120µl mTeSR™1 medium. The cells were incubated at 37˚C with 5% CO_2 and monitored daily. The day after picking, media was replaced with mTesR™1 medium consisting of 50% fresh and 50% conditioned mTesR™1 medium. On the second day media was replaced with 75% fresh and 25% conditioned medium and then media replaced daily with 100% fresh mTesR™1 medium. When the colonies reached 80% confluence, they were passaged to new 96-well plates and frozen. A portion of the iPSCs was retained from each clone for DNA extraction and sequencing.

2.1.1.8 Freezing iPSCs in 96-well plates

Required materials:

- 96-well plates (Cell star®, greiner bio-one)
- Accutase™, 07920, (Stem cell technology)
- Dulbecco's PBS
- mTesR™1 medium, Lot. 14B52766D (STEMCELL™) with 5x supplement, Lot. 13L52098D, (STEMCELL™)
- mFreSR™, 05855, (STEMCELL™)
- Blue tissue roll
- Parafilm (Demis)
- Styrofoam container

Protocol:

When the clones cultured in 96-well plates reached 80% confluence the medium was aspirated and the cells washed twice with 150µl/well PBS. 33µl Accutase, diluted 1:3 in PBS, was added to each well and plates incubated at 37˚C for 7-10 mins until the cells detached and 165µl fresh mTeSR™1 medium was added. 132µl/well of iPSCs (the remaining cells were passaged or lysed for DNA) was transferred into uncoated 96-well plates with 140µl/well of mFreSR™ freezing medium and gently mixed by pipetting. The plate was wrapped using parafilm and blue tissue roll and transferred to Styrofoam containers at -80˚C.

2.1.1.9 Passaging of iPSCs in 96-well plates

Required materials:

- Matrigel™-coated 24-well plate, prepared as described in **the section 2.1.1.2.1**, 24-well plate, (Castar® corning, lifescience)
- Accutase, 07920, (STEMCELL™)
- Dulbecco's PBS
- mTesR™1 medium, Lot. 14B52766D (STEMCELL™) with 5x supplement, Lot. 13L52098D, (STEMCELL™)
- Incubator at 37˚C with 5% CO_2

Protocol:

When clones reached 70-80% confluence the media was removed from the plate by aspiration and the cells washed twice with 150µl/well PBS. 33µl/well Accutase, diluted 1:3 in PBS, was added to the cells and the plates incubated for 5-10 min at 37˚C. 165µl/well mTeSR™1 medium was added and the plates tapped gently on the side for 30-60 sec. 66µl/well of the cell suspension was then transferred to Matrigel™-coated 24-well plates with 500µl fresh mTeSR™ 1 medium. The plate was incubated at 37˚C with 5%CO_2. After 8h the media was replaced with 500µl mTeSR™1 medium in order to avoid any effects of the remaining Accutase. The plate was then incubated at 37˚C with 5% CO_2 with daily feeding until the colonies reach the appropriate confluence for further passaging.

2.1.1.10 Reviving iPSCs frozen in 96-well plates

Required material:

- Matrigel™-coated 24-well plates, as described in **the section 2.1.1.2.1** (24-well plate from Costar®, corning, lifescience)
- Matrigel™-coated 6-well plate, as described in **the section 2.1.1.2.1**
- mTesR™1 medium, Lot. 14B52766D (STEMCELL™) with 5x supplement, Lot. 13L52098D, (STEMCELL™)

- Dulbecco's PBS
- Incubator at 37°C with 5% CO_2
- 70% Ethanol (Sigma)

Protocol:

1ml/well mTeSR™1 in Matrigel™-coated 24-well plates had been prepared before removing 96-well plates from -80°C storage. The 96-well plate was cleaned with 70% Ethanol, parafilm removed and cells thawed by incubating at 37°C for 15-20 mins. The plate was removed from the incubator when wells had only small piece of frozen material remaining. 100µL of warm mTeSR™1 medium was added to each well and the mixture transferred to the Matrigel™-coated 24-well plate. The plate was then incubated at 37°C with 5% CO_2. After 8h the media was replaced with 500µl mTeSR™1 medium. Medium was replaced every 24h. When cells reached approximately 80% confluence they could be expanded into Matrigel™-coated 6-well plates (**section 2.1.1.11**).

2.1.1.11 Expansion of iPSCs from 24-well to 6-well plates

Required materials:

- Matrigel™-coated 6-well plates as prepared in **the section 2.1.1.2.1**
- mTesR™1 medium, Lot. 14B52766D (STEMCELL™) with 5x supplement, Lot. 13L52098D, (STEMCELL™)
- ReLeSR™ enzyme-free reagent, Cat# 05872, (STEMCELL™)
- Dulbecco's PBS

Protocol:

When iPSCs reached 80% confluence the colonies were transferred to Matrigel™-coated 6-well plates. After washing cells twice with 500µl PBS, 300µl of ReLSR™ enzyme was added onto the colonies for 1 min at the room temperature. The ReLSR™ was then aspirated and the plates incubate for 5 min at 37°C. 500µl of mTeSR™1 medium was added to the cells and the plates were tapped gently for 30 secs. The colonies were then resuspended by pipetted twice before being transferred into Matrigel™-coated 6-well

plates. The plates were then incubated at 37°C with daily feeding and monitoring until the confluence of the colonies reached to 70-80% confluence (~4-7 days) and then subcultured at 1:5 or 1:6 in 6-well plates.

2.1.1.12 Extended culture of iPSCs

Long-term culture was required for the investigation of telomere length. Following expansion, iPSCs were maintained in Matrigel™-coated 6-well plates until they had been in culture for 63-64 days. At each passage a portion of cells for each clone was banked and collected for DNA, RNA or flow cytometry.

2.1.2 Quality control of iPSCs

2.1.2.1 Flow cytometry

2.1.2.1.1 Fixing iPSCs for flow cytometry

Required materials:

- BD stemflow™ human pluripotent stem cells sorting and analysis kit, cat no. 560461, (BD Bioscience)
- Paraformaldehyde, PFA, Lot. 072K9283, (Sigma)
- Bovine serum albumin, BSA, Lot #. 040M1673, (Sigma)
- Flow cytometer system with AW08603 software version (Gallios)

Protocol:

Flow cytometric analysis was performed every 10 passages to confirm iPSC phenotype. To achieve a single cell suspension, iPSCs were detached using 1ml/well 1x trypsin/EDTA at 37°C for 10-12 mins and then neutralized by adding DMEM supplemented with 10% FCS and transferred to a 15ml conical tube and centrifuged at 300xg for 5 mins. Cells were counted (as described in **section 2.1.1.5**) before being resuspended in 4% PFA at a concentration of $1x10^7$cells/ml and incubated for 20 mins at

room temperature. The cells were then washed twice with stain buffer (3% BSA in PBS) and resuspended in stain buffer.

The fixed cells were stained with isotype controls or antibodies against cell surface markers (BD stemflow™ human pluripotent stem cells sorting and analysis kit) (**Table 2.2**). The positive and negative beads were also prepared to assist in optimizing application setup and calculating fluorescence compensation (**Table 2.3**). AW08603 software (Gallios) was used to analyse the data. The system has 3 lasers; blue (488), red (638) and violet (405) and there are filters to detect up to 10 emitted wavelengths from fluorescent dyes. **Table 2.4** lists the fluorescent dyes, lasers and filters used in this study.

	Fixed cells	
Antibodies of surface markers	**Specific stain 100µl (1x10^6 cells)**	**Isotype control 100µl (1x10^6 cells)**
SSEA-1	20µl	-
SSEA-3	20µl	-
TRA-1-81	20µl	-
FITC isotype control		20µl
PE isotype control		20µl
Alexa Fluor® 647 isotype control		20µl

Table 2.2: Antibodies for identification of surface markers on hiPSCs. SSEA; stage-specific embryonic antigen, TRA; Trafalgar antigen, FITC; fluorescent isothiocyanate and PE; phycoerythrin.

	Volume added to each tube			
Continent	**negative**	**FITC**	**PE**	**Alexa Fluor® 647**
Negative beads	40µl	40µl	40µl	40µl
Anti-mouse beads	-	40µl	-	40µl
Anti-rat beads	-	-	40µl	-
SSEA-1	-	20µl	-	-
SSEA-3	-	-	20µl	
TRA-1-81	-	-	-	20µl

Table 2.3: The negative and positive beads used to optimize application setup and calculate fluorescent compensation.

Antibody	**Fluorescent dye**	**Laser**	**Filter**
SSA-1	FITC	488nm(blue)	530nm
SSEA-3	PE	561nm(Yellow/green)	582nm
TRA-1-81	Alexa Fluor® 647	647nm(red)	665nm

Table 2.4: Details of fluorescent dye, laser and filter used for each antibody. FITC; fluorescent isothiocyanate and PE; phycoerythrin.

2.1.2.2 Karyotypic analysis of iPSCs

Required materials:

- Nocodazol (Sigma-Aldrich)
- Chromosome resolution additive, CRA (Genial Gentics)
- KCL (Sigma)
- Methanol (Sigma)
- Acetic acid (Sigma)
- Giemsa, #GS500, (Sigma)
- 4′, 6-diamidino-2-phenylindole (DAPI), Lot. 983827 (Invitrogen)
- Mount media, Cat. C9368, (Sigma-Aldrich)
- Fluorescence microscope (Leica) with ImageJ/FIJI program
- Dulbecco's PBS
- 1x Trypsin/EDTA (SLS, Lonza) was prepared by dissolving 5g/l trypsin and 2g/l EDTA to give 10x dilution
- Foetal calf serum, FCS (GIBCO)
- Dulbecco's Modified Engle's Medium, DMEM (Sigma)

Protocol:

Karyotypic analysis was performed every 10 passages. On the day before karyotypic analysis, with iPSCs at 60-70% confluence, the mTesR™1 medium was replaced with fresh mTesR™1 containing 0.1µl/ml medium of Nocodazol diluted 1:1000 with dH_2O and incubated at 37°C for 24h. 20µl of chromosome resolution additive diluted 1:10 with injection quality water was then added and the plate incubated at 37°C for 90 min. The cells were then washed twice with PBS and trypsinized using 1ml/well 1x trypsin/EDTA at 37°C for 10-12 mins. DMEM containing 10% FCS was added and the cells centrifuged at 300xg for 5 min. The pellet and 500µl retained medium was resuspended with 5ml hypotonic solution (0.056M KCL in dH_2O) and incubated at room temperature for 30 min. The supernatant was gently aspirated leaving 1ml for resuspension and then mixed with 1ml Carnoys fixation solution (75% Methanol and 25% Acetic acid), centrifuged again at 300xg for 5 min and the pellet resuspended in 3ml Carnoys fixative. This step

was repeated twice before a final resuspension in 500µl of Carnoys fixative. The suspension has a slight milky appearance at this stage.

Two slides were prepared for each sample and placed in a sloping position in a container in a water bath at 37°C. A small amount of suspension was added drop-wise onto the slides and air dried. The slides were then kept overnight in the dark.

The flowing day the cells were stained with Giemsa. A 1 in 20 dilution of Giemsa in dH_2O was prepared and, 2ml-4ml was added on the cells for 30 mins. Slides were washed twice with dH_2O and air dried before mounting. In some preparations DAPI was also used to stain chromosomes for 5 mins at a 1 in 1000 dilution in PBS before Giemsa staining. A fluorescent microscope (Leica) was used to take images at 100x magnification. Finally, counting chromosome was performed on at least 10 cells per preparation.

2.1.3 iPSC transfection

2.1.3.1 Lipid-based transfection

TransIT®-2020 (Cat#MIR 5410s, Mirus) and lipofectamine®3000 (Lot: 1604127, Invitrogen) were used according to the manufacturer's instructions. For each reagent, both forward and reverse transfections were performed with pCMV-GFP (4.7kbp) was used as a reporter for transfection efficiency.

2.1.3.1.1 Preparation of TransIT®-2020 reagent:plasmid complex

Required materials:

- TransIT®-2020 (Cat#MIR 5410s, Mirus)
- Opti-MEM I (Gibco, Fisher Science)
- pCMV-GFP plasmid

Protocol:

TransIT-2020 reagent was warmed to room temperature and vortexed gently before use. 2.5µg plasmid DNA was added to 250µl of Opti-MEM I reduced serum medium was

in a 1.5ml eppendorf tube and gently mixed by pipetting. 7.5µl TransIT-2020 reagent was then added to the diluted plasmid DNA and mixed gently. Finally, the mixture was incubated at room temperature for 15-30 min to form the plasmid complex.

2.1.3.1.2 Preparation of Lipofectamine®3000 reagent:plasmid complex

Required materials:

- Lipofectamine® 3000, Lot: 1604127, (Invitrogen)
- Opti-MEM I (Gibco, Fisher Science)
- pCMV-GFP plasmid

Protocol:

7.5µl or 3.75µl of Lipofectamine®3000 was diluted in 125µl Opti-MEM I reduced serum medium. In a separate tube, 5µg pCMV-GFP was diluted in 250µl Opti-MEM I reduced serum medium and 10µl P3000™ reagent and mixed gently. The diluted Lipofectamine®3000 and diluted plasmid solutions were then mixed together and incubated at room temperature for 5 mins.

2.1.3.1.3 Lipid-based transfection of iPSCs

2.1.3.1.3.1 Forward transfection assay

When iPSCs reached 70%-80% confluence the TransIT-2020:plasmid complex or Lipofectamine3000:plasmid complex was added to the cells in a dropwise manner. The plate was gently rocked back and forth and side to side to evenly distribute the plasmid complex. The cells were then incubated at 37°C for 24h. Transfection efficiency was determined by counting GFP positive cells on an Evos® fluorescent microscope. Cells were then maintained in culture in order to assess viability following transfection.

2.1.3.1.3.2 Reverse transfection assay

When iPSCs reached 70%-80% confluence they were washed twice with DMEM and detached using Dispase II enzyme (**section 2.2.2.3**) and iPSCs resuspended in mTesR™1 medium. TransIT-2020:plasmid complex or Lipofectamine3000:plasmid complex was added to Matrigel™-coated 6-well plates and the plates rocked back and forth and side to side to evenly distribute the mixture before plating cells. After 24h transfection efficiency was measured by counting GFP positive cells using an Evos® fluorescent microscope. Cells were then maintained in culture in order to assess viability following transfection.

2.1.3.2 Optimization of nucleofection of iPSCs

Required materials:

- Amaxa nucleofector (Lonza)
- Human stem cells nucleofector® starter kit (Amaxa®)
- DMEM (sigma)
- pmaxGFP (3.4kbp) plasmid DNA
- Amaxa nucleofector (Lonza)
- Evos® fluorescent microscope

Protocol:

Nucleofection was conducted using an Amaxa nucleofector (Lonza) and Amaxa® human stem cells nucleofector® starter kit. Five different nucleofection programs (A-012, A-013, A-023, A-027 and B-016) were assessed with each solution nucleofector® starter kit. iPSCs were washed twice in DMEM and then incubated with 1x trypsin/EDTA for 10-12 mins at 37°C to achieve single cell suspension. The reaction was neutralized with DMEM medium supplemented with 10% FCS and then centrifuged at 150xg for 3 mins. The cell pellet was resuspended in mTeSR™1 medium and the concentration adjusted to 6.6-7.7x10^5cells/ml/sample. The cells were pelleted by centrifugation at 150xg for 3 mins and resuspended in 100µl of nucleofection solution. 5µg of pmaxGFP vector was added to each sample and the samples transferred into a Amaxa cuvette. The samples were then nucleofected using each program. Finally, 500µl of pre-equilibrated mTesR™1 media was

added to the cuvette. The sample was transferred to Matrigel™-coated 6-well plates containing 1.9ml mTesR™ 1 media using plastic pipettes. The nucleofection steps were conducted as quickly as possible to minimise adverse effects on the iPSCs. After 24h transfection efficiency was measured by counting GFP transfected cells using an Evos® fluorescent microscope. Cells were maintained in culture in order to assess viability following transfection.

2.1.3.3 Optimization of electroporation of iPSCs

Required materials:

- NEPA21 electroporator (Nepagene)
- EC-002S NEPA electroporation cuvettes (Nepagene)
- Evos® fluorescent microscope
- Opti-MEM I (Gibco, Fisher Science)
- pCMV-GFP

Protocol:

Electroporation was conducted using an NEPA21 electroporator (Nepagene). The electroporator has twenty programs electroporation programs different cells (T**able 2.5**) and programs 5, 6, 7, 8 and 9 were assessed with iPSCs. iPSCs were washed twice in DMEM and then incubated with 1x trypsin/EDTA for 10-12 mins at 37°C to achieve single cell suspension, the reaction was neutralized with DMEM medium supplemented with 10% FCS and then centrifuged at 200xg for 3 mins. The cell pellet was resuspended in Opti-MEM I medium and the concentration adjusted to $1x10^6$ cells/ml/sample. 5µg of DNA pCMV-EGFP plasmid was added to each sample and samples transferred to EC-002S NEPA electroporation cuvettes. The samples were then transfected using each program. Finally, 500µl of warmed mTesR™ 1 media was added to the cuvette. The samples were transferred to Matrigel™-coated 6-well plates containing 1.9ml mTesR™ 1 media using plastic pipettes. The electroporation steps were conducted as fast as possible to minimise adverse effects on the iPSCs. After 24h transfection efficiency was measured

by counting GFP transfected cells using an Evos® fluorescent microscope. Cells were maintained in culture in order to assess viability following transfection.

Type of program	Set parameters											
	Poring pulse						Transfer pulse					
	V	Length (ms)	Interval (ms)	No.	D.Rate (%)	Polarity	V	Length (ms)	Interval (ms)	No.	D.Rate (%)	Polarity
Control	Cells and DNA but no electroporation											
5	120	5	50	2	10	+	20	50	50	5	40	+/-
6	125	2.5	50	2	10	+	20	50	50	5	40	+/-
7	125	5	50	2	10	+	20	50	50	5	40	+/-
8	130	5	50	2	10	+	20	50	50	5	40	+/-
9	135	2.5	50	2	10	+	20	50	50	5	40	+/-

Table 2.5: Parameters of the 5 electroporation programs tested.

2.1.3.3.1 Transfection iPSCs with CRISPR/Cas9

Required materials:

- NEPA21 electroporator (Nepagene)
- EC-002S Nepa electroporation cuvette (Nepagene)
- DMEM (sigma)
- Evos® fluorescent microscope
- 1x Trypsin/EDTA (SLS, Lonza) was prepared by dissolving 5g/l trypsin and 2g/l EDTA to give 10x dilution
- FCS (Gibco)
- Opti-MEM I (Gibco, Fisher Science)
- *ACYP2* gRNA plasmid (Barcode# 6197788), (Horizon Discovery Limited)
- *TSPYL6* gRNA plasmid (Barcode# 6197789), (Horizon Discovery Limited)
- *TERT* gRNA plasmid (Barcode# 6197790), (Horizon Discovery Limited)
- CMV-Cas9 expression plasmid (Barcode# 6176940), (Horizon Discovery Limited)

Protocol:

CRISPR/Cas9 reagents were transfected into iPSCs using a NEPA21 electroporator (Nepagene) and electroporation program 7 as described in **section 2.1.3.3**. Cas9 only transfection was performed as a control. Concentrations of plasmids are detailed in the **table 2.6**.

	Sample	Amount of plasmid µg
	Cas9 only as control	10µg
knock out	Cas9 *g*RNA (*TERT*)	5 µg 5µg
	Cas9 gRNA (*ACYP2*)	5µg 5µg
	Cas9 gRNA (*TSPYL6*)	5µg 5µg

Table 2.6: Concentration of CRISPR/Cas9 plasmids used for transfection.

2.1.4 Molecular biology

2.1.4.1 Primer design

Primers used in this study are listed in **table 2.7**. Intron spanning primer pairs for each gene of interest were designed against the cDNA sequence from Ensembl (www.ensembl.org/index.html) Primer3 version 4.0.0 (primer3.ut.ee). Primers were checked using the *in silico* PCR tool on the UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgPcr). Primers for amplification of gDNA for Surveyor assay were designed by Horizon Discovery Ltd. All primers were ordered from Eurofins.

Name	Assay	Sequence (5́ to 3́)	Annealing temperature (Tm°C)	GC content %	Product size
***TSPYL6* F**	**qRT-PCR**	AGTGAAGGAGCTCAGACACC	59.0	55	**202bp**
***TSPYL6* R**	**qRT-PCR**	GCAGATGACATGTCGGTTCC	59.0	55	
***ACYP2* F**	**qRT-PCR**	CTAGTTCTCGCATTGACCGC	57.0	55	**217bp**
***ACYP2* R**	**qRT-PCR**	GCACGTACAGTCGCTTAGC	57.0	57.9	
***TERT* F**	**qRT-PCR**	CGTCACATCCACCTTGACAA	59.0	50	**212bp**
***TERT* R**	**qRT-PCR**	CGTTCCGCAGAGAAAAGAGG	59.0	55	
***TSPYL6* F**	**Surveyor/ Cloning**	TGAGGAACGTTGACGCTACC	61.2	55	**691bp**
***TSPYL6* R**	**Surveyor/ Cloning**	GATGGACTCCAGGGGGTTCA	64.5	60	
***ACYP2* F**	**Surveyor/ Cloning**	GCAGTCTCATTTGCCGCTTC	63.3	55	**765bp**
***ACYP2* R**	**Surveyor/ Cloning**	GCCAGACATTACAAACTTGGCA	62.3	45	
***TERT* F**	**Surveyor/ Cloning**	CTCCCCTTCCTTTCCGCG	65.3	67	**590bp**
***TERT* R**	**Surveyor/ Cloning**	CACCGTGTTGGGCAGGTAG	62.5	63	

Table 2.7: Primers used in this study. F and R indicator to forward and reverse primer respectively.

2.1.4.2 DNA isolation from Eukaryotic cells

Materials required:

- Mammalian DNA Miniprep Kit (Sigma GenElut)
- 95-100% Ethanol for dilution the washing buffer provided in the kit (Sigma)
- Nanodrop 8000 (Thermo Scientific)

Protocol:

Genomic DNA extraction was performed using the Sigma GenElut Mammalian DNA Miniprep Kit as described in the manufacturer's instructions. Briefly, cell pellets containing approximately $1x10^6$ cells were resuspended in resuspension solution and lysed by incubation with proteinase k (20mg/ml). The cell lysates were added to a GenElute Miniprep binding column and centrifuged at 12,000xg for 1 min. The column was washed twice and DNA eluted in 50-200µl Elution solution. The eluted sample was added back to the same binding column and centrifuged at 12,000xg for 1 min. To increase elution efficiency the column was then incubated at room temperature for 5 mins and then eluted by centrifugation at 12,000xg for 1 min. The concentration and purity DNA of sample were measured using a Nanodrop 8000. The samples were stored at -20°C for further experiments.

2.1.4.3 Extraction of gDNA directly from the iPSCs cultured in 96-well plate

Required materials:

- Direct PCR lysis reagent, Cat#302-C (Viagen)
- Accutase™, Cat#07920, (Stem cell technology)
- Proteinase K, LOT#SLBK1594V (Sigma)
- 96-well PCR plates and adhesive film seals (Geneflow)
- Nanodrop 8000 (Thermo Scientific)

- G_Storm GS4 thermal cycler (G-Storm)

Protocol:

Genomic DNA was extracted iPSCs cultured in 96-well plates using direct PCR lysis reagent. iPSCs were washed twice with 150µl PBS, following by dissociating colonies with 33µl/well Accutase™ diluted in 1:3 PBS at 37°C for 5-10 mins. 165µl/well mTeSR™ was added to the cells. 66µl of cell suspension was transferred to a PCR plate and mixed with 50µl/well lysis buffer contained proteinase K 0.24mg/ml. the PCR plate was sealed with PCR film and centrifuged at 1500 rpm for 3 sec. The plate was then incubated in a G-Storm GS4 thermal cycler at 55°C for 15, 85°C for 45 mins and 10°C on hold.

2.1.4.3.1 Ethanol precipitation of DNA

Required materials:

- Sodium acetate (Sigma)
- 100% Ethanol (Sigma)
- DNA Miniprep column (GenElut Mammalian DNA Miniprep Kit, Sigma)
- Column preparation solution (GenElut Mammalian DNA Miniprep Kit, Sigma)
- Washing buffer (GenElut Mammalian DNA Miniprep Kit, Sigma)
- Nanodrop 8000 (Thermo Scientific)

Protocol:

Precipitation of DNA was conducted in order to concentrate and desalt gDNA extracted from cells cultured in 96-well plates. 1/10 volume of sodium acetate (pH 5.2) was added to the lysates followed by 2-2.5 volumes of cold 100% ethanol. Samples were incubated at -80°C for 20-45 min and then transferred into DNA Miniprep columns pretreated with column preparation solution. The columns were centrifuged at 12,000xg for 1 min,washed with 500µl washing buffer and centrifuged at 12,000xg for 1 min. A second wash with 500µl washing buffer was performed and samples centrifuged at

16,000xg for 3 mins. An additional centrifugation at 16,000xg for 1 min removed residual ethanol. The columns were placed in new collection tubes and 20µl of Milli-Q H_2O added to the filter of the columns and incubated at 60°C for 10 mins in a water bath. Samples were centrifuged at 12,000xg for 1 min and the the concentration and purity of the eluted DNA was measured using a Nanodrop 8000. DNA samples were frozen at -20°C for further experiments.

2.1.4.3.2 Determining the concentration and purity of gDNA

The concentration and purity of gDNA was determined using a Nanodrop 8000 Spectrophotometer (Thermo Scientific). Approximately 1.5µl of each sample was loaded onto the Nanodrop and the Nucleic Acids method on the ND-1000 V2.0.0 software and DNA option was used to analyse samples. The DNA concentration was recorded and the purity of the samples was assessed by the 260/280 and 260/230 ratio. Samples with a 260/280 ratio of around 1.8 and a 260/230 ratio of 2.0-2.2 were stored at -20°C freezer for further use.

2.1.4.4 RNA purification from Eukaryotic cells

Required materials:

- RNA purification kit RNeasy® Plus mini kit, Cat.No. 74134, (QIAGEN)
- B-mercaptoethanol (B-ME) (Sigma)
- 95-100% Ethanol (Sigma) for diluting the buffer RPE provided in the kit
- 70% Ethanol (Sigma)
- RNase- Free DNase Set, Cat no. 79254, (QIAGEN)
- Nanodrop 8000 (Thermo Scientific)

Protocol:

RNA isolation was performed using an RNeasy® Plus mini kit following the manufacturer's protocol. Briefly, approximately $1x10^6$ cella/ml were lysed in 350µl RLT

plus buffer. The homogenized lysate was transferred to a gDNA Eliminator spin column and centrifuged for 30 sec at 8000xg. 350µl of 70% Ethanol to the flow-through and the sample was directly transferred to the RNeasy spin column and centrifuged for 30 secs at 8000xg. The column was retained, 700µl of buffer RW1 was added and the column was centrifuged for 30 secs at 8000xg. The column was washed twice using 500µl of RPE buffer and centrifuged at 8000xg before eluting RNA in 50µl RNase-free water. Gel electrophoresis was performed to check the quality of the RNA and its concentration and purity was measured by using a Nanodrop 8000.

2.1.4.4.1 Determining the concentration and purity of RNA

The concentration and purity of RNA were measured using a Nanodrop 8000 Spectrophotometer (Thermo Scientific). Approximately 1.5µl of each sample was loaded onto the Nanodrop 8000 and the Nucleic Acids method on the ND-1000 V2.0.0 software and RNA option was used to analyse the samples. The RNA concentration was recorded and the purity and integrity of each sample was assessed. A 260/280 ratio of around 2.0 and a 260/230 ratio of 2.0-2.2 was required to lass the RNA as pure. RNA samples were then stored at -80°C freezer for further use.

2.1.4.5 cDNA synthesis

Required materials:

- Reverse Transcription kit, Cat no. 205310, 205311, 205313, (QIAGEN)
- mRNA of the genes of interest
- G-Storm GS4 thermal cycler (G-Storm)

Protocol:

Reverse transcription was performed using a QuantiTect® Reverse Transcription kit following the manufacturer's instructions. The solution was centrifuged briefly and the mastermix was prepared as following:

- xµl RNA or mRNA (1µg)
- 4µl (5x) TransAmp buffer
- 1µl Reverse transcriptase
- Up to 20µl DNase free-water

The mastermix was mixed gently by pipetting and a No Reverse Transcriptase (NRT) control was included in all experiments. Thermal-cycling conditions are shown in **table 2.8.**

Parameters	Temperature (°C)	Duration (min)
Primer annealing	25	10
Reverse transcription	42	15
Inactivation	85	5
Hold	10	-

Table 2.8: Thermal-cycling conditions used for reverse transcription reaction

2.1.4.6 Polymerase chain reaction (PCR)

Required materials:

- cDNA (0.05µg/µl)
- Oligonucleotides (10µM)
- MyTaq HS DNA polymerase kit (Bioline)
- G-Storm GS4 thermal cycler (G-Storm)

Protocol:

PCR of *TERT, ACYP2* and *TSPYL6* genes was performed on cDNA prepared from human iPSCs using MyTaq HS DNA polymerase. The master mix per 25µl reaction contained:

- 5µl 5x Bioline MyTaq buffer
- xµl forward primer, as given in individual experiment (500nM)
- xµl reverse primer, as given in individual experiment (500nM)
- xµl cDNA template, to a final concentration given in individual experiment details
- Bioline MyTaq DNA polymerase
- xµl dH_2O to a final volume of 25µl

Thermal-cycling conditions are shown in **table 2.9.** Following PCR, gel electrophoresis was performed.

	Temperature (°C)	**Duration (Sec)**	
Initial denaturation	95	60	
Denaturation	95	15	x40 cycles
Annealing	*x	60	
Elongation	72	90	
End cycle	-	-	
Final elongation	72	*300	
Hold	10	-	

Table 2.9: Thermal-cycling conditions for MyTaq PCR reactions. *x represents variable annealing temperature. 59°C used for *TERT*, and 57°C *ACYP2* and *TSPYL6* respectively. The final elongation of *TERT* is 90 secs.

2.1.4.7 Gel Electrophoresis

2.1.4.7.1 Agarose gel

Required materials:

- Agarose, Electrophoresis Grade (Melford Laboratories)
- 50x TAE (484g Tris in 1 litre of dH2O, 114.2ml acetic acid, 37.2g EDTA, pH to 7.6 with concentrated HCl, and made up to 2 litres with dH2O)
- GelRed™ Nucleic Acid Gel Stain, 10,000X in Water (Biotium)
- 5x DNA loading buffer blue (Bioline)
- HyperLadder™ I (Bioline)
- Syngene's GeneGenius Bio Imaging System

Protocol:

A 1.5% Agarose gel was prepared by dissolving 1.5g of Agarose in 100ml of 1x TAE running buffer (50x TAE diluted in dH_2O) and heating in the microwave. 10µl of GelRed™ was added to the dissolved mixture. The solution was allowed to cool before pouring into a gel electrophoresis tray and adding a comb. Each PCR sample was mixed with 5x DNA loading buffer blue. Samples and an appropriate DNA Ladder were run at 120 V and 400 ampere until sufficient separation had been achieved. The products were visualized using Syngene's GeneGenius Bio Imaging System and Syngene's GeneSnap software.

2.1.4.7.2 Polyacrylamide gel

Required materials:

- 10% Polyacrylamide gel (Novex TBE Gel)
- Running Buffer 5X (Novex TEB)
- XCell II Surelock Mini-cell vertical electrophoresis tank (Novex, Invitrogen)
- 5X Hi-Density TBE Buffer (Novex, Invitrogen)
- GelRed™ Nucleic Acid Gel Stain, 10,000X in Water (Biotium)

- HyperLadder™ I (Bioline)
- Syngene's GeneGenius Bio Imaging System

Protocol:

PCR samples were mixed with 5X Hi-Density TBE Buffer and loaded onto a 10% Polyacrylamide gel (Novex TBE Gel) with 10 wells in an XCell II Surelock Mini-cell vertical electrophoresis tank with TBE running buffer. An appropriate DNA Ladder was run alongside the PCR samples. A voltage of 200 V was applied and the running duration was determined according to segregation of the sample and ladder. The gel stained was stained with GelRed™ for 30 mins and then washed with dH_2O for 2 min. Samples were visualized using Syngene's GeneGenius Bio Imaging System and Syngene's GeneSnap software.

2.1.4.8 PCR Purification

Required materials:

- MinElute® PCR purification Kit (QIAGEN)
- 96-100% Ethanol (Sigma)
- 3M Sodium acetate (Sigma)
- Nanodrop 8000 (Thermo Scientific)

Protocol:

Purification of PCR reactions for Sanger sequencing was performed using QIAGEN MinElute® PCR purification Kit, as described in the manufacturer's instructions. Briefly, one volume of PCR reaction was mixed with five volumes of buffer PB. At this stage the colour of mixture should be yellow if the sample is at optimal pH. If the colour turns to the purple or orange 10µl of 3M sodium acetate should be added to adjust pH. The sample was then passed through a MinElute column to bind the PCR fragments in the membrane by centrifugation at 17,000xg for 1 min. The column was washed once and the PCR

product eluted by 10µl dH_2O. The concentration and purity of the purified PCR product was measured using a Nanodrop 8000 (as in **section 2.1.4.3.2**). The purified PCR fragments were stored at -20°C until sequencing.

2.1.4.9 Real-time PCR using SensiMix SYBR Green Mix

Required materials:

- 2x Sensi Mix SYBR, No-ROX (lot no. SMT-N313112) (Bioline)
- cDNA (0.05µg/µl)
- Oligonucleotides (10µM)
- Rotor-Gene Q instrument (QIAGEN)
- Rotor-Gene strip tubes and caps (QIAGEN)

Protocol:

RT-PCR was performed to measure the expression of genes using Bioline 2x Sensi Mix SYBR according to the manufacturer's protocol. In short, the specimens ran in duplicate or triplicate.

- 12.5µl Sensi Mix SYBR
- 0.75µl forward primer (300nM)
- 0.75µl reverse primer (300nM)
- xµl Template cDNA (125ng)
- xµl dH_2O to a final volume 25µl

A no template control (NTC) was included for every sample. A mastermix of the above mix was prepared and then aliquoted into strip tubes before addition of appropriate template/dH2O to tubes. The thermal-cycling conditions differed between experiments and the general conditions are summarized in **table 2.10.** The data obtained from each run was analysed by using Rotor-Gene Q Series 2.0.2 software.

Ribosomal phosphoprotein PO (36B4) or proteasome subunit beta 4 (PMSB4) housekeeping controls were used. The thermal cycling conditions of these genes are described in the **table 2.11**.

	Temperature (°C)	Duration (sec)	
Initial denaturation	95	600	
Denaturation	95	15	x30-40 cycles
Annealing	*x	-	
Extension	72	15	
Hold	Ambient	-	

Table 2.10: RT-PCR thermal cycling conditions. *x represents annealing temperature. 59°C for 15secs for *TERT* and 57°C for 15 secs for both genes *ACYP2* and *TSPYL6*.

A

	Temperature (°C)	Duration (sec)	
Initial denaturation	95	600	
Denaturation	95	10	X30 cycles
Annealing with extension	58	60	
Hold	Ambient	-	

B

	Temperature (°C)	Duration (sec)	
Initial denaturation	95	600	
Denaturation	95	15	X40 cycles
Annealing	57	15	
Extension	72	15	
Hold	Ambient	-	

Table 2.11: RT-PCR thermal-cycling conditions for housekeeping controls. Conditions for **(A)** 36B4 and **(B)** PSMB4.

2.1.4.10 Gradient PCR

Required materials:

- gDNA
- Oligonucleotides
- Appropriate DNA polymerase kit
- G-storm GS4 thermal cycler (G-Storm)

Protocol:

A gradient PCR was conducted to determine the optimal annealing temperature for the PCR primers to be used in the Surveyor assay for the *ACYP2*, *TSPYL6* and *TERT* loci. gDNA was used as a template and the reactions were performed in 0.2mL PCR tubes. The reaction mix contained:

- xµL appropriate reaction buffer
- 1µL Forward primer (500nM)
- 1µL Reverse primer (500nM)
- xµL appropriate DNA polymerase
- 100ng DNA
- xµl dH2O up to 25µl

The PCR tubes were vortexed and the samples underwent a 35-cycle PCR reaction on a G-Storm GS4 thermal-cycler. The sequence of the primers was supplied by Horizon Discovery Ltd. The cycling parameters are shown in **table 2.12**. The optimal annealing temperature of each primer pair was used for the Surveyor assay. The duration of the denaturation steps was determined by that recommended for the DNA polymerase according to manufacturer's instructions. NTC was also included as a control.

steps	Temperature (°C)	Duration	
Initial Denaturation	95	2 min	
Denaturation	95	15 sec	x35 cycles
Annealing	63.1	15 sec	
	63.3		
	63.6		
	64.1		
	65.1		
	66.0		
	66.9		
	67.9		
	68.9		
	69.5		
	69.9		
	70		
Elongation	70	1.30 min	
Final elongation	72	5 min	
Hold	4		

Table 2.12: PCR parameters for gradient PCR reactions

2.1.4.11 Preparation of gRNA and Cas9 plasmids

2.1.4.11.1 *E coli* competent cell transformation

Required Materials:

- *Escherichia coli* component cells DH5α (Bioline)
- S.O.C medium, Cat.No.15544-034 (Invitrogen)
- LB agar (Sigma-Aldrich)
- LB broth ((Sigma-Aldrich)

- Appropriate antibiotic, Ampicillin or streptomycin (Sigma)
- 90mm petri-dish (Sarstedt)
- Water bath
- Shaking incubator
- Incubator

Protocol:

Competent cells were thawed on ice and 1-50ng of plasmid DNA added to 100µl of cells and incubated on ice for 30 min. The mixture was heat-chocked at 42°C for 30 seconds and then immediately incubated on ice for 2 mins. 200µl of S.O.C medium was added to the mixture and then incubated at 37°C for 1h in a shaking incubator at 200 rpm. The cells were then plated onto LB agar plate (1.5% agar, 2.5 LB in 90mm plate) containing appropriate antibiotic for selection and inverted plates were incubated overnight at 37°C.

2.1.4.11.2 Growing bacterial colonies

A single colony was picked and transplanted into 10ml of 2.5% LB medium containing appropriate antibiotic and incubated at 37°C overnight in a shaking incubator at 200 rpm.

2.1.4.11.3 Extraction of plasmid DNA

Required materials:

- Plasmid purification maxi kit (QIAGEN)
- 70% Ethanol (Sigma)
- Nanodrop 8000 (Thermo Scientific)

Protocol:

Plasmid DNA was isolated by using QIAGEN plasmid maxi kit according to the manufacturer's protocol with some alteration in conditions. Briefly, after overnight growth in 2.5% LB medium (appropriate volume to the copy-plasmid), the bacteria were harvested at 3000xg for 30 min at room temperature. The cells were resuspended with buffer P1 and lysed with buffer P2 for 5 min. Prechilled buffer P3 was added to neutralize the cells for 20 min. The lysed cells were precipitated at 4255xg for 30 mins at 4°C and the supernatant re-centrifuged under the same conditions before being added to a QIAGEN column. DNA was eluted by addition of 15ml buffer QF and then precipitated with 0.7 volumes of isopropanol at room-temperature then centrifuged at 4255 x g for 30 mins at 4°C. After decanting the supernatant, the DNA pellet was washed with 5ml room-temperature 70% Ethanol and centrifuged at 4255 x g for 10 min at 4°C. The DNA pellet was then air-dried for 5 min and dissolved in 200µl of dH_2O. The plasmid concentration and purity was then determined using a nanodrop (as described in **section 2.1.4.3.2**) and stored at -20°C.

2.1.4.12 Surveyor assay

Required materials:

- Transgenomic Surveryor® mutation detection kit
- PCR product of the knock out and wild type iPSCs
- G_storm GS4 thermal cycler (G-storm)

Protocol:

Screening of targeted iPSCs was performed using a Surveryor assay. First, the gDNA at the taget regions was by PCR using iPSC DNA collected at the first passage following CRISPR/Cas9 transfection. Platinum® Taq DNA, Phusion™ HS DNA and MyTaq HS DNA polymerase (**section 2.1.4.10**) were used for screening experiments for *ACYPT2, TSPYL6* and *TERT* genes respectively (**table 2.13**).

	DNA polymerase		
Continents	**Platinum® Taq DNA polymerase**	**Phusion™ HS DNA polymerase**	**MyTaq HS DNA polymerase**
Nuclease-free water	up to 50µl reaction	up to 50µl reaction	up to 50µl reaction
Polymerase buffer	10x Platinum buffer	5x Phusion® HF buffer	5x MyTaq buffer
Deoxynucleotide triphosphates (dNTPs) (10mM)	1µl	1µl	_
$MgCl_2$ (2mM)	2µl	_	_
Forward primer (10µM)	2.5µl	2.5µl	2.5µl
Reverse primer (10µM)	2.5µl	2.5µl	2.5µl
DNA polymerase	0.2µl	0.5µl	1µl
Template (100-200ng)	xµl	xµl	xµl

Table 2.13: PCR reaction mix for each DNA polymerase.

Thermal-cycling conditions are shown in **table 2.14**. The duration of the denaturation steps varied depending on the DNA polymerase.

Steps	Temperature (°C)	Duration (sec)	
Initial denaturation	98	30	
Denaturation	98	10	35 cycles
Annealing	*x	60	
Extension	72	90	
Final extension	72	300	
Hold	4	-	

Table 2.14: PCR parameters for gDNA amplification. *x represents annealing temperature. 65°C for *TERT* and 65.1°C for both genes *ACYP2* and *TSPYL6*.

5µl of each PCR reaction was separated by gel electrophoresis, as described in **section 2.1.4.7.1**, to confirm the PCR reaction was successful and to estimate the amount of product.

The second step of the Surveryor assay involves hybridization of equal amount of wild-type and mutant DNA to form a heteroduplex. The final volume of each sample should be at least 10µl with a concentration of 200ng DNA. Thermal-cycling conditions are shown in **table 2.15**.

In the final step a transgenomic Surveryor® mutation detection kit was used to identify detect DNA mismatch. The heteroduplex DNA samples were treated with mismatch-specific DNA nuclease as shown in a **table 2.16**.

Temperature (°C)	Time (min)
95	10
85	1
75	1
65	1
55	1
45	1
35	1
25	1
10	Hold

Table 2.15: The parameters of thermocycler to form heteroduplex DNA.

	DNA polymerase		
Reaction mix	Platinum® Taq DNA polymerase	Phusion™ HS DNA polymerase	MyTaq HS DNA polymerase
0.15M MgCl2 solution (1:10)	xµl	–	–
Surveyor enhancer S	1µl	1µl	1µl
Surveryor nuclease S	1µl	2µl	2µl

Table 2.16: Reaction-mix for the Surveyor assay

The samples are mixed gently by vortexing and incubated at 42°C for 60 min. After incubation, a 1:10 volume of Stop solution was added and mixed by pipetting. Agarose or Polyacrylamide gel electrophoresis was carried out as described in **section 2.1.4.7** to identify cleaved products.

2.1.4.13 Sequencing

Sequencing of iPSCs after CRISPR mediated-genome editing was conducted using Sanger sequencing. After purifying PCR products (**section 2.1.4.8**), samples were sent to source bioscience http://www.lifesciences.sourcebioscience.com at a concentration of 1ng/µl per 100bp for sequencing. Sequencing was performed in one or two directions utilizing the forward and reverse primerd at a concentration of 3.2µm/µl. Sequence traces were analysed to identify mutated clones.

2.1.4.14 T-easy cloning

To confirm the genotype of suspected homozygous knockout clones, PCR-amplification of gDNA was performed and cloned into pGEM®-T Easy.

2.1.4.14.1 Preparation insertions

Required materials:

- PCR products of the target sites of each homozygous knockout clone
- Plasmid purification maxi kit (QIAGEN)
- 70% Ethanol (Sigma)
- Nanodrop 8000 (Thermo Scientific)

Protocol:

gDNA at the target site of suspected homozygous clones was amplified by PCR using the Surveyor primers described in **section 2.1.4.12** followed by PCR purification as described in **section 2.1.4.8**. The concentration of purified products was determined using a Nanodrop 8000 as described in **section 2.1.4.3.2**.

2.1.4.14.2 Ligation reaction

Required materials:

- pGEM®-T Easy vector system I
- Taq polymerase kit that produces A-overhang if qPCR products prepared with high fidelity polymerase or any polymerase which does not add adenine overhang (A-overhang).
- G_Storm GS4 thermal cycler (G-Storm)

Protocol:

T-easy vector system contains a 3015bp linearized plasmid with a single 3-terminal thymidine overhang at each end. The PCR product for insertion should have A-overhangs. MyTaq and GoTaq polymerases incorporate A-overhangs, however, blunt PCR products required A-overhang addition using the following conditions:

- 10µl (5x) buffer
- 5µl (1mM) dATP
- 0.5µl GoTaq polymerase
- 4.5µl dH_2O

The mixture was then incubated 95°C at 2min in the G-Storm GS4 thermal-cycler.

The ligation reaction was prepared as follows:

- 5µl (2x) rapid ligation buffer, T4 DNA ligase
- 1µl pGEM®-T Easy vector
- xµl PCR product (to an approximate 1:1 bp/ng ratio with pGEM®-T Easy vector)
- 1µl T4 DNA ligase (50ng)

- dH_2O up to final volume

The reaction was then mixed by pipetting followed by incubation for 1hr and transformation into competent cells as described in **section 2.1.4.11.1**.

2.1.4.14.3 Confirmation of cloning

Required materials:

- M13 Forward and Reverse Oligonucleotides
- MyTaq HS DNA polymerase kit (Bioline)
- G_Storm GS4 thermal cycler (G-Storm)

Protocol:

To ensure that cloning of the expected PCR product had been successful, up to 10 clones were picked from each transformation using a 10µl pipette tip and spread on to a fresh LB/Agar plate. A PCR reactions was prepared using the picked colony as template. The reaction mix is below:

- 4µl (5x) Bioline MyTaq buffer
- 1µl M13 forward primer (10µM)
- 1µl M13 reverse primer (10µM)
- 0.1µl Bioline MyTaq DNA polymerase
- dH_2O up to 20µl

The thermal-cycling conditions are in **table 2.17**

Steps	Temperature (˚C)	Duration (sec)	
Initial denaturation	95	120	
Denaturation	95	15	35 cycles
Annealing	56	30	
Extension	72	60	
Final extension	72	420	
Hold	4	-	

Table 2.17: PCR parameters for screening pGEM T-easy clones.

Agarose gel electrophoresis was conducted as described in **section 2.1.4.7.1** to identify positive clones.

Positive colonies were picked and grown in 5ml of 2.5% LB medium containing appropriate antibiotic as described in **section 2.1.4.11.2**. Plasmid DNA was then extracted as described in **section 2.1.4.11.3** was sequenced by Source Bioscience using the M13 primer. The sequence was compared with reference to determine the effects of CRISPR/Cas9 targeting.

2.1.4.15 Telomere length measurement

2.1.4.15.1 Quantitative polymerase chain reaction (qPCR)

Required materials:

- 2x Sensi Mix SYBR, No-ROX (lot no. SMT-N313112) (Bioline)
- Modified iPSC DNA
- Oligo forward and reverse primer of telomere length
- Oligo forward and reverse primer of single copy gene

- K562 DNA
- Qiagility machine
- Rotorgene Q (QIAGEN)

Protocol:

Telomere length was measured by qPCR using the protocol previously described by Richard Cawthon (Cawthon, 2002). The sequence of the primers used in this experiment are shown in **table 2.18**. Telomere length was determined by dividing the quantity of telomeric repeats (T) to the quantity of a single copy gene (S), to give the T/S ratio. The single copy gene utilised to normalize for cell count was 36B4.

Name of primers	Primer sequences (5'- 3')
Tel1b	CGGTTTGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTT
Tel2b	GGCTTGCCTTACCCTTACCCTTACCCTTACCCTTACCCT
36B4F	CAGCAAGTGGGAAGGTGTAATCC
36B4R	CCCATTCTATCATCAACGGGTACAA

Table 2.18: Primer sequences for telomere length measurement. Tel1b and Tel2b indicate the forward and reverse primers for amplifying telomeric DNA. 36B4F and 36B4R were used to amplify 36B4. Primers are 5'-3'.

T and S runs were conducted separately. To ensure the same enzyme mix was used for T and S, Master mixes for both were made up at the same time. In addition, to keep conditions uniform the T run was performed immediately before the S run. The master mix is shown in **table 2.19**. Reactions were performed using a Qiagility machine and 22µl of master mix and 3µl of gDNA (10ng/µl) in a Rotordisc containing 100 wells. 48 samples, NTC and a calibrator sample (K562 gDNA) were plated in duplicate. The rotordisc was then sealed and transferred to a Rotorgene Q (QIAGEN) to run. The thermal-cycling profile on the Rotorgene Q was for T 95˚C for 10 min followed by 20 cycles of 95˚C for 15 sec and 58˚C for 1min and for S as following, 95˚C for 10 min followed by 30 cycles of 95˚C for 15 sec and 58˚C for 1min. to achieve better runs, the Qiagility was warmed for 30 min before the first run of the day.

Volumes / Materials	Telomere (T)	Single copy gene (S)
2x SensiMix	12.5µl	12.5µl
Telb and Tel2b	300nM of each	-
36B4F	-	300nM
36B4R	-	500nM
gDNA template	30ng	30ng
dH_2O	Up to 25µl	Up to 25µl

Table 2.19: Master mixes of telomere and single copy gene.

2.1.4.15.2 Data analysis

Analysis was conducted utilizing the Rotorgene Comparative quantitation software. The mean amplification of all samples in the run was calculated (MAE) and the concentration of each sample relative to the calibrator (reference) sample was calculated using the equation:

$$\text{Sample concentration} = \text{MAE}^{(\text{Calibrator take-off} - \text{sample take-off})}$$

This was determined for both the T run and S run for each sample duplicate. The relative concentrations of all samples except NTC were then exported to Microsoft Excel to calculate average telomere length. Telomere length was expressed as T relative concentration / S relative concentration (T/S). All data were relative to the K562 calibrator sample. The duplicate values that were more than 0.2 of a cycle different for the take-off value were excluded in this assay.

2.1.5 Statistical Analysis

Data from all experiments was analysed using GraphPad Prism 6.00 (GraphPad software Inc.) except for the analysis of telomere length using a mixed model in **sections 4.2.3.4** and **5.2.2.5** which was performed by Dr. Chris Nelson, Department of Cardiovascular Sciences/University of Leicester. Results are shown as mean ± Standard Deviation (SD) unless otherwise stated. Statistical significance of the data was assessed by a Student's t-test. A one-way Analysis of Variance (ANOVA) was used to compare more than two groups.

Chapter 3

Optimisation of human induced Pluripotent Stem Cells for Genome Editing

3.1 Introduction

In 2012 the Nobel Prize in medicine or physiological was awarded to Shinya Yamanaka and Sir John Gurdon for the discovery that mammalian somatic cells can be reprogrammed into stem cells. Since their discovery induced pluripotent stem cells (iPSCs) have become the experimental model of choice for regenerative medicine, disease modelling of both monogenic and polygenic disorders and the emergence of genome editing technologies has increased their utility, especially for functional genomic studies. For this study, the reactivation of telomerase during reprogramming and resulting maintenance of telomere length by iPSCs means they provide a system for the investigation of both known and novel genes involved in the regulation of telomere length.

3.1.1.2 Culture of iPSCs

Pluripotent stem cells (PSCs) require careful culture so that they maintain their pluripotency. Originally, hiPSCs were cultured in similar conditions to those used for ESC culture with inactivated mouse embryonic fibroblasts (MEFs) as a supporting or "feeder" layer and Dulbecco's Modified Engle's Medium (DMEM) containing 10% fetal bovine serum (FBS) (Takahashi and Yamanaka 2006, Takahashi et al. 2007). Such conditions prevent spontaneous differentiation and support PSC expansion (Thomson et al. 1998). Alternative human derived feeder cells including fetal fibroblasts, foreskin fibroblasts, human placenta fibroblasts, and adult marrow cells have also been used to support PSC culture (Richards et al. 2002, Amit et al. 2003, Hovatta et al. 2003, Simón et al. 2005, Genbacev et al. 2005, Tecirlioglu et al. 2010, Cheng et al. 2003). These cell lines all support the maintenance of human pluripotent stem cells and allow prolonged culture in an undifferentiated state. However, these culture conditions are problematic for certain applications because of secretion of undefined factors into the media by MEFs and potential batch effects of FBS. The presence of a feeder layer also raises additional challenges for genome editing due to the long-term culture and selection involved in the technique. Because of these reasons feeder independent culture techniques that employ solubilized extracellular matrix (ECM) components in combination with a defined culture media have been developed. The most commonly used matrix is Matrigel™ an ECM

mixture isolated from Englebreth-Holm-Swarm mouse tumor cells consisting primarily of collagen IV, laminin and entactin, and considered to be a reconstituted basement membrane preparation (Kleinman and Martin 2005, Kleinman et al. 1982, Orkin et al. 1977). A number of defined serum-free culture media have been developed, such as mTesR®1 (STEMCELL™ technologies) and Essential 8™ (Gibco), to overcome the potential batch effects and contamination issues of FBS and ensure more consistent and reproducible stem cell culture. PSCs are passaged at a ratio between 1:3 and 1:9 depending on the growth rate of the cells by specific enzymes such as collagenase, Dispase and Accutase or mechanical approaches such as cell scraper or other passaging tools. When colonies become too large and dense, they are prone to differentiation.

3.1.2 Genome editing in iPSCs

The combined potential of iPSCs and genome editing has been apparent for many years with some of the earliest applications of ZFNs and TALENs (Hockemeyer et al 2009, Zou et al 2009, Hockemeyer et al. 2011) and the first use of the CRISPR/Cas9 (Mali et al 2013, Cong et al 2013) being in hiPSCs. Indeed, since the first reported editing of hiPSCs using the CRISPR/Cas9 system there have been editing of patient-derived iPSCs for a number of diseases, loss-of-function studies and double and triple gene knockouts (Xie et al. 2014, González et al. 2014, Wang et al. 2015) In addition, there are now several cases where edited hiPSCs are moving towards clinical applications. For example, Huang *et al*. (2015) successfully corrected a mutation in *HBB* gene in hiPSCs from Sickle cell anaemia patients, which they then differentiated in to erythrocytes (Huang et al. 2015), while Li and colleagues corrected mutations in hiPSCs from Duchenne muscular dystrophy patients and differentiated these in to skeletal muscle cells (Li et al. 2015).

3.1.2.1 Technical considerations for genome editing iPSCs

There are a number of technical aspects that need to be considered for genome editing experiments from initial design through successful targeting, which are shared regardless of the target cell line or organism. However, two specific procedures that are potentially

problematic for hiPSCs and require careful optimisation are the delivery of the genome editing machinery and the clonal isolation and expansion of the targeted cells. Successful targeting requires high nuclear expression of the genome editing machinery, however, hiPSCs can be less amenable and more sensitive to standard transfection and electroporation protocols than other cell lines and many protocols therefore suggest enriching for more highly expressing cells antibiotic selection or fluorescence-activated cell sorting (Chatterjee et al. 2011, Peters et al. 2013). The clonal isolation of hiPSCs is also less than straightforward, as successful culture requires cells to be maintained between 20-80% confluence with increased spontaneous differentiation and cell death if these limits are exceeded. To overcome this the most frequent approach is to optimise the seeding and culture of clonal hiPSC colonies followed by colony picking.

3.1.3 Aims and objectives

This study involves the knockout of known and novel genes involved in the maintenance of telomere length in hiPSCs. This is the first time hiPSCs have been cultured by our group and the first use of CRISPR/Cas9 at the University of Leicester. The objective of this part of the study was therefore to establish long-term hiPSC culture and optimise CRISPR/Cas9 mediated gene knock-out.

Aims;

- To establish culture of hiPSCs and quality-control procedures.
- Optimization of hiPSCs for CRISPR/Cas9 genome editing.

3.2 Results

3.2.1 Human iPSC culture

3.2.1.1 iPSC cell line

Wild-type hiPSCs (GM23720*B, NIGMS collection) were purchased from the Coriell institute for Medical Research, USA. These cells are episomally reprogrammed B-lymphocytes from a healthy 22 year-old woman and have been evaluated for cell surface antigen expression and pluripotency. The culture conditions provided by the Coriell institute were followed with further optimisation and quality control procedures detailed below.

3.2.1.2 Optimisation of iPSC culture

iPSCs were cultured in feeder-free conditions as described in **sections 2.1.1.2.1** and **2.1.1.2.2** using Matrigel™ (BD Biosciences)-coated plates and mouse embryonic fibroblast serum-conditioned media (MEF-CM) mTesR®1 (STEMCELL™ technologies). Visual monitoring was performed daily in order to check confluence and cell morphology. Healthy, undifferentiated iPSCs grow in tightly associated compact colonies and individual cells have a high nuclear to cytoplasmic ratio with prominent nucleoli as seen under an inverted microscope **(Figure 3.1)**. Differentiated cells with the appearance of neuronal or fibroblast cells tended to be present as separate colonies or at the edges of iPSC colonies. Suspected embryoid bodies were also noted and there were occasional clusters of dead cells at the edge and centre of undifferentiated and differentiated colonies **(Figure 3.2).** In general, the percentage of the undifferentiated iPSCs was estimated to be approximately 80%.

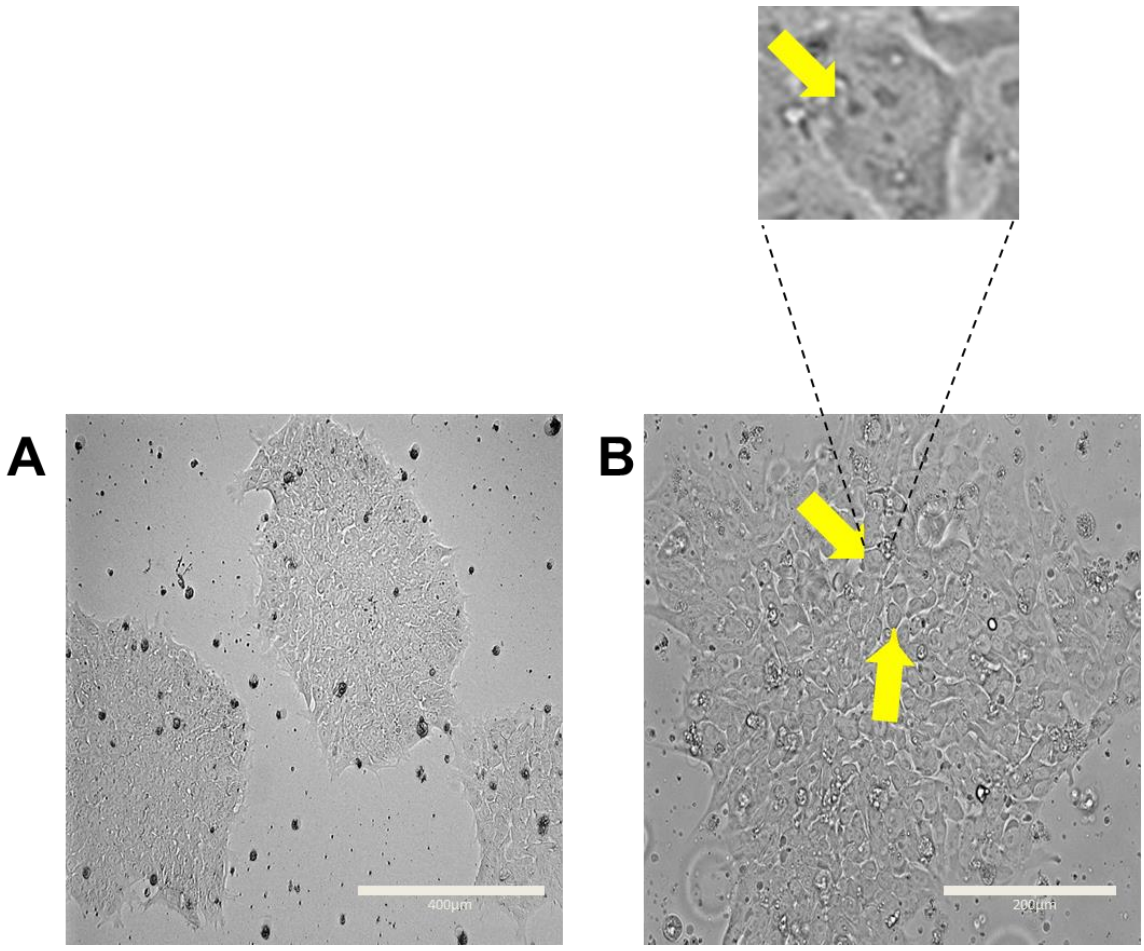


Figure 3.1: Morphology of healthy iPS colonies in feeder-free culture. (**A**) Three undifferentiated compact iPSC colonies **(the scale bar indicates 400µm)**. (**B**) A single iPSCs colony showing tightly associated cells with high nuclear to cytoplasmic ratio and prominent nucleoli (yellow arrows), **(the scale bar indicates 200µm).**

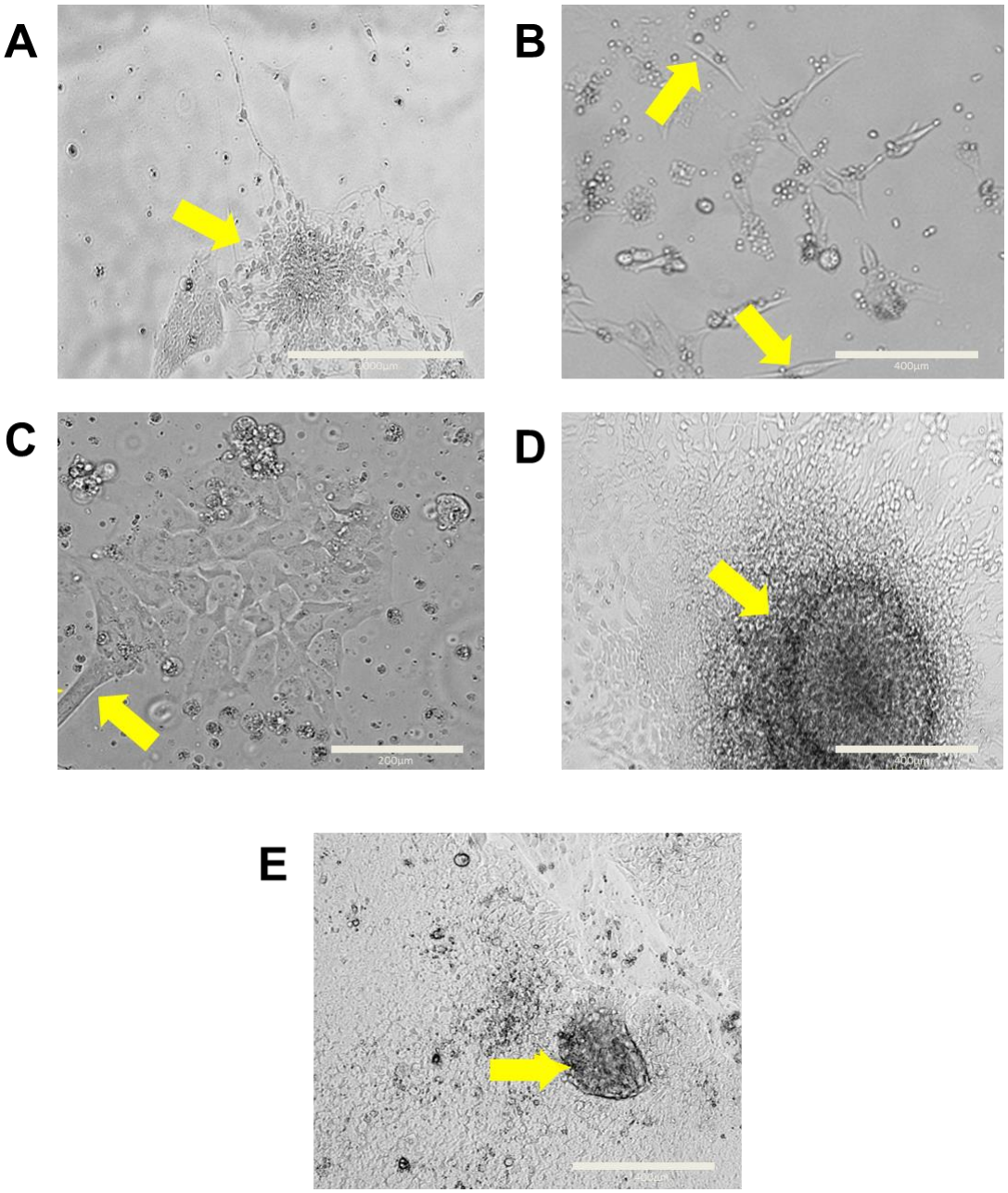


Figure 3.2: Examples of iPSC differentiation and death. (**A**) A differentiated colony containing a neuronal like cell (yellow arrow) **(the scale bar indicates 1000µm)**. (**B**) Fibroblast like cells **(the scale bar indicates 400µm).** (**C**) Early differentiation around the edge of an iPSC colony **(the scale bar indicates 200µm)**. (**D**) Embryonic body **(the scale bar indicates 400µm)** (**E**) Dead cells within an overgrown colony (yellow arrow) **(the scale bar indicates 400µm)**.

3.2.1.3 Passaging techniques and enzymes

When cell confluence reached 70-80% the cells were passaged using the Dispase II enzyme (stem cells™) as described in the **section 2.1.1.2.3**. Briefly, the culture media was removed from the plate and the cells washed twice with Dulbecco's Modified Engle's Medium (DMEM) (Sigma) or Phosphate buffer saline (PBS) (DulbeccoA, Oxid). There was no noticeable difference between PBS or DMEM. Dispase II was added and the plates incubated at 37°C with frequent observation. When the edges of colonies appeared slightly folded back, but the colonies were still attached to the plate (approximately 7 mins) the Dispase II was aspirated. The cells were then washed twice with DMEM, mTesR™1 added, and iPSC colonies carefully detached using a cell scraper. A drawback of using this method is the retention of differentiated cells and the potential for a gradual increase in the proportion of differentiated cells with prolonged culture **(Figure 3.2)**. In an attempt to decrease the percentage of differentiated colonies we tested the non-enzymatic reagent ReLeSR™ (Stem Cell Technologies), which preferentially dissociates hiPSCs and does not require scraping **(Figure 3.3)**. As described in **section 2.1.1.2.4**, when hiPSCs reached 70-80% confluence the media was removed and the colonies washed twice with PBS or DMEM and 1 ml/well of ReLeSR™ added for 1 min, aspirated and the plates then incubated for 5 mins at 37°C. Following incubation, 1ml/well mTeSR™1 medium was added and the plate was tapped gently on the side of the plate for 30-60 sec. The cell suspension was gently mixed by pipetting up and down in a 2 ml serological pipette to produce cell aggregates of approximately 50-200 μm for plating. 5-6 ml mTeSR™1 medium was added to the cell suspension and 1ml/well was replated on 6-well Matrigel™-coated plates. An additional 1ml of mTeSR™1 was then added to each well to achieve a final volume of 2ml/well medium. ReLeSR™ resulted in a substantial decrease in the amount of differentiation and was therefore used for all general iPSC maintenance. Passaging was performed at a ratio of 1:5 to 1:6, which ensured appropriate spacing between colonies. Approximately $1x10^6$-$1x10^{12}$ cells/ml were regularly banked (as described in **section 2.1.1.4)** in liquid nitrogen. On revival, **(section 2.1.1.2.2)** approximately 70-85% cells were generally viable.

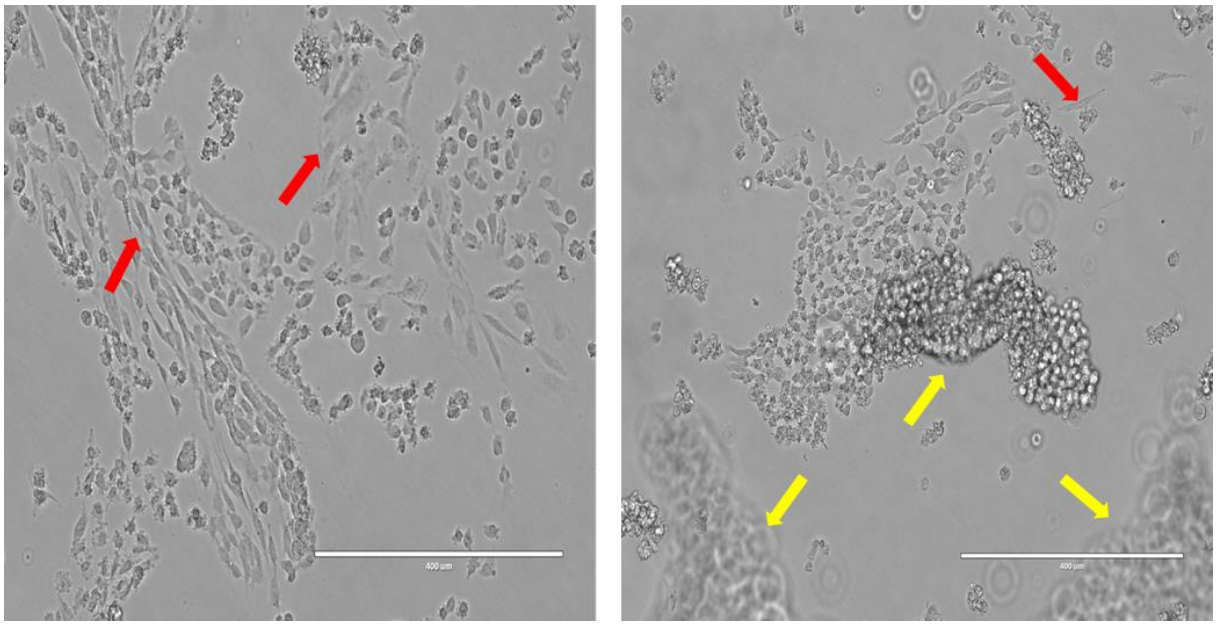


Figure 3.3: Undifferentiated and differentiated iPSCs after incubation with ReLeSR™ enzyme. The yellow arrows indicate detached iPS colonies (right figure) and the red arrows indicate the adherent differentiated iPSCs of different morphology (right and left figures). **(the scale bar indicates 400µm)**.

3.2.2 iPSC Quality Control

3.2.2.1 Flow cytometry analysis of iPSC surface markers

iPSCs can be characterised by measuring expression of specific cell-surface markers such as glycolipid antigens (SSEA3 and SSEA4), keratan sulfate antigens (TRA-1-60, TRA-1-81, GCTM2 and GCT343), and protein antigens (CD9, Thy1(CD90), tissue–nonspecific alkaline phosphatase and class 1 HLA) (Adewumi et al. 2007). iPSCs can also be characterised by expression of the stem cell self-renewal transcription factors NANOG and POU5F1 (OCT4) (Schwarz et al. 2014).

The expression of stem cell markers was measured every 10 passages to confirm that iPSCs were being maintained in an undifferentiated state. hiPSC colonies were harvested as a single cell suspension using 1x trypsin/EDTA, followed by fixation in 4% paraformaldehyde (PFA). The fixed cells were stained with isotype controls and antibodies against cell surface markers using the BD stemflow™ human pluripotent stem cells sorting and analysis kit. The positive and negative beads were also prepared to assist in optimizing application setup and calculating fluorescence compensation. The positive beads were coated with antibodies that bind to specific cell surface markers; SSEA-1, SSEA-3 and TRA-1-81. Negative beads had no binding capacity. The iPSC population was characterised by two parameters; forward scatter (FS), which is a measure of cell size and side scatter (SS), which is a measure of cell granularity. The entire population was analysed for surface marker expression (**Figure 3.4**), cells debris presented towards the origin of the graph. Finally, the surface markers were detected by flow cytometry with 3 lasers; blue (488), red (638) and violet (405) and there are filters to detect up to 10 emitted wavelengths from fluorescent dyes.

The AW08603 software (Callios) was used to analyse the data. S**ection 2.1.2.1** describes in these methods in more detail. Passage 10 iPSCs incubated with antibodies against SSEA-1, SSEA-3 and TRA-1-81 were 96.3% positive for SSEA-3 detected with PE fluorescent dye and 75.66% for TRA-1-81 detected with Alexa Fluor fluorescent dye (**Figure 3.5, A, B)**. Only 2.11% of cells stained for the negative surface marker SSEA-1 detected with FITC (**Figure 3.5, C**) These results demonstrated that at P10 the iPSCs remained pluripotent and were being cultured successfully.

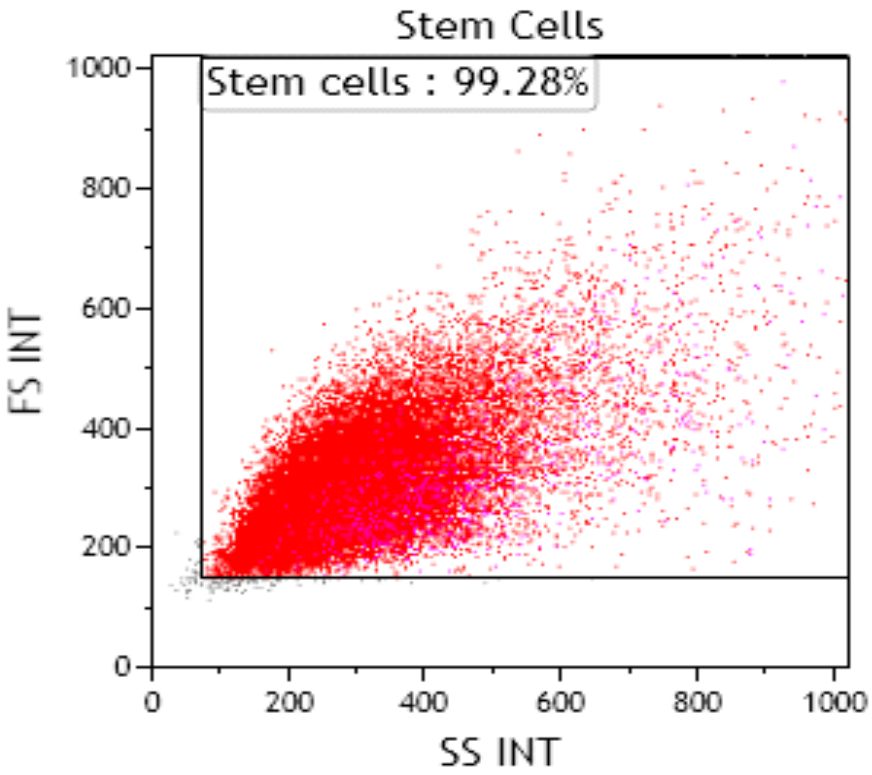


Figure 3.4: Contour plot of the hiPSC population. iPSCs were identified by their forward scatter (FS, Y axis) and side scatter (SS, X axis). This population was used to determine the cell surface marker expression.

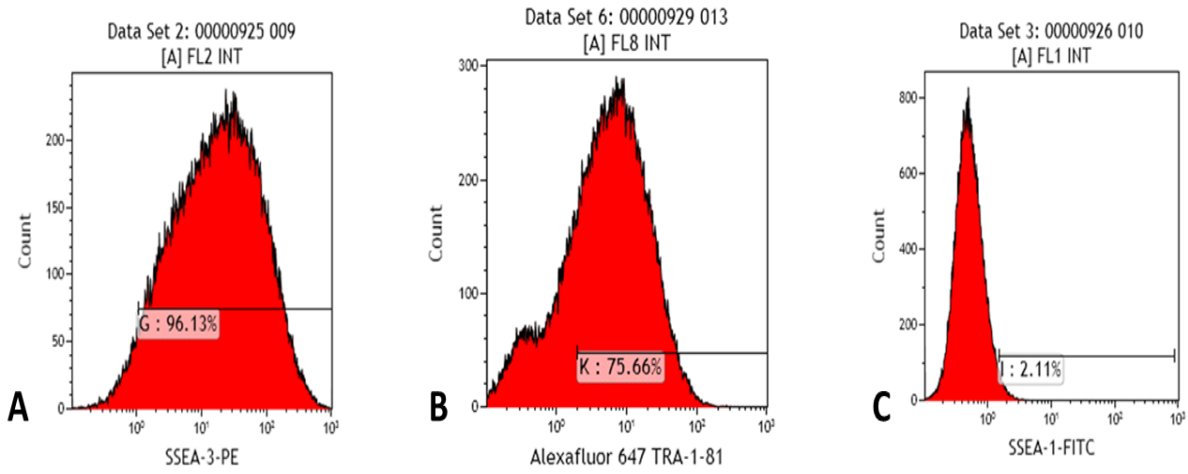


Figure 3.5: hiPSCs at passage 10 were analyses by flow cytometry to determine the surface marker expression. iPSCs were positive for **(A)** SSEA-3 antibody coated with PE fluorescent dye and **(B)** TRA-1-81 antibody coated with Alexa Fluor® fluorescent dye. The cells were negative for **(C)** SSEA-1 antibody coated with FITC fluorescent dye.

3.2.2.2 Karyotype analysis

The iPSCs used in this study had previously been characterized by the Coriell institute and reported to be genetically stable. Chromosome integrity is affected by long-term culture and culture conditions therefore karyotyping needs to be regularly performed (Draper et al. 2004, Caisander et al. 2006, Grandela, Wolvetang 2007).

Karyotyping was initially performed using Giemsa (result not shown) and DAPI staining, however, the resolution achieved with both reagents was too low to properly analyse the cells (**Figure 3.6, A**). Further karyotypic analysis was carried out using chromosome resolution additive (CRA), which increases chromosome length and provides greater separation (**Figure 3.6, B, C**). As described in the **section 2.1.2.2**. Briefly, mTesR™1 medium was replaced with fresh mTesR™1 containing 2µl of 0.1 mg/ml Nocadazol, in order to arrest the cells in prometaphase and incubated for 24 hours. The cells were then trypsinized and the pellet was resuspended in hypotonic solution (0.056M KCL in dH2O) and incubated at room temperature for 30 mins. This causes the cells to swell resulting in greater separation of the chromosomes. The pellet was washed three times with fresh Carnoy's fixation solution (75% Methanol and 25% Acetic acid) and then resuspended in Carnoy's solution. Two slides were prepared for each sample by warming to 37°C and the sample applied in a dropwise manner to slides held at a 45° angle. The slides were then left to air-dry overnight before staining with Giemsa (Sigma) and 4′, 6-diamidino-2-phenylindole (DAPI) (Invitrogen). Images were taken at 100x magnification and chromosomes counted. Chromosome counting revealed all cells analysed to have a normal karyotype.

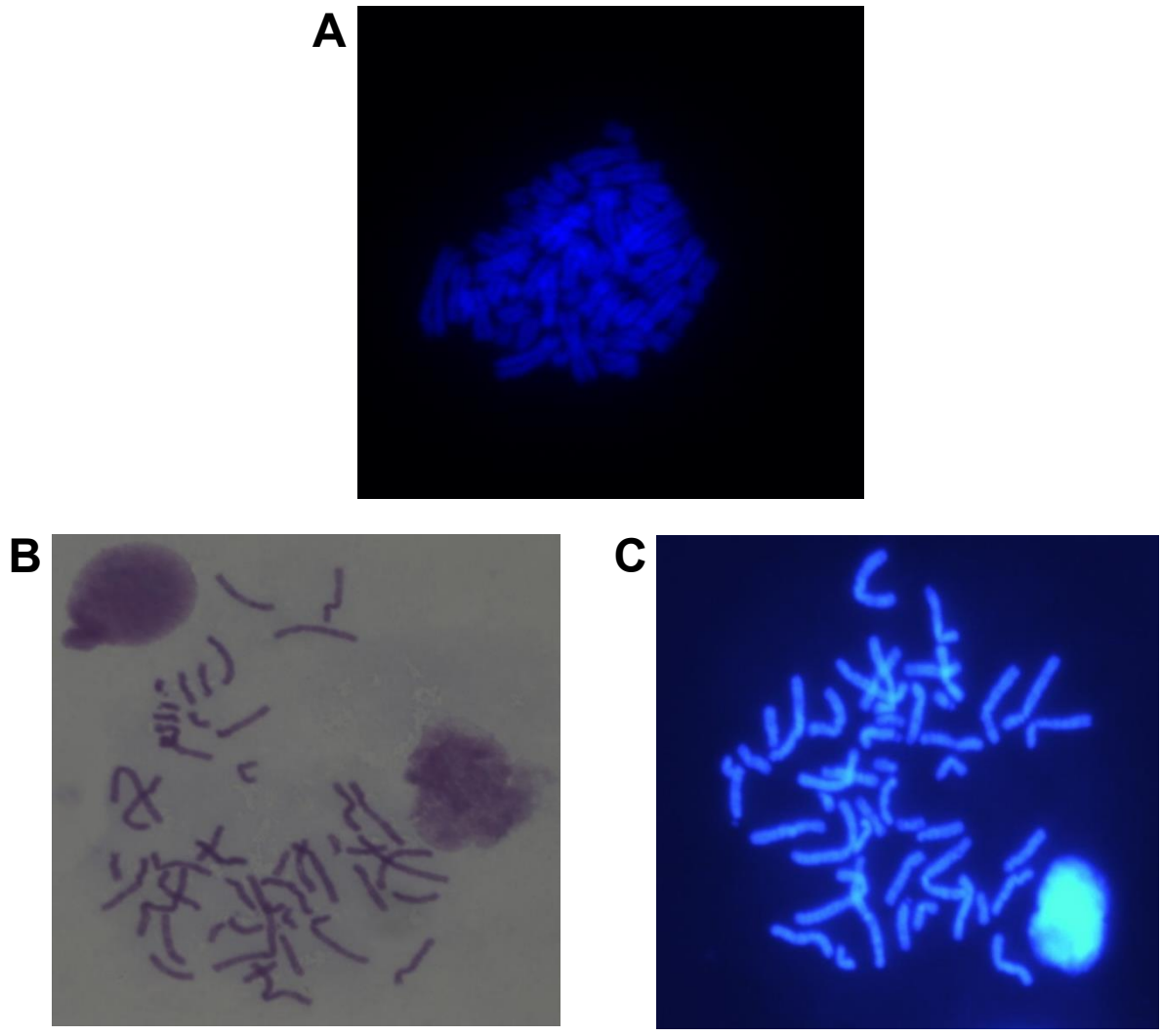


Figure 3.6: Karyotypic analysis of iPSCs. (A) iPSCs stained with DAPI, the resolution achieve was insufficient to accurately analyse the karyotype of the iPSCs before using CRA. **(B)** iPSCs stained with Giemsa. **(C)** iPSCs stained with DAPI.

3.2.3 iPSC transfection

3.2.3.1 Analysis of lipid-based iPSC transfection

CRISPR/Cas9 genome editing requires high expression of the genome editing machinery in the target cell nucleus. Therefore, several different transfection methods were tested in order to determine the optimal delivery method. Initially, forward and reverse lipid-based transfection using TransIT®2020 (Mirus) and Lipofectamine®3000 (Invitrogen) reagents was investigated. Forward transfection involves the addition of the lipid-plasmid complex onto the cells while for reverse transfection a cell suspension is added to the lipid-plasmid complex. Reverse transfection results in prolonged exposure of the cells with the transfection reagent compared to forward transfection (Villa-Diaz et al. 2010). A green fluorescent protein (GFP) plasmid pCMV-GFP (4.7kbp) was used to determine transfection efficiency. Cells were analysed visually using an Evos® fluorescent microscope 24h after transfection by counting the percentage of cells expressing GFP cells. Control plasmid free transfections were used for comparison **(Figure 3.8, B)**. Reverse transfection of iPSCs using TransIT®2020 produced the highest transfection efficiency, however, the proportion of transfected cells was still only 23.3% and forward transfection resulting in just 5.3% of cells expressing GFP **(Figure 3.7)**. The results with Lipofectamine®3000 were even less successful with the highest efficiency of reverse transfection using the recommended 3.7µl of Lipofectamine®3000 being only 3.7% and forward transfection only slightly higher at 5.2%. Increasing the amount of Lipofectamine®3000 to 7.5µl increased efficiency to 10.7% with reverse transfection and 15.8% with forward transfection **(Figure 3.7)**. It was concluded that the relatively low efficiencies achieved with lipid-based methods was too low for CRISPR/Cas9 genome editing. Cell viability wasn't noticeably affected with either method

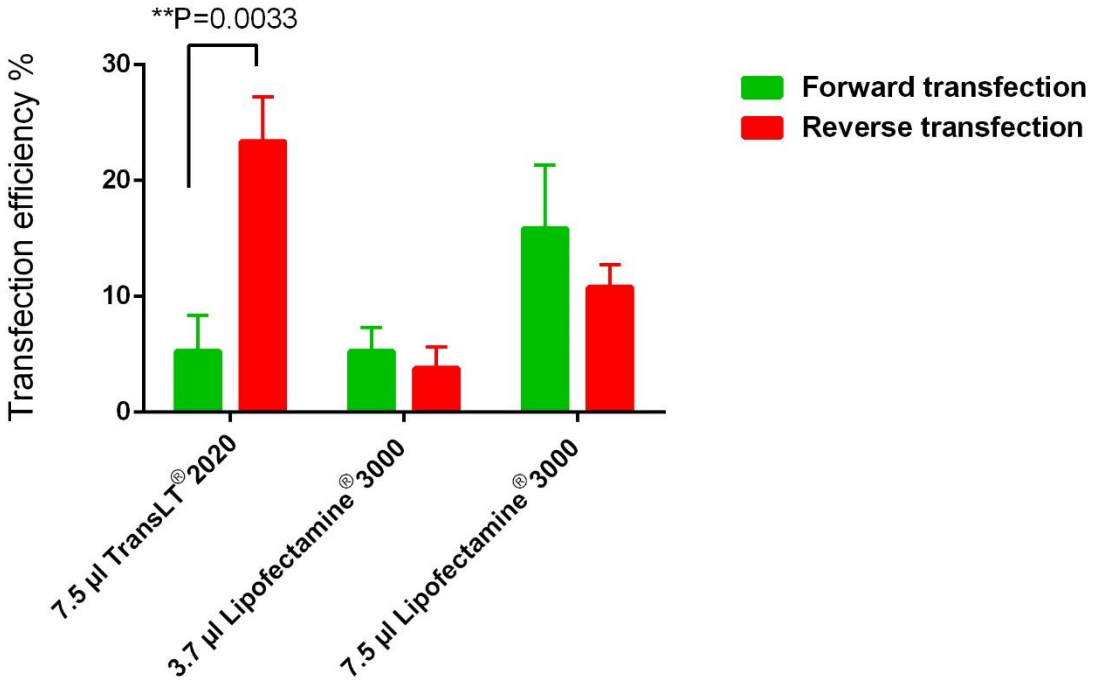


Figure 3.7: hiPSC transfection efficiency with lipid-based transfection. Forward and reverse transfection was conducted with TransIT®2020 and Lipofectamine®3000**.** The graph shows percentage of cells expressing GFP 24 hours after transfection (n=3).

3.2.3.2 Analysis of iPSC nucleofection

The next method tested was the Amaxa® nucleofection system, which introduces plasmid DNA directly into the nucleus and has been reported to produce high transfection efficiency in in a range of different cell types (Hamm et al. 2002, Dityateva et al. 2003, Lenz et al. 2003). A single cell suspension of 6.5-7.7x10^5 cell/mL iPSCs was prepared by using 1x trypsin/EDTA and five nucleofection programs (A-012, A-013, A-023, A-027 and B-016) were tested using the Amaxa® human stem cells nucleofector® starter kit. pmaxGFP (3.4kbp) was used as a reporter for transfection efficiency. Of note, the first time the iPSCs were nucleofected a fault with the system caused the shock to be applied twice resulting in increased efficiency. In subsequent testing two shocks were therefore applied for each of the nucleofector programs tested. There was no observable difference in cell viability between the single or double applications of the shock or between the different programs with all conditions tested showing approximately 60% confluence after 24h and full confluence after 5 days **(Figure 3.8, C)**. Transfection efficiency was determined by counting the percentage of cells expressing GFP 24h after electroporation using an Evos® fluorescent microscope **(Figure 3.8, D)**. The highest transfection efficiency achieved was 81.3% by applying two shocks of program B-016 with Amaxa® human stem cells nucleofector® solution 2 **(Figure 3.9)**.

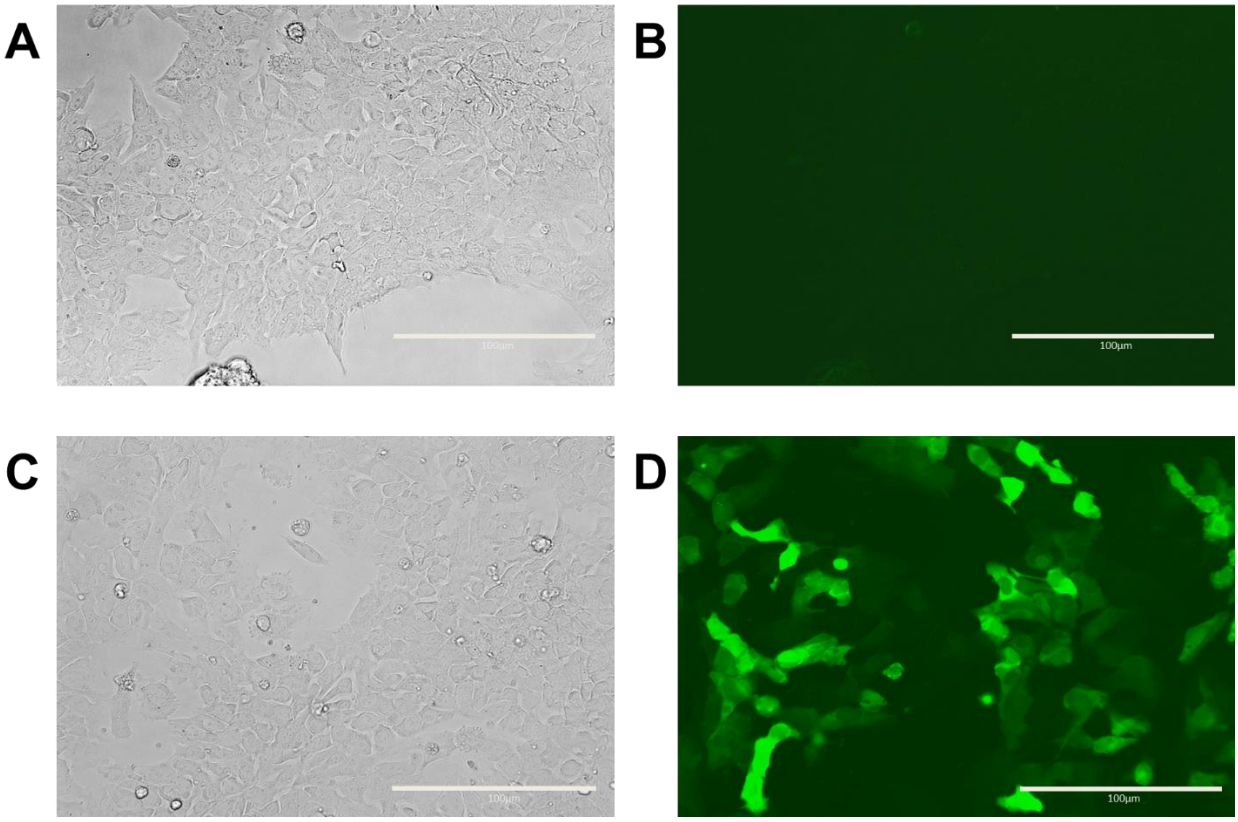


Figure 3.8: Example of iPSCs following nucleofection. A and **B** show the empty-pmaxGFP control after 24h of nucleofection, **(the scale bar indicates 100µm)**. **C** and **D** show the 24h post-transfected iPSCs with pmaxGFP, **(the scale bar indicates 100µm)**.

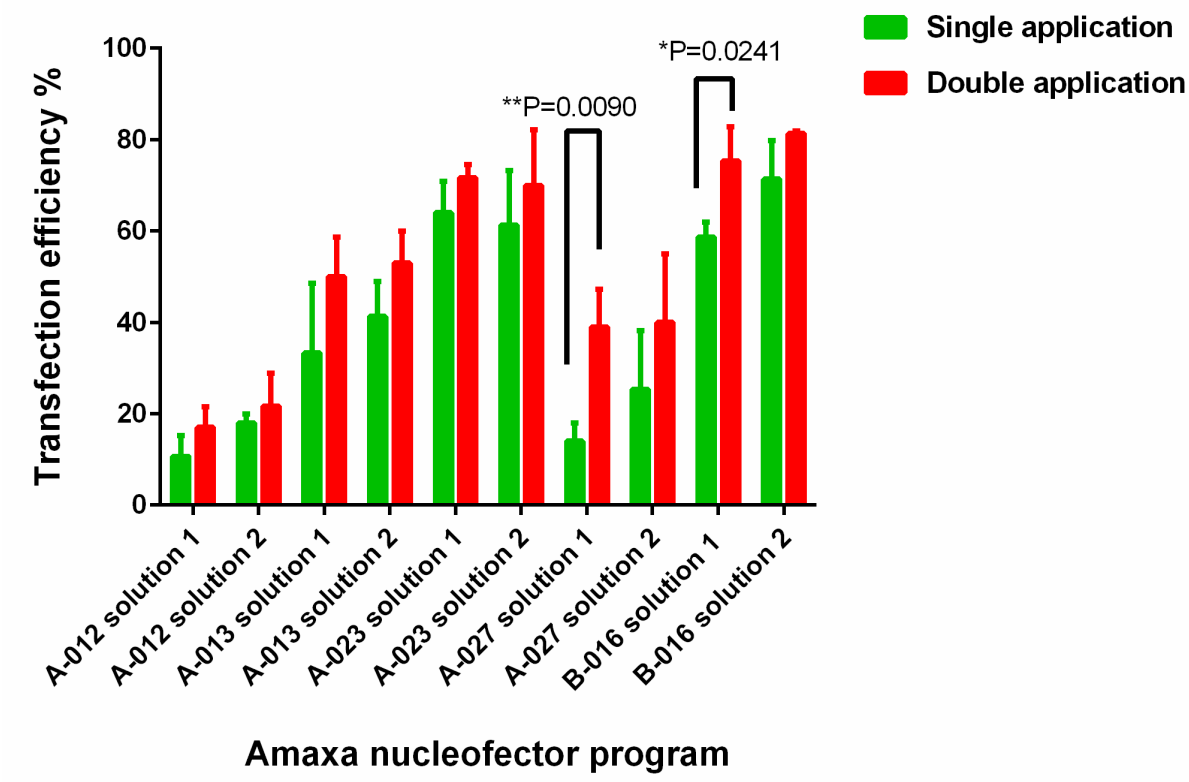


Figure 3.9: Optimization of iPSC nucleofection. Five programs (A-012, A-013, A-023, A-27 and B-016) with two solutions (sol1 and sol2) were tested; double application indicates the shock was delivered to iPSCs twice (n=3).

3.2.3.3 Analysis of electroporation

Efficiency of electroporation was tested using the NEPA21 system (Nepagene), which has cost benefits over the Amaxa® system. A single cell suspension of 6.5-1.0x10^5 cell/mL iPSCs was prepared by using 1x trypsin/EDTA. After centrifuging, the pellets were washed twice with Opti-MEM I medium (Gibco, Fisher Science) and resuspended using 100µl with the same medium of washing as in **the section 2.1.3.3**, the cells then presented to the electroporation process. Electroporation efficiency was assessed using the pmaxGFP reporter plasmid and 5 programs that had been optimised by the manufacturer for iPSCs **(Table 2.5, chapter 2)**. Cell viability was high with 70% confluence 24h **(Figure 3.10)** and full confluence after 3 days. The highest transfection efficiency achieved was approximately 81% using program 7 **(Figure 3.11)**. Due to increased cell viability, lower cost and ease of use compared with to the Amaxa® system the NEPA21 was used for all downstream experiments.

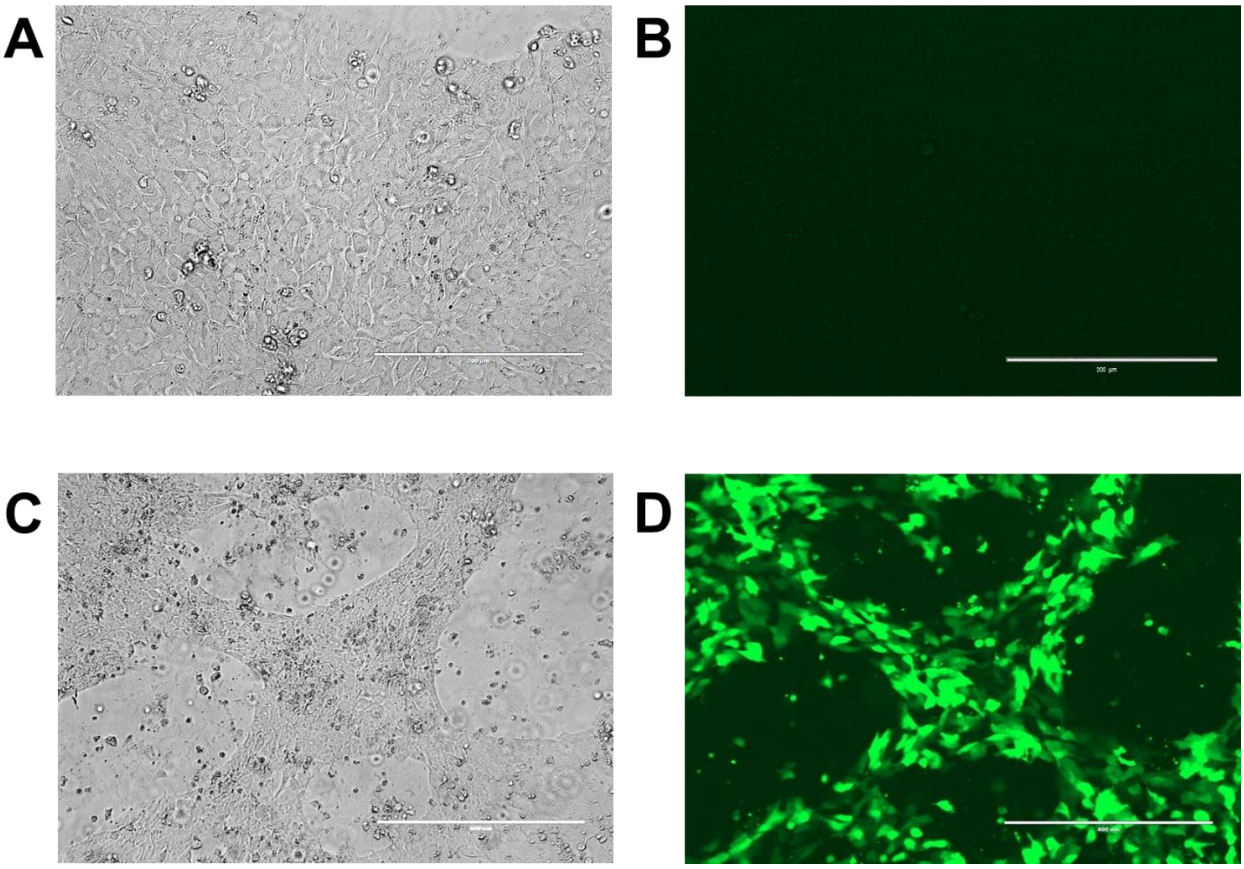


Figure 3.10: Electroporation of hiPSCs with pmaxGFP plasmid by using program seven with Opti-MEM medium. **A** and **B** exhibit the observation after 24h of electroporation without the plasmid as a control, **(the scale bar indicates 200µm)**. **C** and **D** exhibit 24h post-transfected iPSCs with pmaxGFP plasmid, **(the scale bar indicates 400µm).**

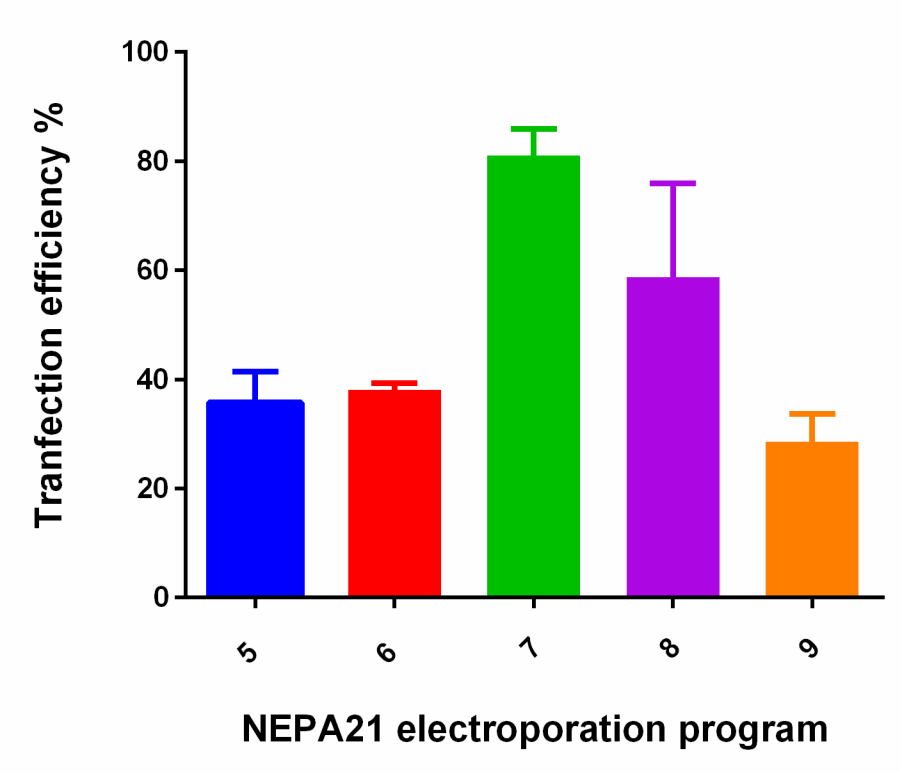


Figure 3.11: Optimization of iPSC electroporation. Five programs (5, 6, 7, 8 and 9) were tested using the NEPA21 electroporator (n=3) **(Refer to table 2.5)**

3.2.4 Clonal isolation of iPSCs

The endpoint of iPSC targeting using genome editing requires the clonal isolation of successfully edited cells. To optimise the clonal isolation of our iPSCs, the method reported by Peters *et al*. (2013) was followed with some modifications. Matrigel™-coated 9cm plates were coated with Matrigel™ containing ROCK inhibitor Y-27632 (Lot. Doo155678, China) at 10µM, in order to increase the plating efficiency of iPSCs (Pakzad et al. 2010), and iPSCs seeded at 10^3, 10^4 and 10^5 cells/ml **(section 2.1.1.6)**. The day after seeding the media of the 10^4 and 10^3 cell/ml seeded dishes was replaced with a mixture of 50% mTesR™1 media and 50% conditioned mTesR™1 media (collected from 40-50% confluent iPSCs and filtered through a 0.22µm membrane filter (Millex®GP) and then stored at 20°C. The media of the 10^5 cell/ml seeded dish was replaced with mTesR™1 media. After approximately 10 days the iPSCs formed colonies roughly 2.5mm in diameter **(Figure 3.12)**.

Seeding at 10^3 cells/ml **(Figure 3.12, D)** produced isolated colonies of an appropriate size for picking. Individual colonies were disrupted with the tip of a 200µl pipette (adjusted to 45µl) and then transferred to one well of a 96-well Matrigel™-coated plate containing 120µl fresh mTesR™1 media **(Figure 3.13, D)**. The following day picked colonies had settled and started to grow **(Figure 3.13, E, F)**. After 5-10 days the colonies reached 80-95% confluence **(Figure 3.13, G)**. In the genome editing experiments DNA would be harvested and the cells frozen to allow sequence identification of correctly targeted cells. To test the likely rate of attrition using this process the cells were frozen and revived.

In this process, due to the size of the well of the 96-plate, and to avoid losing of apportion of the clone and fasteing the process, the detaching enzyme and freezing solution were replaced. the freezing was conducted as in **the section 3.2.4** using mFreSR™ (Stem Cell technology) solution, after detached the cells from the culturing plate by using Accutase™ (Stem cell technology), about two third of the cells mixed with the freezing solution and then placed in new 96-well plate then stored at 80°C. Revival and expansion was performed as described in **sections 2.1.1.10** and **2.1.1.11** respectively. Approximately 92.5% of wells were recovered upon revival and considered sufficient for these experiments.

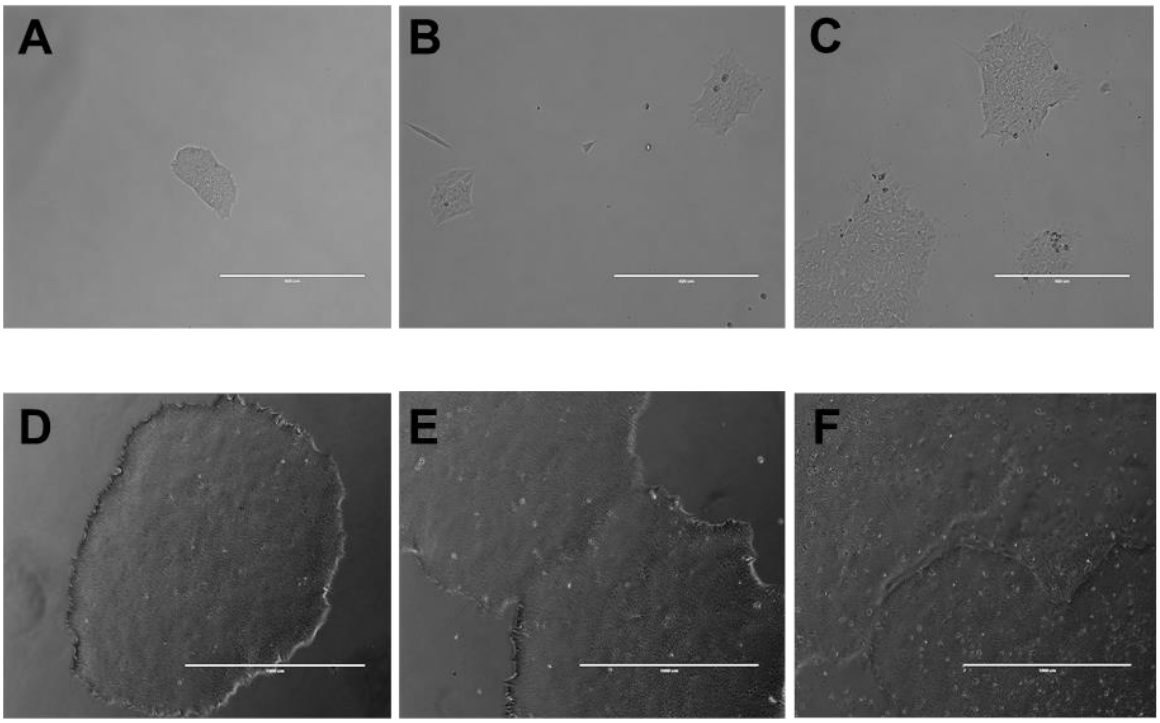


Figure 3.12: iPS Colonies grown from single cells in 90mm Petri dishes. **A, B** and **C** show iPS colonies at day 5 following seeding at 10^3, 10^4 and 10^5 cells/ml respectively **(the scale bar indicates 400µm). D, E and F** show iPS colonies at day 10 following seeding at 10^3, 10^4 and 10^5 cells/ml respectively **(the scale bar indicates 1000µm).**

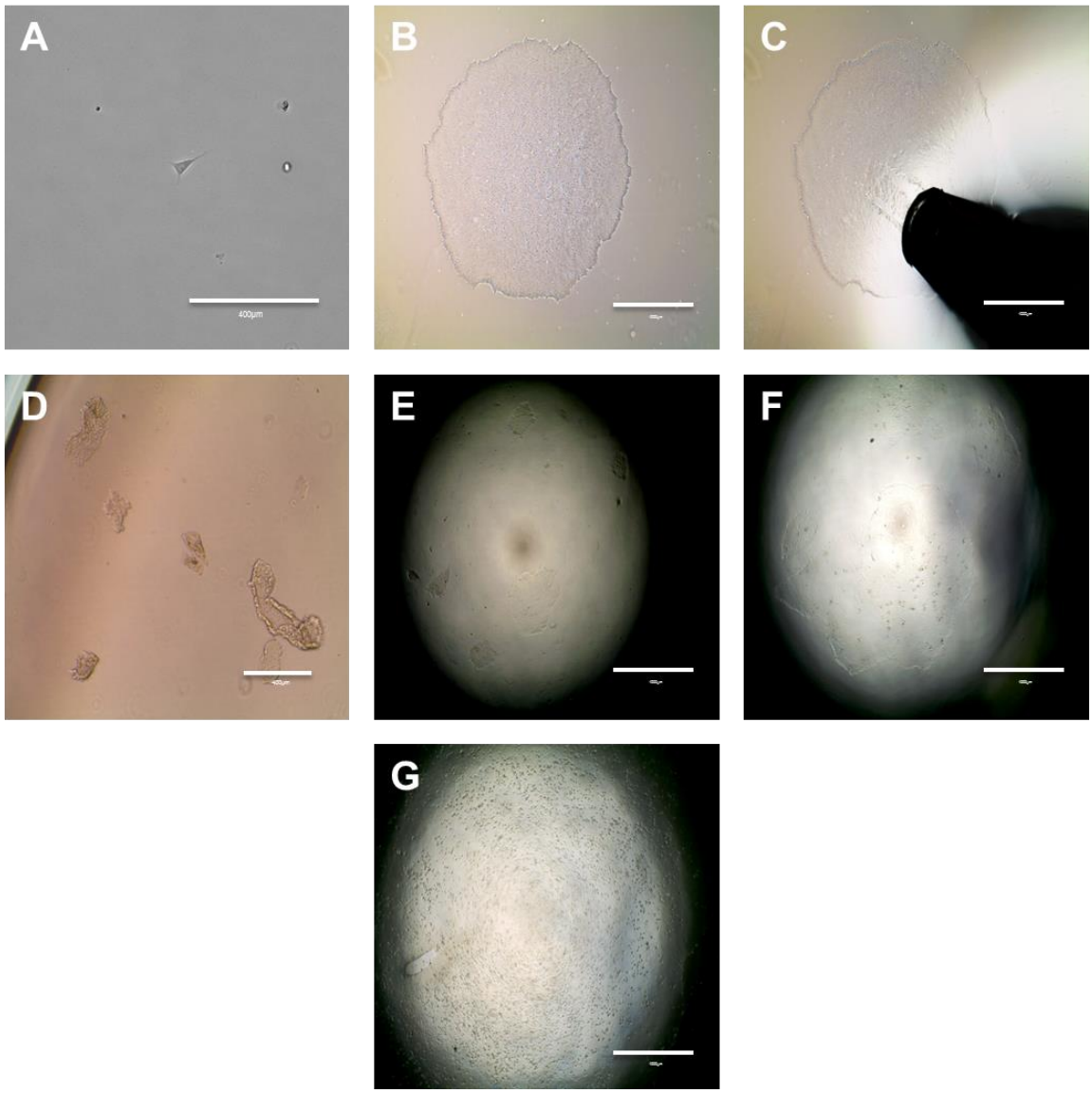


Figure 3.13: Clonal isolation of iPSCs. (A) A single cell seeded on the 90mm Matrigel™ coated Petri dish **(the scale bar indicates 400µm, Evos Fl Microscope)**. **(B)** A single colony on the 90mm Matrigel™-coated Petri dish **(the scale bar indicates 1000µm, Evos Xl Microscope)**. **(C)** Colony picking by pipette **(the scale bar indicates 1000µm, Evos Xl Microscope)**. **(D)** Picked colony after transfer to Matrigel™-coated 96-well plate **(the scale bar indicates 400µm, Evos Xl Microscope)**. **(E)** Established and growing clone the second day following picking **(the scale bar indicates 1000µm, Evos Xl Microscope)**. **(F)** A single growing clone on the Matrigel™ 96-well plate 4 days after picking **(the scale bar indicates 1000µm, Evos Xl Microscope)**. **(G)** A single clonal colony **(the scale bar indicates 1000µm, Evos Xl Microscope)**.

3.3 Discussion

The overall objective of this study requires the CRISPR/Cas9 genome editing of hiPSCs. As this was the first time these experiments will be performed by our group the aim of the experiments described in this chapter was to establish iPSC culture and quality control measures and perform optimization in preparation for the delivery of the CRISPR/Cas9 machinery. A single iPSC cell line, GM23720*B (Coriell), was obtained for use in this study and initially cultured using the same reagents and conditions in which they had previously been maintained. The iPSCs were successfully maintained in an undifferentiated state under these culture conditions with the cells growing in characteristic tightly associated compact colonies with individual cells having a high nuclear to cytoplasmic ratio with prominent nucleoli. The hiPSCs were very similar in morphology to undifferentiated human ESCs (Xu 2006). Differentiated cells were observed at an estimated percentage of approximately 20% and efforts were made to further optimize iPSC culture in order to maintain a higher proportion of undifferentiated cells. Changing the method of cell passage to incorporate the enzyme free reagent ReLeSR™, which causes the preferential dissociation of undifferentiated colonies rather than differentiated cells (Mitalipova et al. 2005), rather than the enzyme dispase II increased the proportion of undifferentiated cells to around 95%. Morphology of iPSCs alone does not confirm the undifferentiated iPSCs, therefore, cell-surface markers of PSCs were assessed by flow cytometry. iPSCs were positive for expression of SSEA-3 (96.13%) and TRA-1-81 (75.66%), markers of differentiation whereas they exhibited very little expression of SSEA-1 (2.11%), a marker of differentiation. To confirm iPSCs were maintained in an undifferentiated state they were visually inspected every day and by flow cytometry every 5-10 passages. Karyotyping of iPSCs was performed to show that they retained a normal chromosomal complement and ensure that abnormal iPSCs, which may have greater proliferation capacity (Draper et al. 2004, Caisander et al. 2006), did not take over the culture. However, the current karyotyping is an inappropriate method and a full chromosome alignment would be need to ensure 22 pairs of chromosome plus the sex chromosome are present, as well as absent of any change in chromosome content and structure such as translocation, inversion and a large duplication and deletion.

The genome editing machinery (CRISPR gRNA and Cas9 nuclease) used to generate knockout iPSCs will be delivered as plasmid DNA in this study and high expression is

required to achieve a reasonable targeting efficiency. Various transfection techniques were therefore tested to identify the best approach and over 80% efficiency was achieved using both a NEPA21 electroporator and Amaxa® nucleofector. Due to cost, cell viability and ease of use the NEPA21 was chosen for the experiments discussed in **chapters 4** and **5**. Flow cytometry could be used to quantify both GFP positive population and the absolute cell number comparison with transfected control cells and increased genome editing efficiency can also be achieved by transfecting a fluorescently-tagged Cas9 protein and enriching the cells with high levels of expression by fluorescence-activated cell sorting (Ran et al. 2013), however, this wasn't available for this study. Optimization of the clonal isolation showed seeding of iPSCs at 10^3 cells/ml to produce isolated colonies of an appropriate size for picking and the colonies showed good survival through each stage of the protocol.

In conclusion, hiPSCs culture has been successfully established and the techniques required for CRISPR/Cas9 genome editing have been optimized. In the next two chapters CRISPR/Cas9 will be used to generate hiPSC models to investigate the functional consequences of known (*TERT*) and candidate (*ACYP2*, *TSPYL6*) genes involved in telomere maintenance.

Chapter 4

CRISPR/Cas9-mediated Knockout of *TERT* in hiPSCs

4.1 Introduction

In somatic cells telomeres shorten with every cell division and this progressive shortening of telomeres eventually results in cell senescence and death (Dimri et al. 1995, Oeseburg et al. 2009, Shelton et al. 1999). In stem cells, and other cell types capable of self-renewal, telomere length is maintained by the action of the enzyme telomerase. This chapter will describe CRISPR/Cas9 targeting of *TERT* in hiPSCs and the effects on telomere length maintenance and hiPSC phenotype.

4.1.1 Telomerase in Stem Cells

The indefinite proliferation and self-renewal capacity of iPSCs is due to their constitutive telomerase activity and telomerase can therefore be considered a stem cell marker. Consequently, the expression, regulation and importance of telomerase has been the focus of a number of studies (Marion et al. 2009, Huang et al., 2011, Wang et al., 2012).

ESCs have long telomeres, high telomerase activity and maintain their telomere length, whilst telomerase expression increases and is an important component of the reprogramming of somatic cells into hiPSCs (Takahashi et al., 2007, Stadtfeld et al. 2008). Upregulation of telomerase activity occurs during the initial reprogramming of somatic cells into miPSCs and precedes the expression of pluripotency genes such as Oct4, Nanog, and Sox2 (Wang et al., 2012). During reprogramming telomeres lengthen and undergo epigenetic changes and acquire chromatin marks, including a reduction in trimethylated histones H3K9 and H4K20, characteristic of the more open chromatin structure of ESC telomeres (Marion et al. 2009). In addition, both the levels of telomerase activity and telomere length is heterogenous between miPSCs and longer telomeres correlate with increased self-renewal capacity and pluripotency (Huang et al., 2011, Wang et al., 2012).

4.1.2 Telomerase deficient Stem cells

The importance of telomerase activity in stem cells has been demonstrated by several studies using telomerase deficient mouse and human ESCs and iPSCs. Marion *et al* (2009) demonstrated that miPSCs could be derived from fibroblasts of *Terc*$^{-/-}$ mice, showing that telomerase is not essential for reprogramming although the reprogramming of fibroblasts from later generations, which have shorter telomeres, was less efficient (Marion et al 2009). The results of Marion *et al* (2009) also indicated that telomerase is the primary mechanism of telomere length elongation during reprogramming as the telomeres of *Terc*$^{-/-}$ miPSCs were shorter than the parental fibroblasts and reintroduction of telomerase resulted in lengthening of telomeres and rescued reprogramming efficiency. In similar experiments, Wang *et al* (2012), observed comparable effects on reprogramming and observed a reduction in telomere length of up to 50% during long-term culture (27 passages), while in experiments with *Tert*$^{-/-}$ miPSCs, Kinoshita *et al* (2014) observed telomere shortening and increased incidence of chromosomal abnormalities with extended culture. In addition, Wang *et al* (2012), found evidence of telomerase-independent telomere elongation during reprogramming and through early passaging of *Terc*$^{-/-}$ miPSCs. At these stages, Wang *et al* (2012) found higher rates of telomere sister chromatid exchange, a sign of active recombination based telomere lengthening seen in ALT, as well as changes in expression of *Dnmt3a* and *b* and *Zscan4*, which encode key regulators of this process (Wang et al 2012, Allsopp 2012). Furthermore, knockdown of *Zscan4*, in wild-type ESCs resulted in telomere shortening of around 10% with increased passaging (Zalzman et al 2010). These studies show that telomerase and ALT are both involved in telomere lengthening during reprograming and maintenance of telomere length during long-term culture.

In terms of pluripotency, telomerase is not necessary for miPSCs to generate chimeras and teratomas, however, cells with longer telomeres show higher efficiency than those with short telomeres (Huang et al 2011, Kinoshita et al 2014). Together, these observations suggest that telomerase plays a critical but none essential role in the induction, self-renewal capacity and pluripotency of miPSCs and is necessary for maintaining telomere length, which is essential for true pluripotency, while the ALT pathway is also involved but less important.

The above studies involved the investigation of telomerase activity in miPSCs. Although the sequence of telomeres and their function is the same between mouse and human there are important differences (Calado and Dumitriu et al. 2013). Mouse telomeres are up to 10 times longer than those in humans and homozygous loss of telomerase has relatively minor phenotypic consequences in mice (at least during early generations), whereas heterozygous mutation in components of telomerase cause premature aging disorders such as dyskeratosis congenital (Blasco et al. 1997, Chiang et al. 2004, Hao et al. 2005, Calado et al. 2013). To the best of our knowledge there have been no investigation of the effects of knockout of telomerase components in human stem cells, however, knockdown experiments have been performed (Yang et al 2008) and hiPSCs have been derived from patients with the telomere deficiency syndromes dyskeratosis congenita and aplastic anemia (Batista et al. 2011, Winkler et al 2013). Over expression of a dominant-negative *TERT* expressing construct was reported to remove the self-renewal capacity of ESCs while siRNA reduced ESC proliferation, increased apoptosis, altered morphology and increased expression of differentiation markers, suggesting human stem cells might be more sensitive to telomerase deficiency in comparison to mouse stem cells (Yang et al 2008). Batista and colleagues studied the effects of telomerase deficiency in hiPSCs derived from dyskeratosis congenital patients with mutations in *TERT*, *TCAB1* and *DKC1* (Batista et al 2011). Mutant hiPSCs were described as being resistant to reprogramming. *TERT* mutant hiPSCs exhibited impaired telomere lengthening during reprogramming compared to wild-type hiPSCs while the *TCAB1* and *DKC1* mutant hiPSCs were shorter than their parental fibroblasts and showed no evidence of telomere elongation. Extended culture of *DKC1* mutant hiPSCs was reported and telomere length was shown to decrease until passage 36 when the hiPSCs exhibited spontaneous differentiation, which Batista *et al* (2011) concluded was due to critically short telomeres. Prolonged culture of the *TERT* and *TCAB1* mutant hiPSCs was not assessed beyond passage 10. Winkler *et al* (2013) generated hiPSCs from patients with aplastic anemia caused by mutations in *TERT* and *TERC*. Similar to the study by Batista *et al*, telomere elongation during reprogramming was reduced. Telomere length increased over the first 15 passages following reprogramming, before stabilising. Prolonged culture beyond this point was not investigated, so it is not clear if these cells would be able to maintain telomere length indefinitely. Winkler *et al* (2013) also

investigated ALT activity, but unlike the studies in miPSCs, found no evidence that ALT contributed telomere extension during iPSC reprogramming

4.1.4 Aims and objectives

As described in the introduction, the CRISPR/Cas9 genome editing has attracted intense interest as an affordable, easy to engineer and programmable method for the precise targeting in cells and organisms. The first reports of successful engineering of mammalian cells emerged less than one year before the start of our study (Mali et al 2013, Cong et al 2013) and since then there have been a vast number of reports incorporating CRISPR/Cas9 technology to answer a range of biological questions. The objective of this chapter was to use CRISPR/Cas9-mediated genome editing to generate *TERT* knockout hiPSCs and investigate the effects on telomere maintenance and hiPSC phenotype. The results of these experiments would then inform the targeting of novel genes genetically associated with telomere length.

Aims;

- To use CRISPR/Cas9 genome editing to generate *TERT* knockout hiPSCs
- To investigate the effects of *TERT* knockout on telomere length and iPSC phenotype.

4.2 Results

4.2.1 Expression of telomerase reverse transcriptase (*TERT*) gene in iPSCs

We first performed RT-PCR to confirm *TERT* expression in the GM23720*B hiPSC cell line. Optimised primers were used to amplify *TERT* from cDNA alongside no template control (NTC) and reverse transcriptase negative (NRT) controls. A product of the expected size was detected (**Figure 4.1**).

4.2.2 Design of sgRNA for *TERT* knockout

The *TERT* sgRNA used in this study was designed by our collaborators Horizon Discovery Group plc (Cambridge, UK) using proprietary software. Horizon designed and tested five sgRNAs in HEK293 cells using the *Streptococcus pyogenes* Cas9 expressing plasmid PL1. The most efficient sgRNA, 5'- GCTGCGCAGCCACTACCGCG-3', had an estimated cutting efficiency of approximately 39%, as determined using the Surveyor assay **(Figure 4.2)**. The sgRNAs were designed to target the gene early in its coding sequence, thereby making it more likely that any frameshift mutation would remove protein function. **(Figure 4.3)**. Analysis of potential off-target effects using the Desktop Genetics CRISPR design software (www.deskgen.com) predicts zero coding or non-coding off-targets with up to two mismatches. All other possible off-target sites included at least one mismatch in the sgRNA seed region.

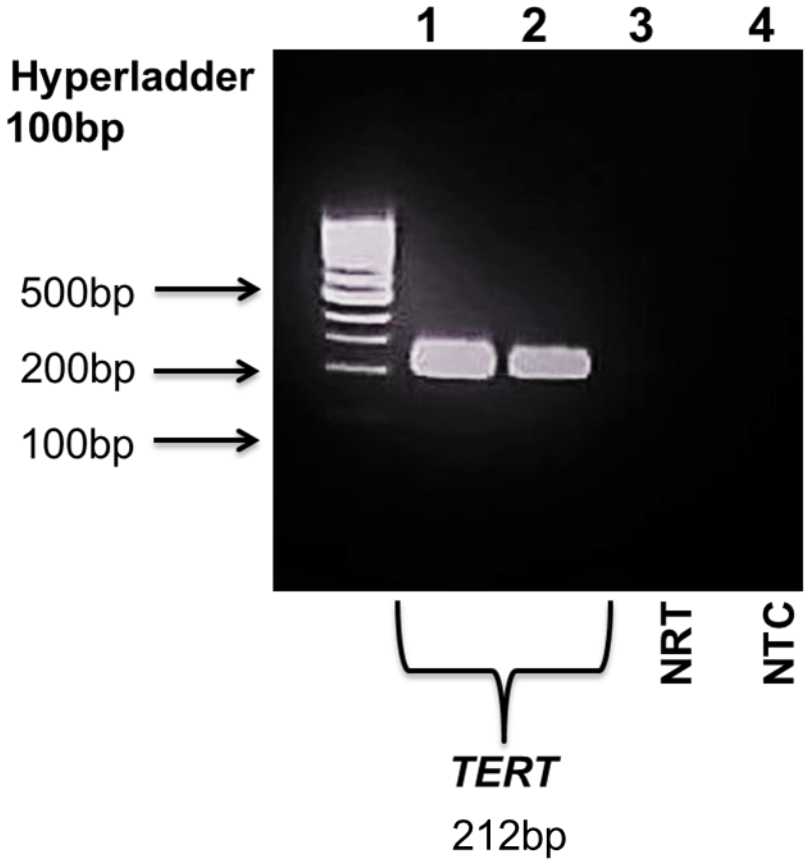


Figure 4.1: RT-PCR shows the expression of *TERT* in GM23720*B iPSCs. Lane 1 and 2, no reverse transcriptase (NRT, lane 3) and no template control (NTC, lane 4).

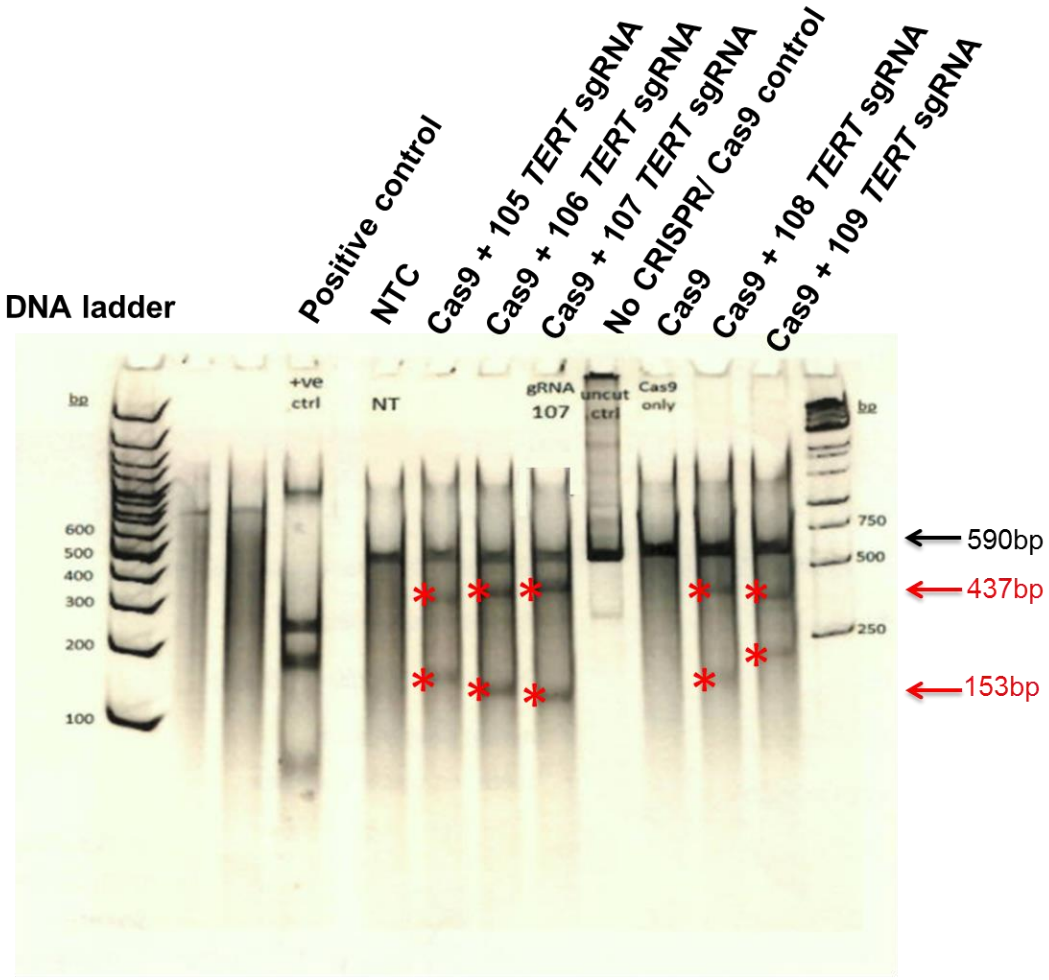


Figure 4.2: Surveyor screening assay of Horizon Discovery Group plc for *TERT*. Horizon Discovery Ltd performed a Surveyor assay to select the most efficient sgRNA. Each candidate sgRNA was co-transfected into HEK293 alongside Cas9. Controls comprise a Cas9 only transfection, a no template control (NTC), No CRISPR/ Cas9 control and a mismatch positive control. PCR products of the target genomic region (or positive control) were subjected to a Surveyor assay and resolved on an acrylamide gel. Digestion products (red asterisks) of the expected size were detected for each sgRNA. Densitometry was performed to estimate cutting efficiency.

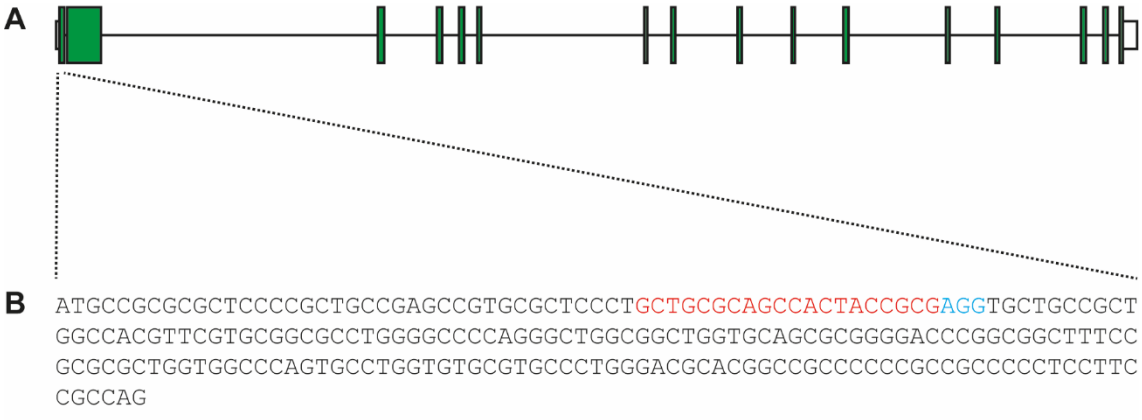


Figure 4.3: sgRNA target site in *TERT*. (**A**) Schematic illustration of the *TERT* transcript showing the exons and introns; green bars are exons separated by the black line which represents introns. **(B)** Genomic DNA sequence of exon one of *TERT*. The guide RNA sequence is highlighted in Red. The protospacer adjacent motif (PAM) is highlighted in blue.

4.2.3 CRISPR/Cas9 targeting of *TERT*

hiPSCs were transfected with Cas9 PL1-plasmid on its own or with Cas9 and CRISPR *TERT* sgRNA using the NEPA21 electroporator (Nepagene). After transfection, the cells were allowed to recover and culture normally. When the cells reached confluence, they were passaged and a portion of cells collected for DNA extraction and screening. When the cells next reached confluence they were banked in liquid nitrogen.

4.2.3.1 Screening using the Surveyor Assay

PCR primers were designed to amplify a 590bp region of genomic DNA that contained the sgRNA target sequence and gradient PCR was performed in order to determine the optimal annealing temperature **(Figure 4.4)**. The 35 cycle PCR reaction parameters were set up included annealing temperature gradient step which was ranged from 63-70°C. The rest of parameters were as in **section 2.1.4.10**. A smaller non-specific band was observed at lower annealing temperatures, while at higher temperatures there was less amplification of the expected product. Therefore, an intermediate temperature of 65°C was chosen for all further reactions.

Genomic DNA from targeted and control cells was then used as template for PCR and the Surveyor assay was then carried out to confirm successful targeting of *TERT* by CRISPR/Cas9. The Transgenomic Surveyor® mutation detection kit contains the Surveyor nuclease, a proprietary T7 endonuclease, which digests double-stranded DNA that contains mismatches. The position of the *TERT* sgRNA target site in the PCR product should result in two digestion products of 153bp and 437bp and products of the expected size were seen in the samples from cells transfected with Cas9 and the *TERT* sgRNA, but not in Cas9 only transfected cells **(Figure 4.5)**. A control CG with mismatch provided with the kit was included as a positive control.

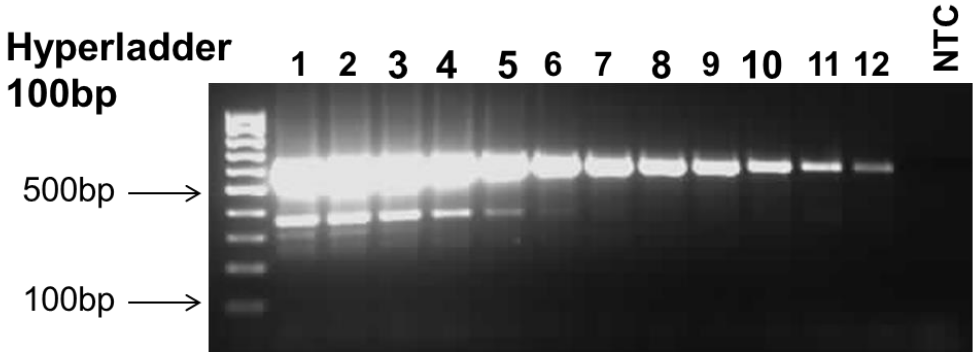


Figure 4.4: Gradient PCR reactions of annealing temperature for *TERT*. shows gradient annealing temperature for *TERT* gene visualized by gel electrophoresis on 1.5% agarose. 100bp ladder as marker for DNA bands and no template control (NTC) was used. Gradient annealing temperature range lane1; 63.1ºC, lane2; 63.3ºC, lane3; 63.6ºC, lane4; 64.1ºC, lane5; 65.1ºC, lane6; 66.0ºC, lane7; 66.9ºC, lane8; 67.9ºC, lane9; 68.9ºC, lane10; 69.5ºC, lane11; 69.9ºC and lane12; 70.0ºC.

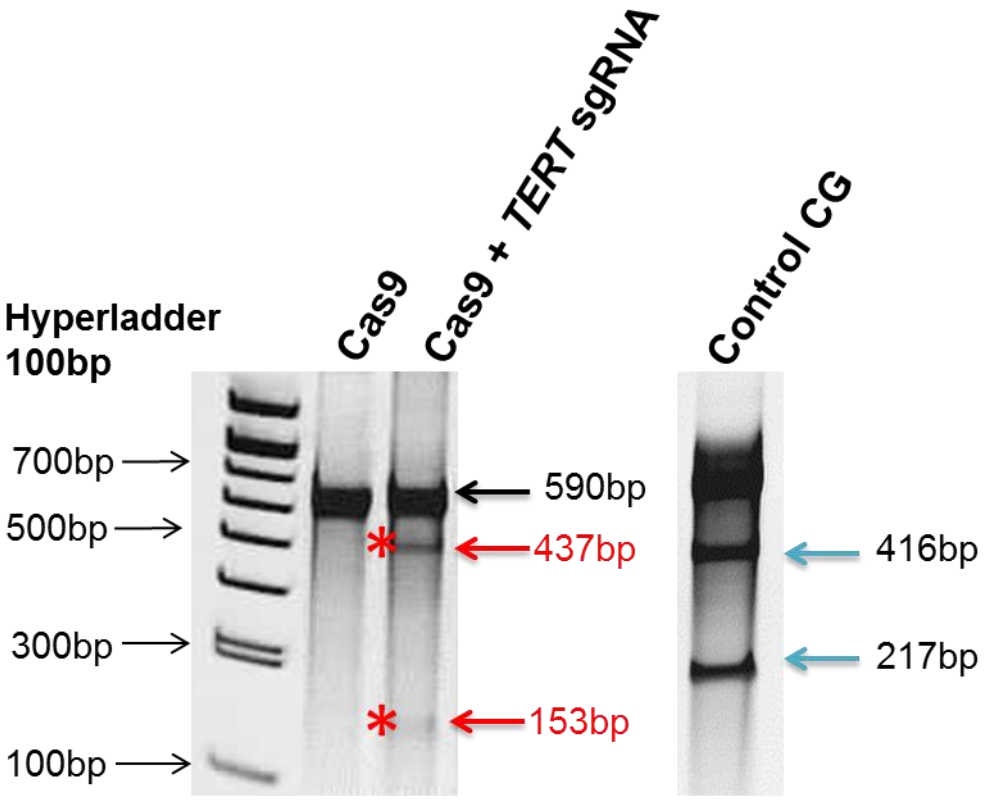


Figure 4.5: Surveyor screening assay. Surveyor assay of Cas9 only, Cas9 plus *TERT* sgRNA and CG mismatch control resolved by electrophoresis on an acrylamide gel. Digestion products (red asterisks) of the expected size were detected in the Cas9 plus *TERT* sgRNA lane, but not with Cas9 only. Blue arrows indicate to the digested products of the CG mismatch control.

4.2.3.2 Clonal Derivation and Sequence Confirmation of CRISPR/Cas9 targeted cells

As described in **section 3.2.4**, the targeted cells were seeded at 10^3 cells/ml on a 90mm petri dish. When the colonies reached approximately 2.5mm in diameter (around 7-10 days) they were picked and transferred to a Matrigel™-coated 96-well plate. When the colonies reached confluence, they were passaged into fresh plates and frozen at -80°C and a portion of cells was retained for DNA extraction and Sanger sequencing. As summarised in **Table 4.1**, 35 of 36 colonies survived transfer to 96-well plates and were sequenced. Sequence traces were analysed by the TIDE sequence deconvolution tool https://tide-calculator.nki.nl/ (Brinkman et al. 2014) which generally suggested that 37.1% generated frameshift mutations. In details, 21 of the sequenced colonies were homozygous for both wild-type alleles, 6 colonies contained 1 wild-type allele and 8 colonies had indels in both alleles **(Table 4.1)**. The genotype predicted by TIDE analysis for each clonal line is shown in **supplementary table 1**. Targeting efficiency (proportion of targeted alleles) was therefore 31.4%. All clones were successfully expanded to 6-well plates and archived in liquid nitrogen.

Description	Number of clones
Transferring into 96-well plate from Petri-dish	36
Growing colonies in 96-well plate	35
After sequencing • Wild type • Homozygous • Heterozygous	21 8 6
After culturing in 24-well plate • Wild type • Homozygous • Heterozygous	21 8 6
After passaging and transferring into 6-well plate • Wild type • Homozygous • Heterozygous	21 8 6

Table 4.1: Number and type of mutation in each targeted *TERT* iPSC line. Number and predicted genotype of clonal iPSCs during expansion.

4.2.3.3 Measurement of Telomere Length in *TERT* deficient cells

Following expansion, genomic DNA was purified from n=5-6 wild-type, heterozygous and mutant cells. This time-point corresponded to 30 days post-transfection of the CRISPR/Cas9 components. This time-point was chosen as Wang *et al*. (2012) observed telomere shortening over 27 passages in mouse *Terc* deficient iPSCs. qRT-PCR was performed to determine the effect of *TERT* deficiency on telomere length **(Figure 4.6)**. Heterozygous cells (mean 2.872 ± SD 0.2296) showed a trend towards shorter telomeres compared to wild-type cells (mean 3.733 ± SD 0.4110) (P=0.1189, student t-test) while the homozygous mutant cells (mean 2.522 ± SD 0.3515) had a borderline significant (P=0.049, students t-test) reduction in telomere length compared to wild type.

4.2.3.4 Extended culture of *TERT* knockout cell lines

Two mutant and three wild-type cell lines were taken forward for further experimentation. To confirm mutation of both alleles of the knockout cell lines PCR amplification of the region of genomic DNA (gDNA) containing the sgRNA target site was performed, cloned into pGEM-T easy and at least 4 colonies sequenced **(Section 2.1.4.14)**. *TERT*$^{-/-}$ clone 3 contained a 1bp guanine (G) insertion in one allele and 1bp deletion of same nucleotide in the other allele **(Figure 4.7)**. Both mutations cause a frameshift in the coding sequence and premature stop codons **(Figure 4.8)**. *TERT*$^{-/-}$ clone 25 was homozygous for a 1bp G insertion **(Figure 4.7, 4.9)**. Of note, a third homozygous mutant line, *TERT* clone 11, was also investigated, however, subsequent sequence analysis revealed contamination with wild-type cells, and this line was therefore excluded from the study.

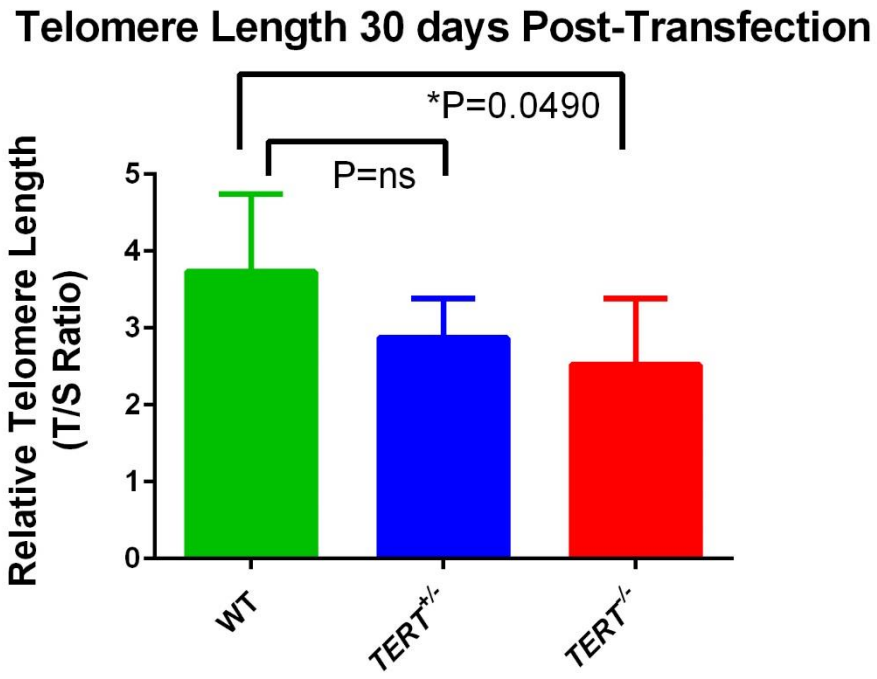


Figure 4.6: Relative telomere length of knockout TERT clones 30 days post-transfection. Relative telomere length (T/S Ratio) of homozygous wild-type (WT), and heterozygous (*TERT*$^{+/-}$) and homozygous (*TERT*$^{-/-}$) (n=5-6).

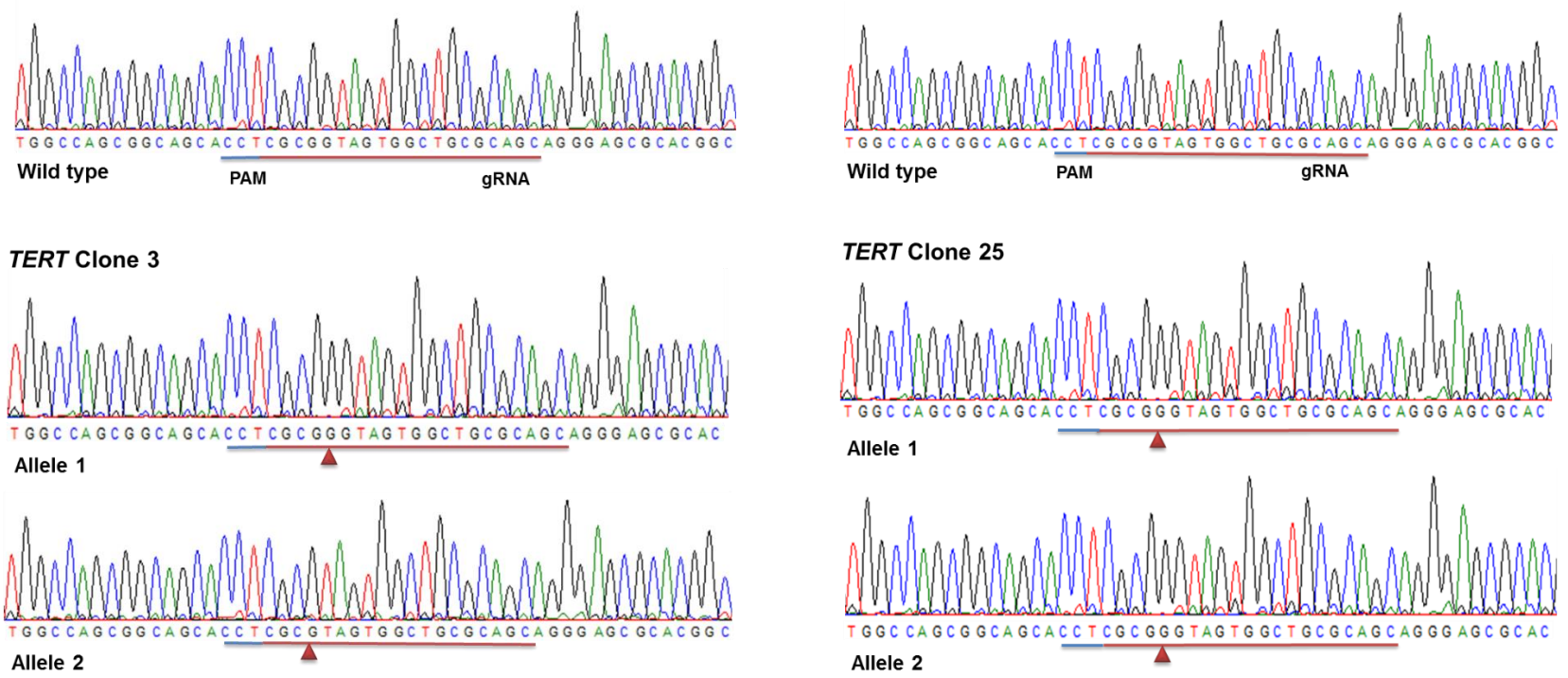


Figure 4.7: Sequence Traces for Wild-type and *TERT*$^{-/-}$ Clones 3 and 25. Sequence traces of each allele of the *TERT*$^{-/-}$ clones. The red bar indicates the sgRNA and the blue bar indicates the PAM, The cut site and consequent indel is shown by the red triangle.

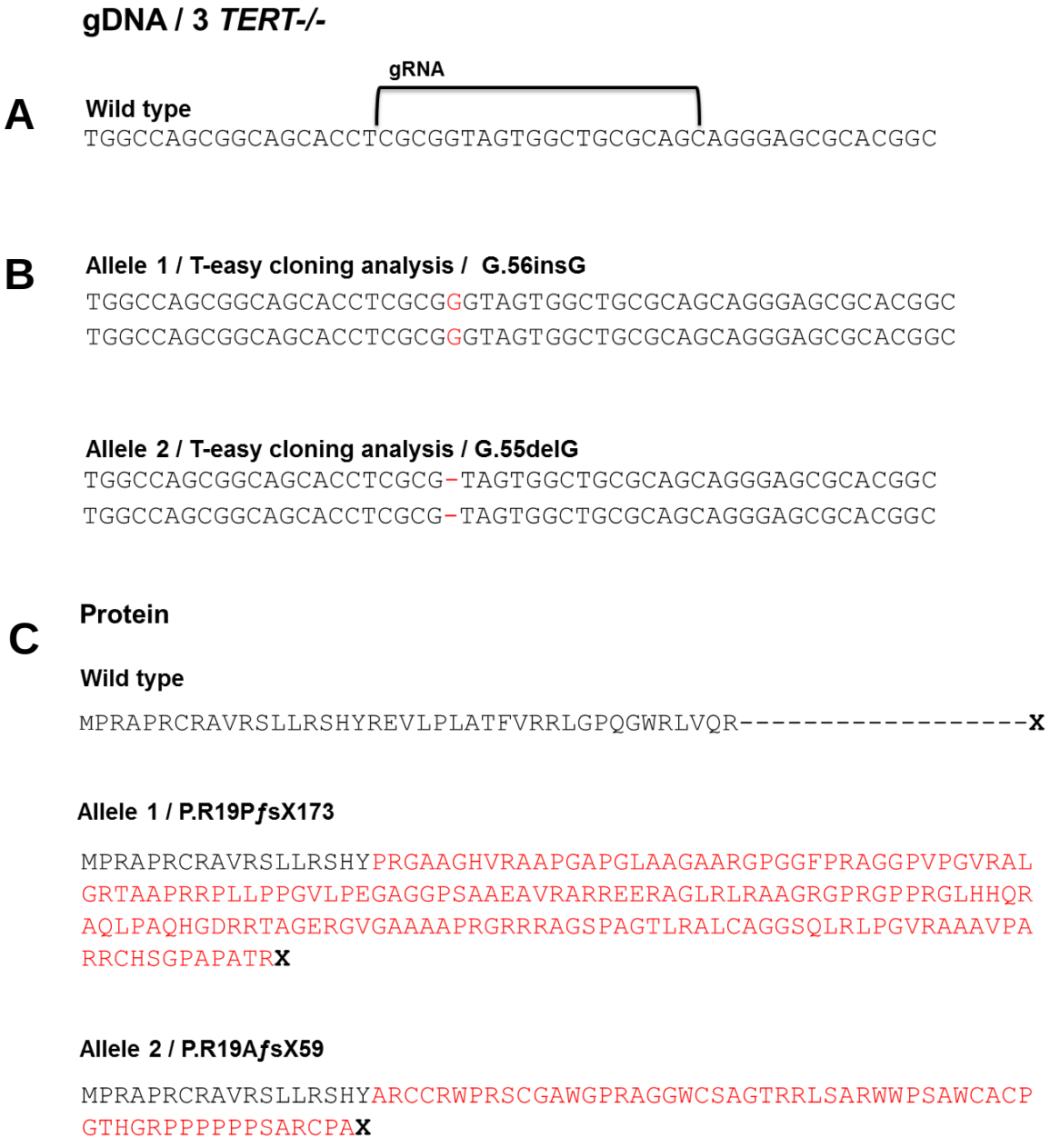


Figure 4.8: gDNA and Protein Sequence of *TERT*$^{-/-}$ Clone 3. (A) gDNA sequence of parental cells. **(B)** gDNA sequence of allele 1 and 2 *TERT*$^{-/-}$ clone 3 showing a 1bp G insertion at cDNA position 56 in allele one and 1bp G deletion at cDNA position 55 in allele two. **(C)** Predicted protein sequencing of each allele of *TERT*$^{-/-}$ clone 3, the frameshift sequencing is highlighted in red and X refers to stop codon. Only the first 41 amino acids of wild type TERT is shown.

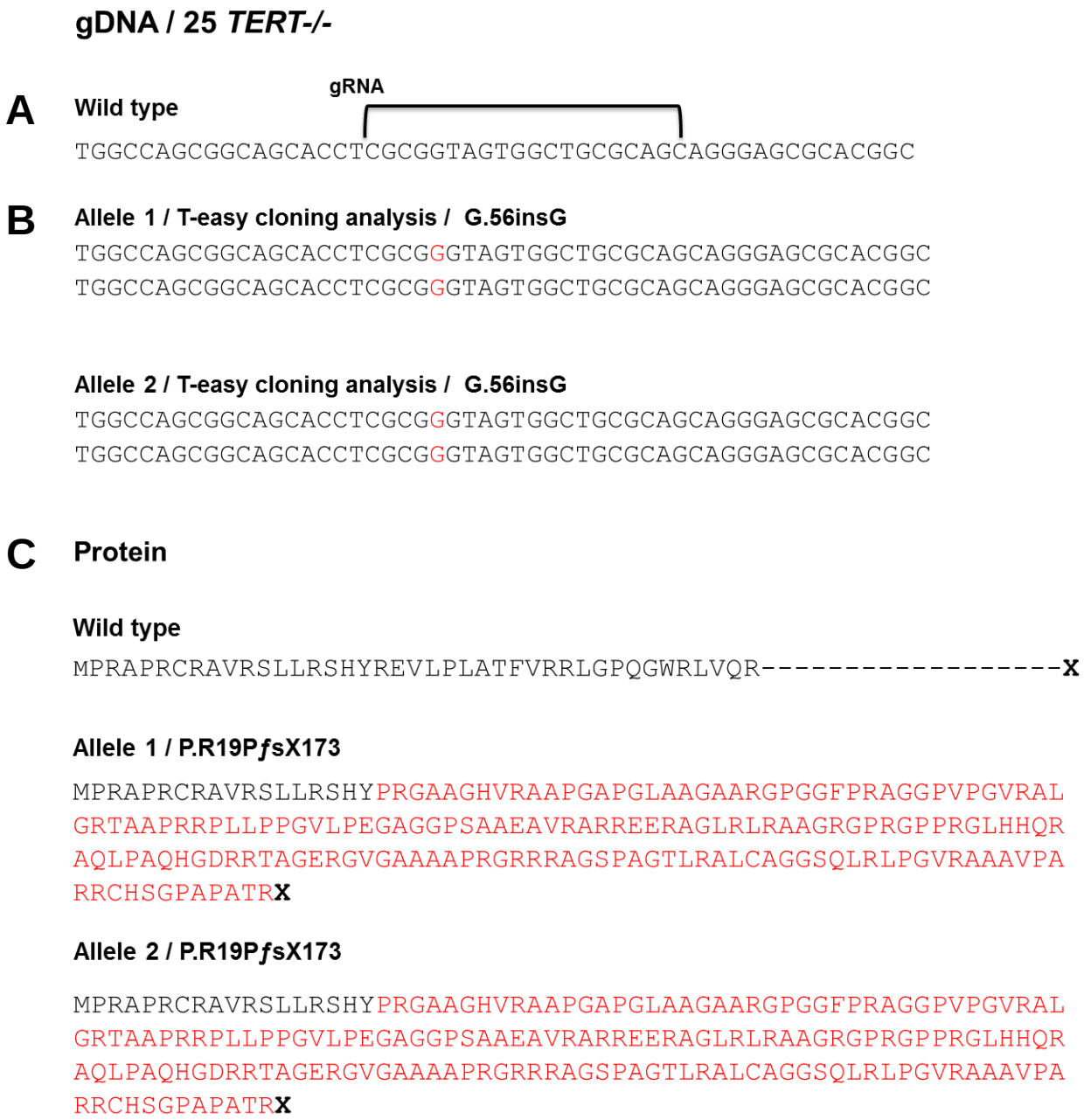


Figure 4.9: gDNA and Protein Sequence of *TERT*$^{-/-}$ Clone 25. **(A)** gDNA sequence of parental cells. **(B)** gDNA sequence of allele 1 and 2 *TERT*$^{-/-}$ clone 25 showing a 1bp G insertion at cDNA position 56 in both alleles. **(C)** Predicted protein sequence of each allele of *TERT*$^{-/-}$ clone 25, the frameshift sequencing is highlighted in red X refers to stop codon. Only the first 41 amino acids of wild type TERT is shown.

The cell lines were cultured for an additional 12 passages under standard iPSC culture conditions with a portion cells retained at each passage for DNA and RNA extraction and archiving. The total time spent in culture was 64 days post-targeting and was equivalent to about 15 population doublings. The mutant lines were resequenced at the end of the experiment to confirm genotype and telomere length was measured at seven time-points **(Figure 4.10)**. We assessed the difference between the wild-type and mutant iPSCs over time by applying a mixed model with sample as a random effect ($P<0.001$, $\beta=-0.304$).

4.2.3.5 Characterization of *TERT* knockout iPSCs

To determine whether the pluripotency of *TERT* knockout human iPSCs was altered, approximately $1x10^7$ cell/ml were fixed at several time-points and analysed by flow cytometry (as described in **section 2.1.2.1 and section 3.2.2.1**). There was no difference in the levels of the mean expression values of the SSEA-3 or TRA-1 cell surface markers between wild-type and $TERT^{-/-}$ iPSCs. However, the differentiated marker SSEA-1 of the clone 3 and 25 homozygous null was unstable with dramatically increasing in clone 3 at day 64. Especially for *TERT* 3, which had a very large increase in SSEA-1 at the later time-points **(Figure 4.11)**. Towards the end of the experiment, the growth rate of clone 3 was also observed to decrease in relation to the wild-type clones and $TERT^{-/-}$ 25, which might be a consequence of the shorter telomere length in this clone. The actual T/S ratio for clone 3 at day 64 was 1.02 compared to 1.45 for clone 25 and between 1.84 and 3.25 for the wild-type clones.

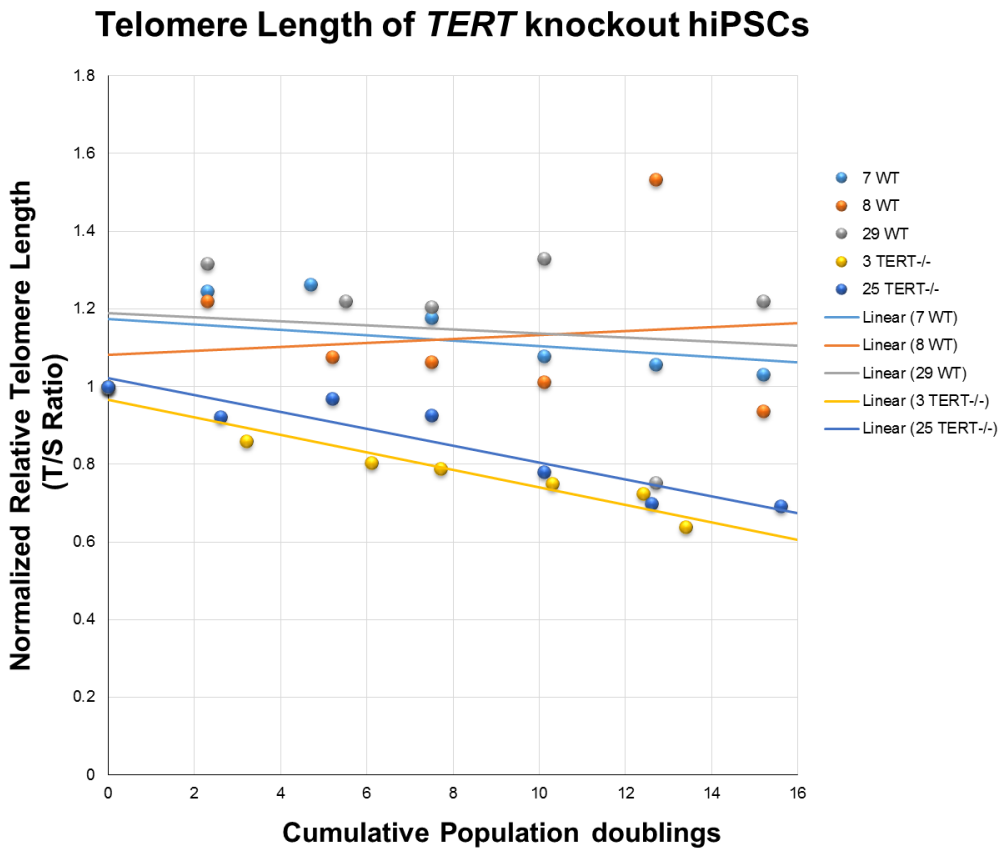


Figure 4.10: Relative telomere length during extended culture of *TERT* knockout cells. Relative telomere length (T/S ratio) between 34 and 64 days post-targeting in WT and *TERT*$^{-/-}$ mutant iPSCs. Telomere length remains constant in wild-type iPSCs, but decreases by approximately 30% over the 30 days of culture in mutant iPSCs ($P<0.001$, $\beta=-0.304$). Samples are normalized to T/S ratio at day 34 post-targeting.

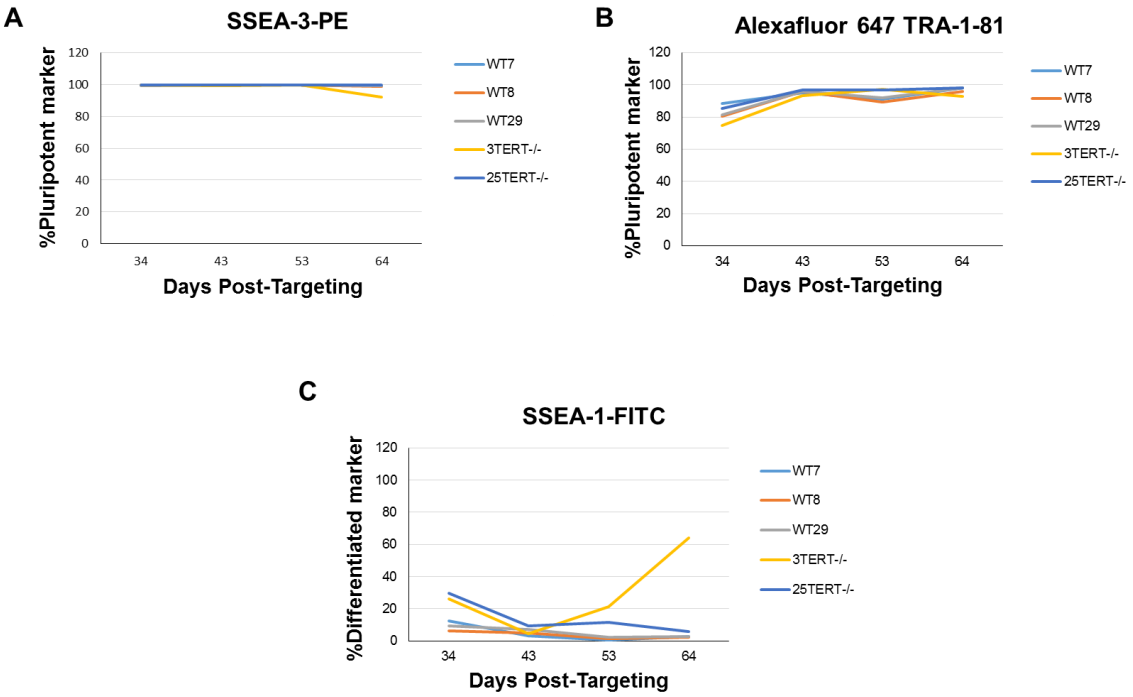


Figure 4.11: Analysis of *TERT*$^{-/-}$ iPSC surface markers. Flow cytometric analysis of SSEA-3 **(A)**, TRA-1-81 **(B)**, and SSEA-1 **(C)** expression of across 4 time-points between 34 and 64 days post-targeting. The expression levels of the markers SSEA-3, TRA-1-81 and SSEA-1 were detected using the antibodies of the surface markers coated with PE, Alexa Fluor® and FITC fluorescent dyes respectively.

4.3 Discussion

The main objective of this chapter was to use CRISPR/Cas9-mediate genome editing to knock out *TERT* in hiPSCs and study the effects on telomere length and iPSC phenotype. We used a validated sgRNA designed against exon one of *TERT* to target iPSCs and achieved a targeting efficiency of 31.4%. The majority of mutated alleles had 1bp insertions or deletions, about 37.1% in frameshift mutation. Analysis of telomere length revealed a decrease of ~30% for every 30 days in culture in homozygous *TERT* knockout iPSCs compared to wild-type controls. We also observed a non-significant reduction in telomere length in hiPSCs with heterozygous mutation in *TERT*, however, this prolonged culture of these lines would be needed to confirm the effects on telomere attrition.

As mentioned in the introduction of this chapter, there have been no previous studies of the effect of knockout of telomerase components in hiPSCs. Knockdown of *TERT* has been reported to be incompatible with ESC self-renewal, resulting in reduced proliferation, increased apoptosis and increased differentiation Yang *et al* (2008). It is not clear why our results, which involve complete knockout of *TERT*, result in a less severe phenotype, but could be due to differences in sensitivity between the ESCs in that study compared to the hiPSCs in our experiments or perhaps to differences in the initial telomere length in the parental cells. Telomerase deficient patient-derived iPSCs and mouse knockout iPSCs studies have also been investigated, however, these tend to investigate the effects of telomerase on telomere length during reprogramming rather than during prolonged culture. The attrition in telomere length in our hiPSCs is comparable with the data of Wang *et al*. (2012) who observed telomere shortening of 20-50% over 27 passages. Kinoshita *et al*. (2014) reported that *Tert* knockout miPSCs exhibited telomere length reduction by southern blotting, but the extent of shortening was not quantified.

Previous studies have reported that telomerase deficiency and telomere length can affect pluripotency. Huang *et al* (2011) and Wang *et al*. (2012) found that telomerase deficient miPSCs showed reduced ability to form chimeras and teratomas. Kinoshita *et al*. (2014) also examined the expression of pluripotency markers including Oct3/4, Sox2, Klf4, c-Myc, Nanog, ECAT and Eras and found that even though their expression was

maintained, and the cells could form embroid bodies *in vitro*, they showed reduced teratoma forming ability when injected *in vivo*, which they suggested was due to reduced proliferation capacity. We analysed the expression of the pluripotency markers SSEA-3 and TRA-1 as well as the differentiation marker SSEA-1 during prolonged culture. We saw no difference in expression of pluripotency markers suggesting that telomerase deficiency and telomere shortening did not alter hiPSC pluripotency during the 64 days of culture. However, expression of the differentiation marker was increased in the knockout lines, especially in *TERT*$^{-/-}$ clone 3, which showed a substantial increase in expression of SSEA-1 at later passages. At the same time the *TERT*$^{-/-}$ clone 3 was observed to have reduced growth rate compared to wild type lines and *TERT*$^{-/-}$ clone 25 at the end of the extended culture. These effects could be a result of the shorter telomere length in clone 3 (T/S of 1.02) compared to the other *TERT* mutant line (T/S of 1.45) and wild-type clones (T/S of 1.84-3.25). Batista *et al* (2011) reported a similar effect with prolonged culture of a *DKC1* deficient iPSC line and suggested that the critical shortening of telomere length led to loss of self-renewal was due to critical shortening of telomeres. We would need to confirm our observations with extended culture of the other *TERT*$^{-/-}$ lines and further investigation of the pluripotency should address their ability to form teratomas and chimeras at both early and late passages.

In conclusion, CRISPR/Cas9-mediated genome editing was successfully used to generate *TERT* knockout hiPSC lines. Telomere length was markedly decreased in *TERT* null iPSCs during extended culture. These results provide an outline for identifying the requirement of other genes in telomere length maintenance and the next chapter describes the use of the same strategy to investigate the effect of candidate telomere length associated genes on telomere maintenance.

Chapter 5

CRISPR/Cas9-mediated Knockout of Genes at the *ACYP2* locus

5.1 Introduction

5.1.1 GWAS to identify genes involved in the maintenance of telomere length

The result of the first GWAS for telomere length was performed in 2009 and identified two candidate regions associated with telomere length on chromosome 18q12.2 near the *VPS34/PIK3C* gene (Mangino et al 2009). The lead variants did not reach genome-wide significance, but the yeast orthologue of *VPS34* has been linked to telomere maintenance (Rog et al. 2005). In 2010, Codd *et al*. conducted a GWAS of mean leukocyte telomere length in approximately 12,000 individuals and identified a genome-wide significant signal near *TERC.* Later the same year, Levy *et al.* (2010), reported a locus on chromosome 10 near *OBFC1*, which encodes a component of the CST complex. In 2013, Bojesen and colleagues reported an association between SNPs at the *TERT* locus (Bojesen et al. 2013) and, in the largest study to date, Codd *et al.* (2013) reported seven loci following a meta-analysis and replication in a total of nearly 50,000 individuals (Codd et al. 2013). Five of the telomere length associated loci contain strong candidate genes that have known roles in telomere biology (*TERC*, *TERT*, *OBFC1*, *RTEL1*, and *NAF1*), however, the causal gene(s) at the other two loci (*ACYP2* and *ZNF208*) is unclear. In this chapter CRISPR/Cas9 genome, editing was used to target hiPSCs in an effort to identify the responsible gene at the *ACYP2* locus.

5.1.1.1 The *ACYP2* telomere length locus

The *ACYP2* telomere length locus is on chromosome 2p16 **(Figure 5.1)**. The effect allele of the lead variant, rs11125529, has an allele frequency of 0.88 and is associated with an age-related attrition in telomere length of 2.23 years per allele, which is equivalent to a per allele decrease in telomere length of 66.9 base pairs. There are 33 variants in high linkage disequilibrium ($r^2 > 0.8$) with the lead variant. These span a region of 36,567 base pairs within the gene body of *ACYP2* and overlap *TSPYL6*, which is transcribed from the antisense strand of *ACYP2* intron 3. All of the variants in high-linkage disequilibrium with the lead telomere length associated SNP are intronic except for two variants that fall in the 3' untranslated region of *TSPYL6*, one variant, rs6740641, that causes a

synonymous change at amino acid position 100 in *TSPYL6* and another variant, rs76397255, that causes a two amino acids deletion at amino acid position 194-195. Variants at the *ACYP2* locus have also been reported to associate with diseases including sudden cardiac death but require confirmation by other larger studies and stroke (Aouizerat et al. 2011, Liang et al. 2016), however, these are not in high linkage disequilibrium with the SNPs linked to telomere length.

ACYP2 and its homologue *ACYP1* encode the enzyme acyl phosphatase, which catalyses the hydrolysis of the high energy carboxyl-phosphate bond of acyl phosphates of membrane pumps (Modesti et al. 1993). Both enzymes are widely expressed in mammalian tissues, but are also called muscle acylphosphatase (ACYP2) and erythrocyte acylphosphatase (ACYP1) due to the predominant site of expression (Fiaschi et al 1995). It is not clear how this might link ACYP2 to telomere length, however, there have also been reports of ACYP2 migrating into the nucleus, where it might be involved in the hydrolysis of DNA (Stefani, Ramponi 1995, Chiarugi et al. 1997).

TSPYL6 encodes a member of the *TSPY/TSPYL/SET/NAP-1* (TTSN) superfamily (Vinci et al. 2009). There have been no studies into TSPYL6 protein function, however, the TSPYL6 protein does contain a nucleosome assembly protein (NAP) domain. Proteins that contain NAP domains, such as NAP and SET, directly bind nucleosomes and can function as regulators of gene expression and cell division (Rodriguez et al., 1997, Carlson et al., 1998). Paralogues of TSPYL6 include TSPYL2, which is involved in gene regulation and cell-cycle regulation following DNA damage (Tao et al. 2011, Tsang et al. 2014, Epping et al. 2015) and TSPYL5, which supresses levels of P53 (Epping et al. 2011).

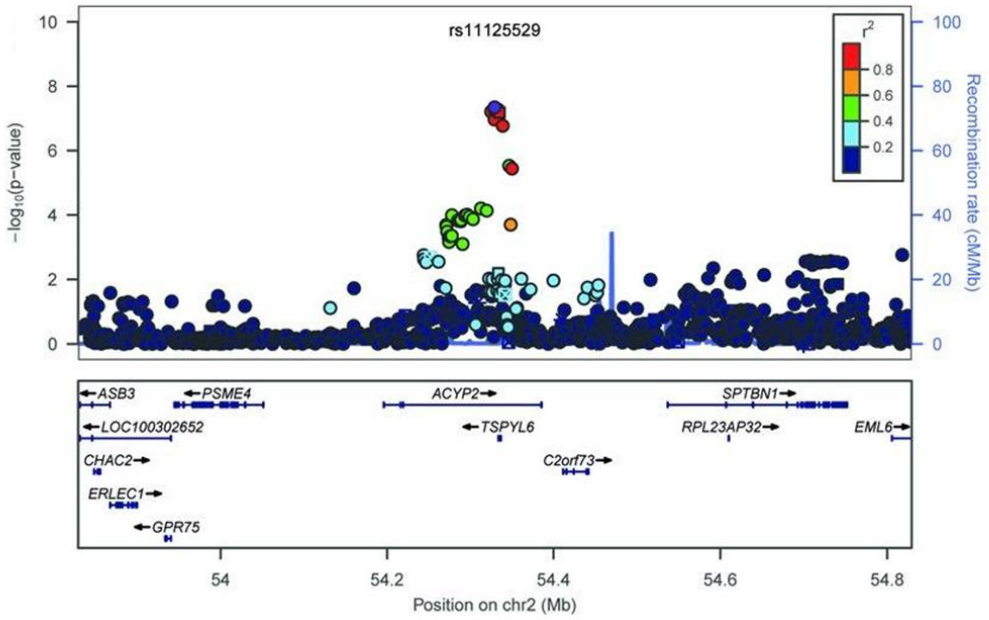


Figure 5.1: Regional association plots for the *ACYP2* locus. The log10 P value is plotted against base pair position for each SNP. The lead SNP is shown in purple with the other SNPs coloured according to linkage disequilibrium (r2). Blue beak indicates the recombination rate (HapMap2) and RefSeq genes present in the locus are shown at the bottom. This figure is taken from Codd *et al*. (2013).

5.1.2 Aims and objectives

The objective of the work presented in this chapter was to use CRISPR/Cas9 gene targeting to determine whether a gene at the *ACYP2* locus is involved in the maintenance of telomere length.

The aims are;

- To generate *ACYP2* and *TSPYL6* hiPSC knockout cells using CRISPR/Cas9
- Investigate the effects of gene knockout on telomere length

5.2 Results

5.2.1 Expression of telomere length associated genes in iPSCs

We first investigated whether *ACYP2* and *TSPYL6* are expressed in iPSCs using RT-PCR. Absence of expression would mean altering the downstream screening to include differentiation of the iPSCs into a cell-type that does express the gene of interest. Optimised primers were used to amplify *ACYP2* and *TSPYL6* in cDNA prepared from iPSCs along with a no template control (NTC) and reverse transcriptase negative (NRT) controls. Both genes were found to be expressed in iPSCs **(Figure 5.2)**.

5.2.2 Design and validation of sgRNAs targeting genes at the *ACYP2* locus

As described in the previous chapter, sgRNAs for *ACYP2* and *TSPYL6* were designed by Horizon Discovery Ltd (Cambridge, UK) to target each gene early in its coding sequence, thereby making it more likely that any frameshift mutation would remove protein function. Horizon Discovery Ltd designed two *ACYP2* sgRNAs and six *TSPYL6* sgRNAs and cloned these into pGE12-001. The sgRNAs were then transfected into HEK293 cells alongside the S. *pyogenes* Cas9 -PL1 plasmid. Cutting efficiency was measured using Surveyor assay and the most efficient sgRNA (**Figure 5.3**) for each gene was supplied by Horizon. The most efficient sgRNA for *ACYP2,* 5'-GTCCGAACACCTCGTAGTCCA-3', had an estimated cutting efficiency of approximately 72%, no predicted off-target effects within a gene and two predicted off-target sites in non-genic regions. For *TSPYL6,* 5'-GCGGGGCGGTGGGAACACGAT-3' had an estimated cutting efficiency of 19% and eight potential off-target sites, two of which are in genes. **Figure 5.4** shows the position of the sgRNAs in each gene.

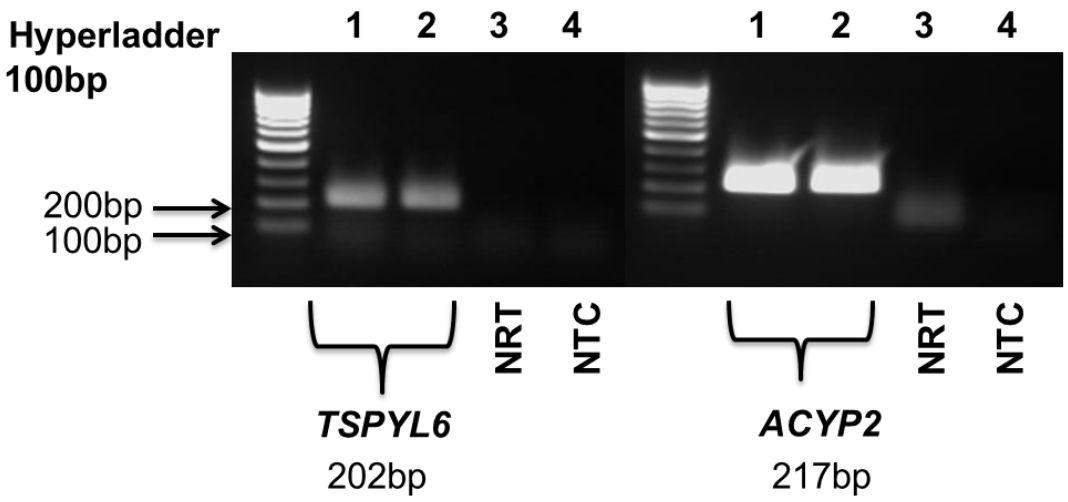


Figure 5.2: RT-PCR shows the expression of *TSPYL6* and *ACYP2* in GM23720*B iPSCs. Lane 1 and 2, no reverse transcriptase (NRT, lane 3) and no template control (NTC, lane 4).

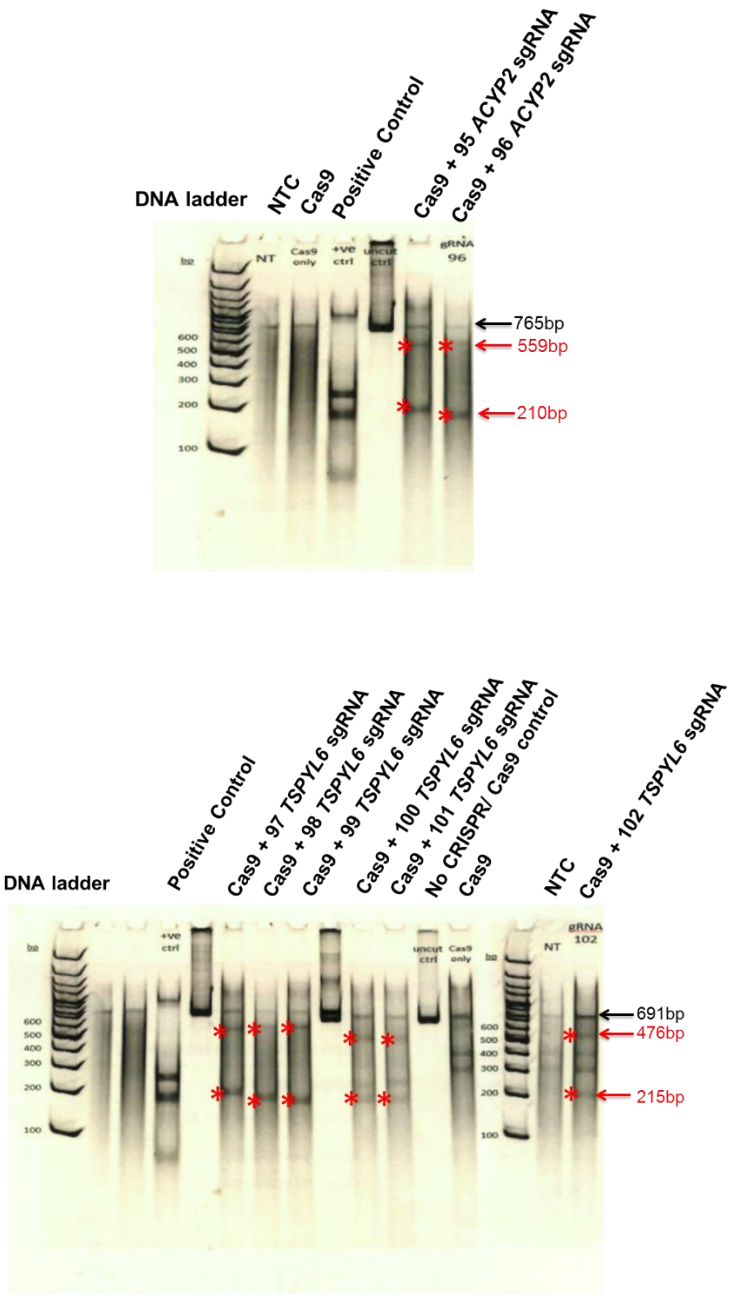


Figure 5.3: Surveyor screening assay for *ACYP2* and *TSPYL6* sgRNA selection. Horizon Discovery Ltd performed a Surveyor assay to select the most efficient sgRNA for each gene. Each candidate sgRNA was co-transfected into HEK293 alongside Cas9. Controls comprise a Cas9 only transfection, a no template control (NTC), No CRISPR/ Cas9 control and a mismatch positive control. PCR products of the target genomic region (or positive control) were subjected to a Surveyor assay and resolved on an acrylamide gel. Digestion products (red asterisks) of the expected size were detected for each sgRNA. Densitometry was performed to estimate cutting efficiency.

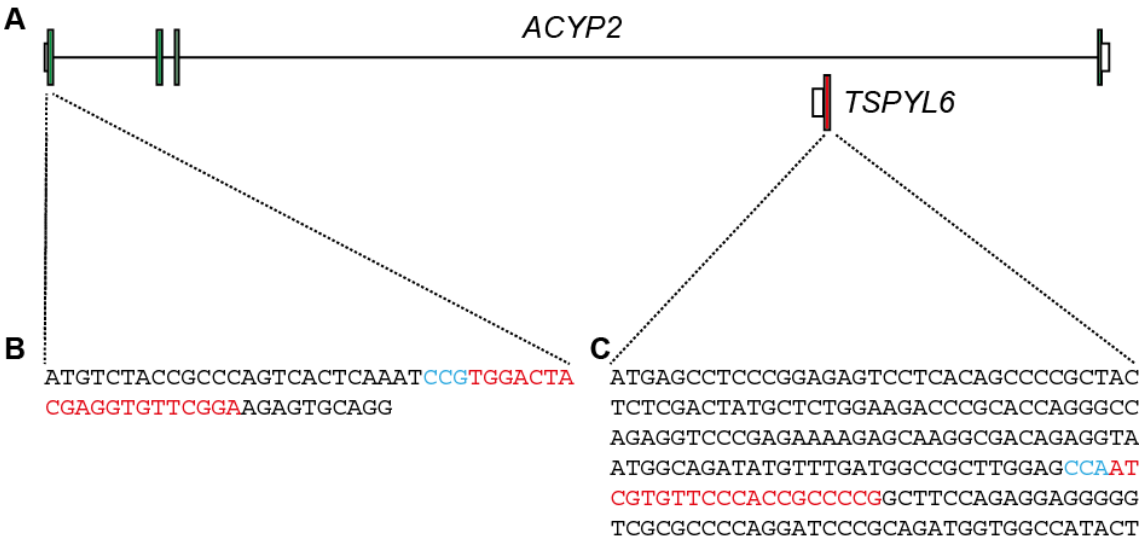


Figure 5.4: Protein encoding transcripts and exon sequencing of *ACYP2* and *TSPYL6*. (**A**) Schematic illustration of the *ACYP2* and *TSPYL6* transcripts; *ACYP2* exons are shown as green bars separated by a black line representing introns. The single *TSPYL6* exon is shown as a red bar. (**B)** Shows the first 58 coding nucleotides of exon one of *ACYP2*. The sgRNA target sequence is shown in red with the PAM in blue. **(C)** Shows the first 210 coding nucleotides of *TSPYL6.* The sgRNA target sequence is in red and the PAM is in blue.

5.2.2.1 CRISPR/Cas9 knockout of genes at the *ACYP2* locus

iPSCs were transfected with Cas9 PL1-plasmid on its own or with the *ACYP2* or *TSPYL6* sgRNA using the NEPA21 electroporator (Nepagene) and program seven as described in **section 2.1.3.3.1**. After transfection, the targeted cells were allowed to recover and were cultured normally. When the cells reached confluence, they were passaged and a portion of cells collected for DNA extraction and screening. When the cells next reached confluence they were banked in liquid nitrogen.

5.2.2.2 Screening of CRISPR/Cas9 targeted iPSCs using the Surveyor assay

Gradient PCR was used to optimise the PCR amplification of the region of genomic DNA surrounding the target site of each sgRNA **(Figure 5.5)** followed by amplification of each region in targeted and control cells. The primer sequences are shown in **section 2.1.4.1** and the optimised PCR conditions in **section 2.1.4.12**.

The Transgenomic Surveyor® mutation detection kit was then used to detect if targeting had been successful. The expected PCR products of 765bp and 691bp were seen for *ACYP2* and *TSPYL6* respectively **(Figure 5.6)**. The presence of mismatches should cause the Surveyor nuclease to digest the PCR product and produce fragments of 210bp and 559bp for *ACYP2*, and 215bp and 476bp for *TSPYL6*. In the experimental samples bands of the expected size were clearly seen suggesting successful targeting of both genes **(Figure 5.6)**.

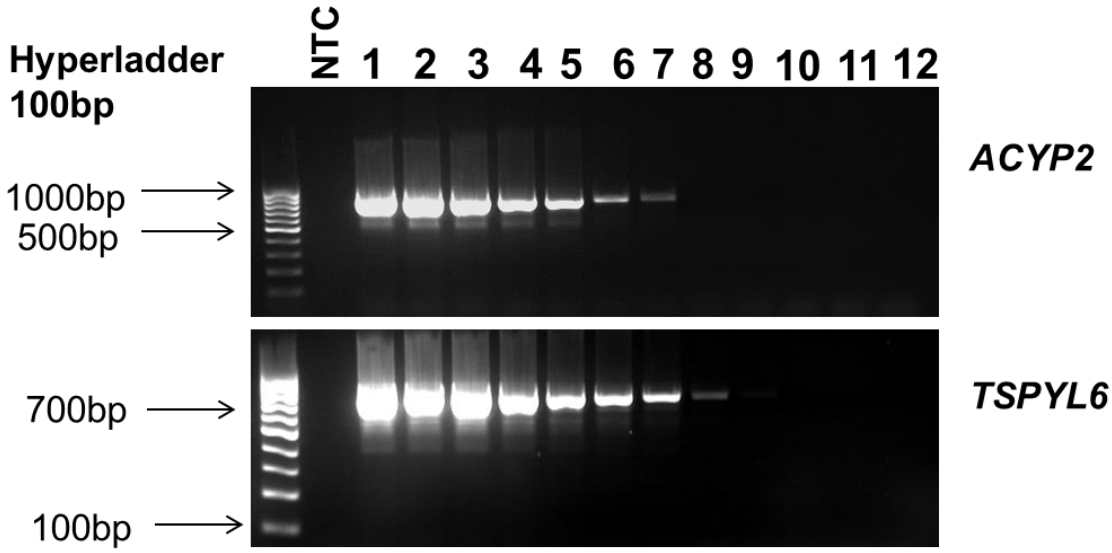


Figure 5.5: Gradient PCR reactions of annealing temperature for *ACYP2* and *TSPYL6*. shows gradient annealing temperature for *ACYP2* and *TSPYL6* genes visualized by gel electrophoresis on 1.5% agarose. 100bp ladder as marker for DNA bands and no template control (NTC) was used. Gradient annealing temperature range lane1; 63.1°C, lane2; 63.3°C, lane3; 63.6°C, lane4; 64.1°C, lane5; 65.1°C, lane6; 66.0°C, lane7; 66.9°C, lane8; 67.9°C, lane9; 68.9°C, lane10; 69.5°C, lane11; 69.9°C and lane12; 70.0°C. The optimal annealing temperature for both genes was 65.1°C.

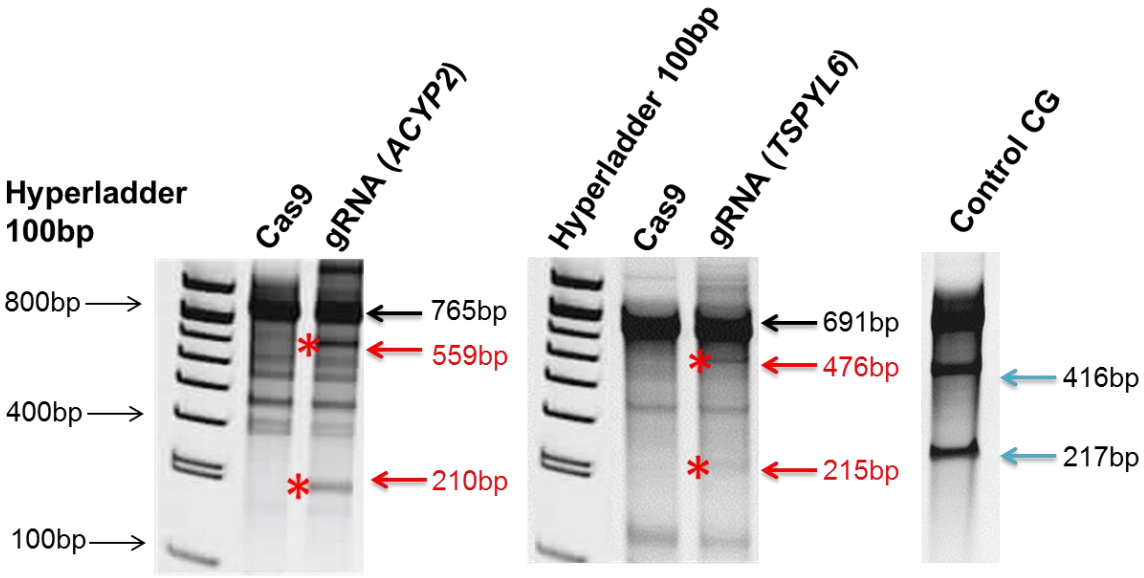


Figure 5.6: Surveyor nuclease detection of products of hybridized amplicons of *ACYP2* and *TSPYL6*. Shows the acrylamide gel images to allow visualization of both the upper and lower bands Products of the expected size are indicated by red asterisks, no product of the expected size is seen in Cas9 only targeted cells. Blue arrows indicate to the digested hybridized PCR of the mismatch positive control CG.

5.2.2.3 Clonal Derivation and Sequence Confirmation of CRISPR/Cas9 targeted hiPSCs

As described in **section 3.2.4**, the transfected cells were seeded at 10^3 cells/ml on a 90mm petri dish. When the colonies reached approximately 2.5mm in diameter (around 7-10 days), they were picked and transferred to a Matrigel™-coated 96-well plate. When the colonies reached confluency, they were passaged into fresh plates and frozen at -80°C. A portion of cells was retained for DNA extraction and Sanger sequencing. As summarised in **Table 5.1**, 28 of 36 and 30 of 36 colonies of *ACYP2* and *TSPYL6* respectively survived transfer to 96-well plates and were sequenced to identify potential mutant clones. Sequence traces were analysed by the TIDE sequence deconvolution tool https://tide-calculator.nki.nl/ (Brinkman et al. 2014) which suggested that 9 *ACYP2* clonal lines and 10 *TSPYL6* clones were homozygous for both wild-type alleles, 7 *ACYP2* clones and 19 *TSPYL6* clones contained 1 wild-type allele and 12 *ACYP2* clones and 1 *TSPYL6* clone had indels in both alleles. Therefore, targeting efficiency (proportion of targeted alleles) was 55.3% for *ACYP2* and 35% for *TSPYL6*. The genotype predicted by TIDE analysis for each clonal line is shown in **supplementary table 2** and **3**. Unfortunately, TIDE sequence analysis suggested that the single *TSPYL6* homozygous targeted clone (clone 13) contained an in-frame 6 base pair deletion in one allele while all of the *ACYP2* homozygous lines contained in-frame deletions of one or both alleles. It is not known whether the in-frame deletions are deleterious. As heterozygous loss of telomerase impacts telomere length maintenance we decided to continue our experimentation using these lines. Wild-type control and homozygous targeted clones were expanded to 24-well plates and then to 6-well plates. DNA was then extracted from a subset of clones (*TSPYL6* clone 13, *ACYP2* clones 4,9 and 18) for further sequence analysis and initial telomere length measurements, and hiPSCs archived in liquid nitrogen. The region of genomic DNA surrounding the CRISPR/Cas9 targeted mutation was amplified by PCR and cloned into pGEM®-T Easy vector and at least four colonies sequenced for each clone **(section 2.1.4.14)**. For *ACYP2*, one allele of clone 4 had a 1bp adenine (A) insertion at position c.32 in the coding sequence, which would cause a frameshift and premature stop at amino acid position 30 **(Figure 5.7, 5.8 and 5.9)**. The other allele caused a 9bp deletion between c.33-41, resulting in a 3 amino acid deletion. Clone 9 also contained a 1bp insertion A in one allele, while the second allele consists of a 7bp deletion and 16bp insertion resulting

in a change in the protein sequence and 3 amino acid insertion **(Figure 5.7, 5.8 and 5.9)**. Again, one allele of clone 18 contained a 1bp A insertion at position 32, while the second allele consisted of a 9bp deletion between positions 32-40 and 3 amino acid deletion. **(Figure 5.7, 5.8 and 5.9)**. One allele of *TSPYL6* clone 13 consisted of a 6bp deletion at c.1087-1091 causing a two amino acid deletion **(Figures 5.10, 5.11, A and B).** The other allele was an apparent a 1bp deletion cytosine (C) at c.998 and 18bp deletion at c.1080-1097, resulting in a frameshift and premature stop at amino acid positon 84. It is not clear how this complex mutation was generated. It is possibly a result of a 99pb deletions coupled with an 80pb insertion **(Figures 5.10, 5.11, A and B)**.

Description	Number of clones of *ACYP2*	Number of clones of *TSPYL6*
Transferring into 96-well plate from Petri-dish	36	36
Growing colonies in 96-well plate	28	30
After sequencing • Wild type • Homozygous • Heterozygous	 9 12 7	 10 1 19
After reviving and culturing in 24-well plate • Wild type • Homozygous • Heterozygous	 8 11 Not done	 8 1 Not done
After passaging and transferring into 6-well plate • Wild type • Homozygous • Heterozygous	 8 11 Not done	 8 1 Not done

Table 5.1: Number and type of mutation in *ACYP2* and *TSPYL6* in each targeted iPSC line.
Number and predicted genotype of clonal iPSCs during expansion.

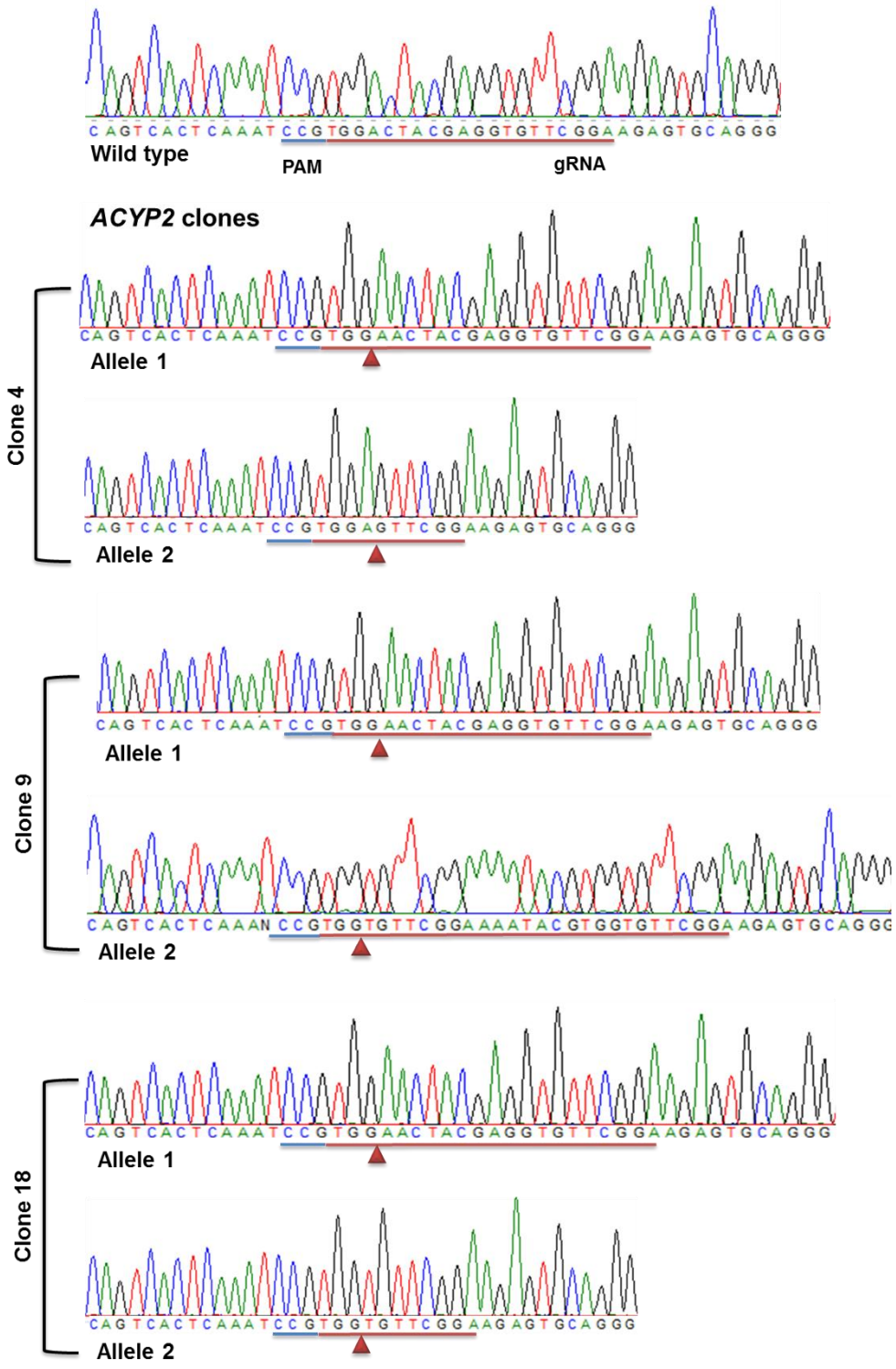


Figure 5.7: Sequence Trace Data for *ACYP2* knockout clones 4, 9, and 18. Sequence traces of wild-type and each allele of the *ACYP2* homozygote targeted clones. The blue bar indicates to PAM sequence, the red bar represents the sequence targeted by the sgRNA and the red triangle the specific site of cutting by CRISPR/Cas9.

Figure 5.8: Genomic DNA sequence of *ACYP2* homozygote targeted clones. Genomic DNA sequence of wild-type and *ACYP2* mutant clones. Insertions are shown in red text and deletions represented by dashes.

Figure 5.9: Protein sequence of *ACYP2* homozygote targeted clones. Predicted protein sequence of wild-type and mutant *ACYP2*. Changes in the protein sequence are highlighted in red.

5.2.2.4 Measurement of Telomere Length in *ACYP2* and *TSPYL6* deficient cells

Following expansion, the DNA of wild-type controls and *ACYP2* and *TSPYL6* mutant clones at 47 and 48 days post-targeting respectively and telomere length was measured by qRT-PCR (**Figure 5.12**). There was no difference in telomere length of *ACYP2* knockout cells compared to wild-type clones (mean 1.616 ± SD 0.04783 versus mean 1.865 ± SD 0.3456, P= 0.5028, student's t-test). Telomere length in *TSPYL6* clone 13 was longer than the wild-type controls, but was within the range seen with the wild-type clones in chapter 4 and is probably a result of the normal variation between individual wild-type cells.

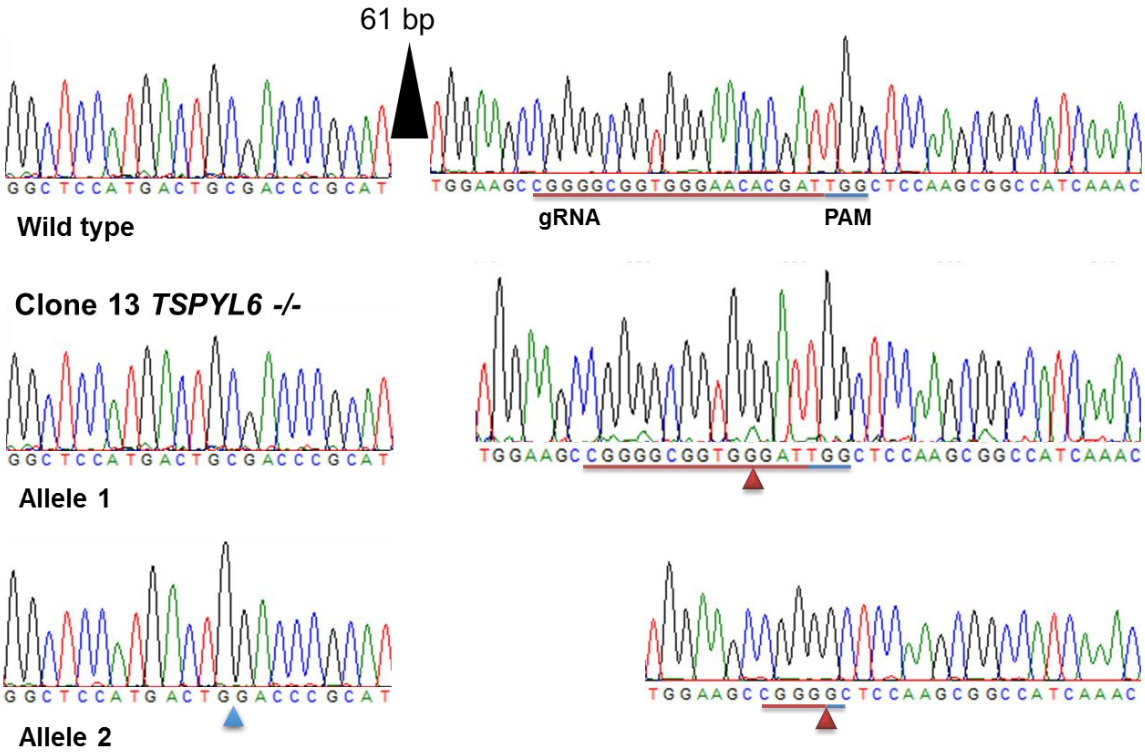


Figure 5.10: Sequence Trace for *TSPYL6* clone 13. Sequence traces of wild-type and each allele of *TSPYL6* knockout clone 13. The blue bar indicates the PAM sequence, the red bar represents the target sequence of the gRNA and the red triangle shows the cut-site of the CRISPR/Cas9 while blue triangle shows a cut out gRNA.

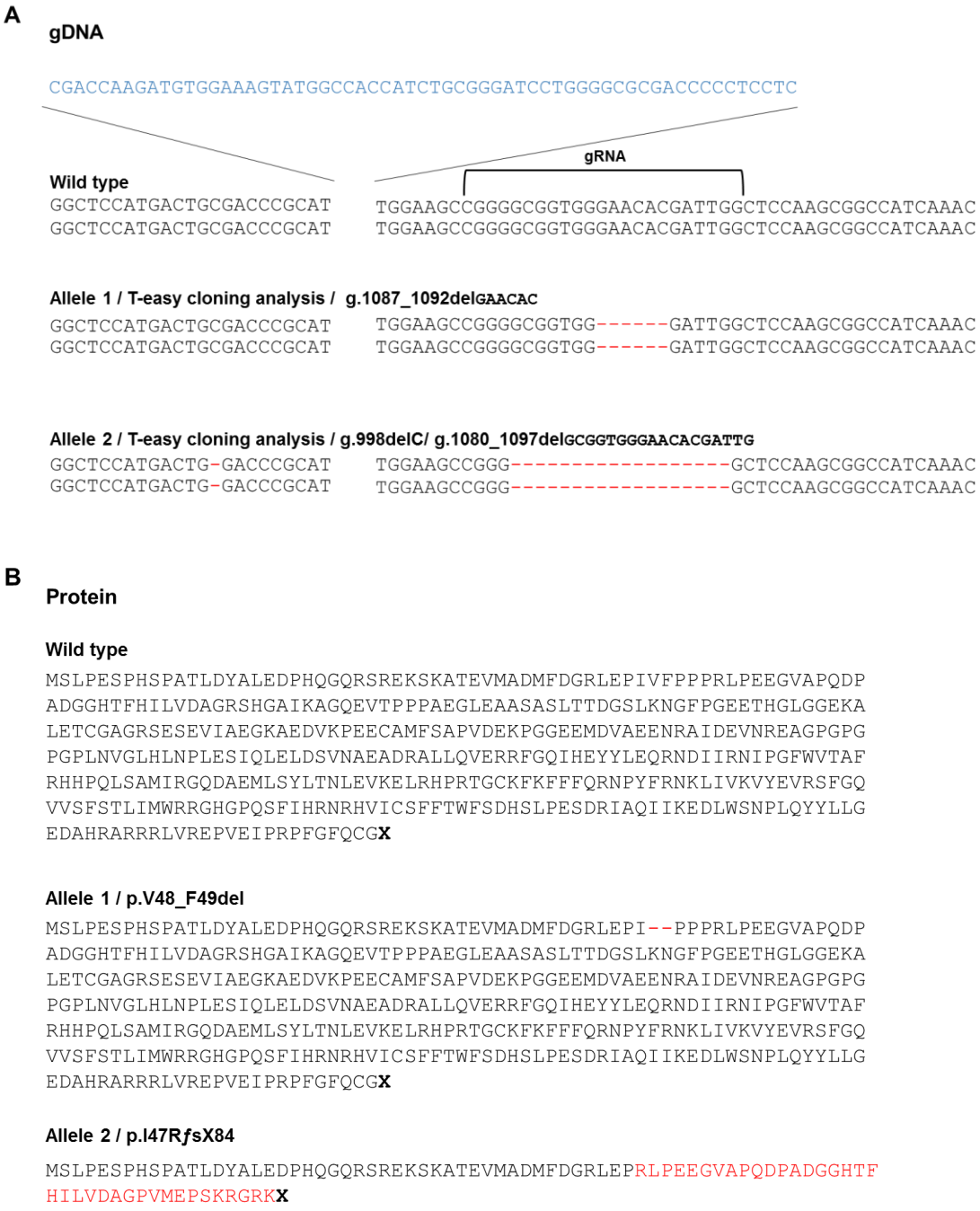


Figure 5.11: Genomic DNA and Protein sequence of TSPYL6 homozygote targeted clone 13. (A) Genomic DNA sequence showing the 6bp at coding position 1087-1092 in allele 1 of *TSPYL6* clone 13 and 1bp at c.998 and 18bp deletion at c.1080-1097 in allele 2. Deletions are shown with red dashes. The blue sequence shows c.1008-1079. **(B)** Protein sequence of wild-type and TSPYL6 mutant proteins showing the two amino acid deletion in allele1 (red dash) and the frameshift (highlighted in red) caused by the two deletions in allele 2.

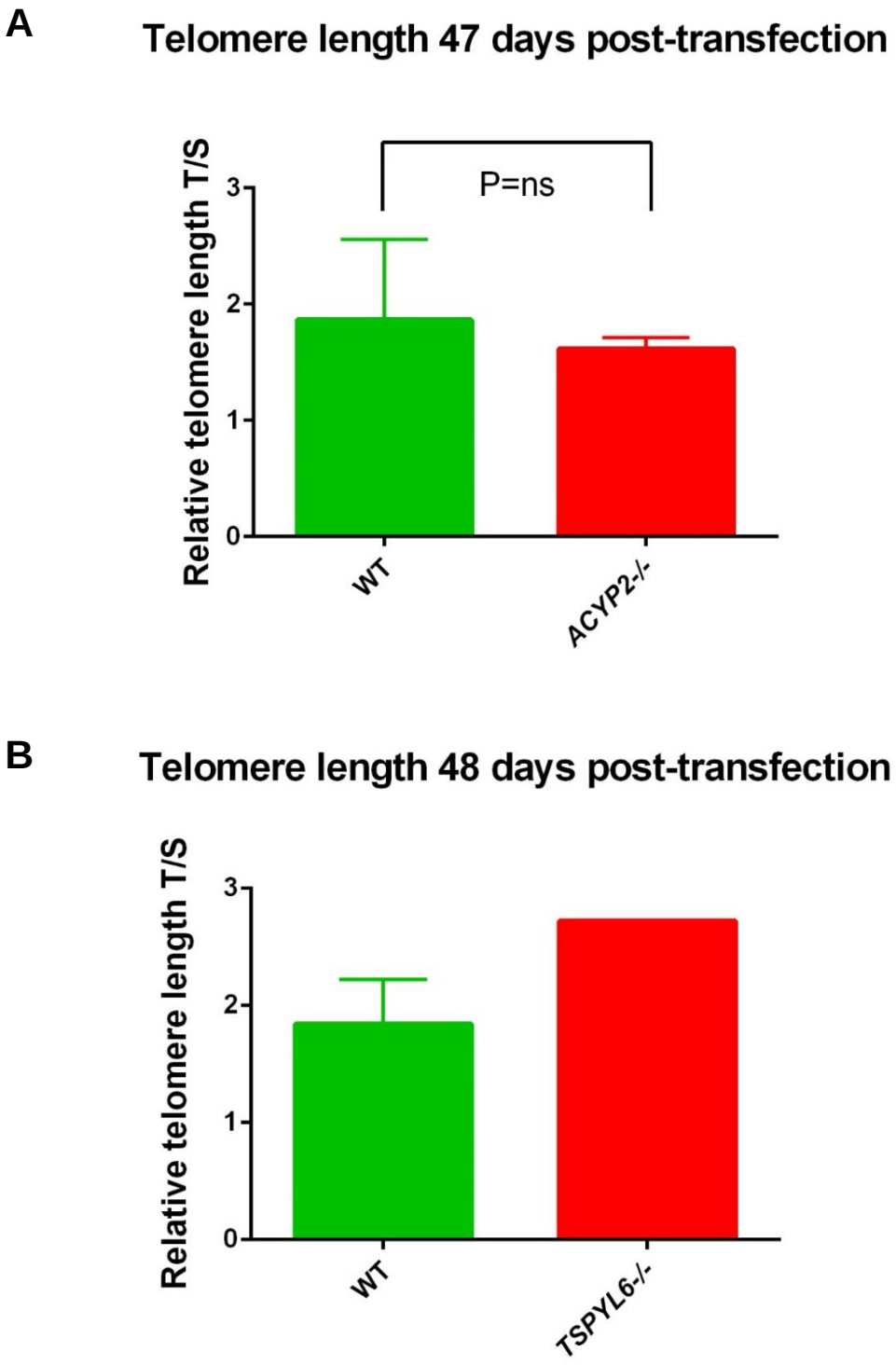


Figure 5.12: Relative telomere length of *ACYP2* and *TSPYL6* homozygote targeted cell lines. (A) Telomere length of *ACYP2* homozygote targeted clones at day 47 post-targeting (n=4 in each group). **(B)** Telomere length of the *TSPYL6* homozygote targeted clone at day 48 post targeting (n of WT=3, n of *TSPL6*$^{-/-}$=1). T/S indicates telomere length divided by single-copy gene 36B4. WT represents wild-type cell.

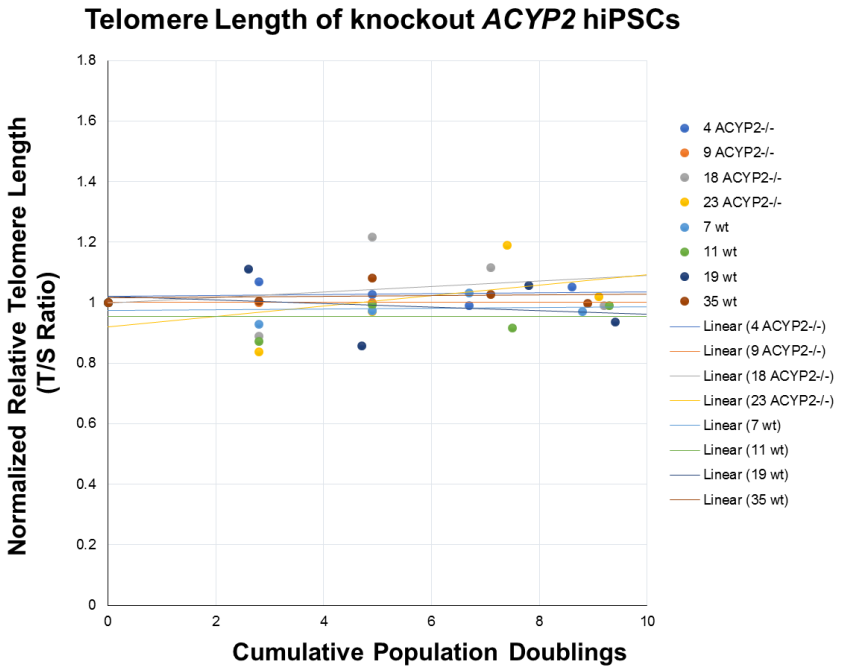
5.2.2.5 Extended culture of *ACYP2* and *TSPYL6* homozygote targeted cell lines

Wild-type and mutant hiPSCs were maintained as described in **section 2.1.1.2.4** until they had been in culture for approximately 60 days post CRISPR/Cas9 targeting. The total time of culture was equivalent to about 9 population doublings. At each passage a portion of the cells from each passage was archived in liquid nitrogen and a portion of cells retained for DNA and RNA extraction. The telomere length of the *ACYP2* and *TSPYL6* homozygote cell lines was measured by qPCR across several time-points. We assessed the differences between mutant and control iPSCs by applying a mixed model with sample as a random effect. No change in telomere length was seen for either the *ACYP2* (P=0.7) or *TSPYL6* mutant cell lines in comparison to wild-type controls **(Figure 5.13, A and B**).

5.2.2.6 Characterization of *ACYP2* and *TSPYL6* homozygote targeted hiPSCs

The *ACYP2* and *TSPYL6* mutant hiPSCs were characterized for pluripotency during extended culture by measuring the expression of cell surface markers using flow cytometry. As previously described **(Sections 2.1.2.1 and 3.2.2.1)**, approximately $1x10^7$ cells/ml were fixed in 4% paraformaldehyde (PFA) in 20 min at room temperature at several time-points during the extended hiPSC culture and stained for SSEA-3, TRA-1-81 and SSEA-1. There was no observable difference in marker expression between the mutant cell lines and the wild-type controls **(Figure 5.14, A, B)**.

A



B

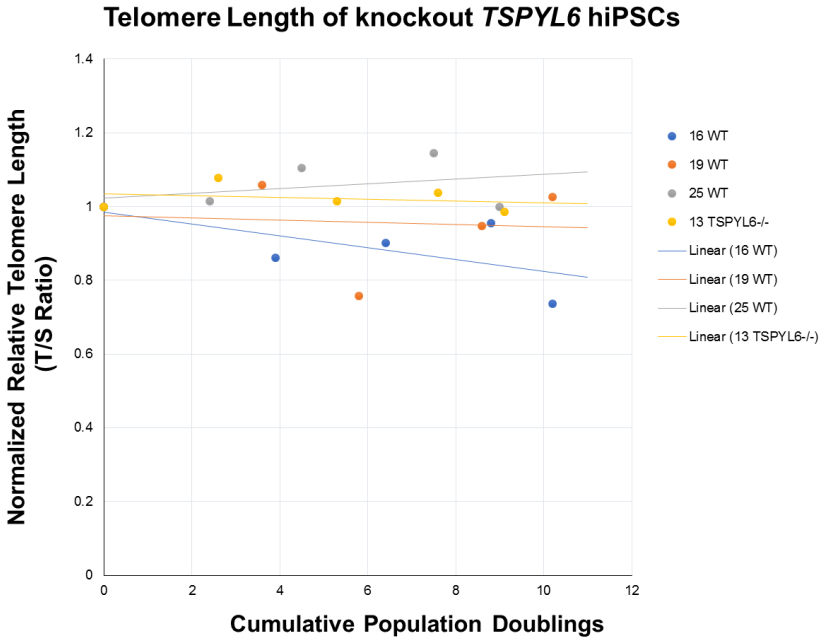


Figure 5.13: Relative telomere length of *ACYP2* and *TSPYL6* wild type and knockout cell lines. **(A)** Normalized relative telomere length over 63 days post CRISPR/Cas9 targeting of *ACYP2*. **(B)** Normalized telomere length over 64 post CRISPR/Cas9 targeting of *TSPYL6*. T/S indicates to telomere length divided by single copy gene 36B4. WT represents wild-type cells.

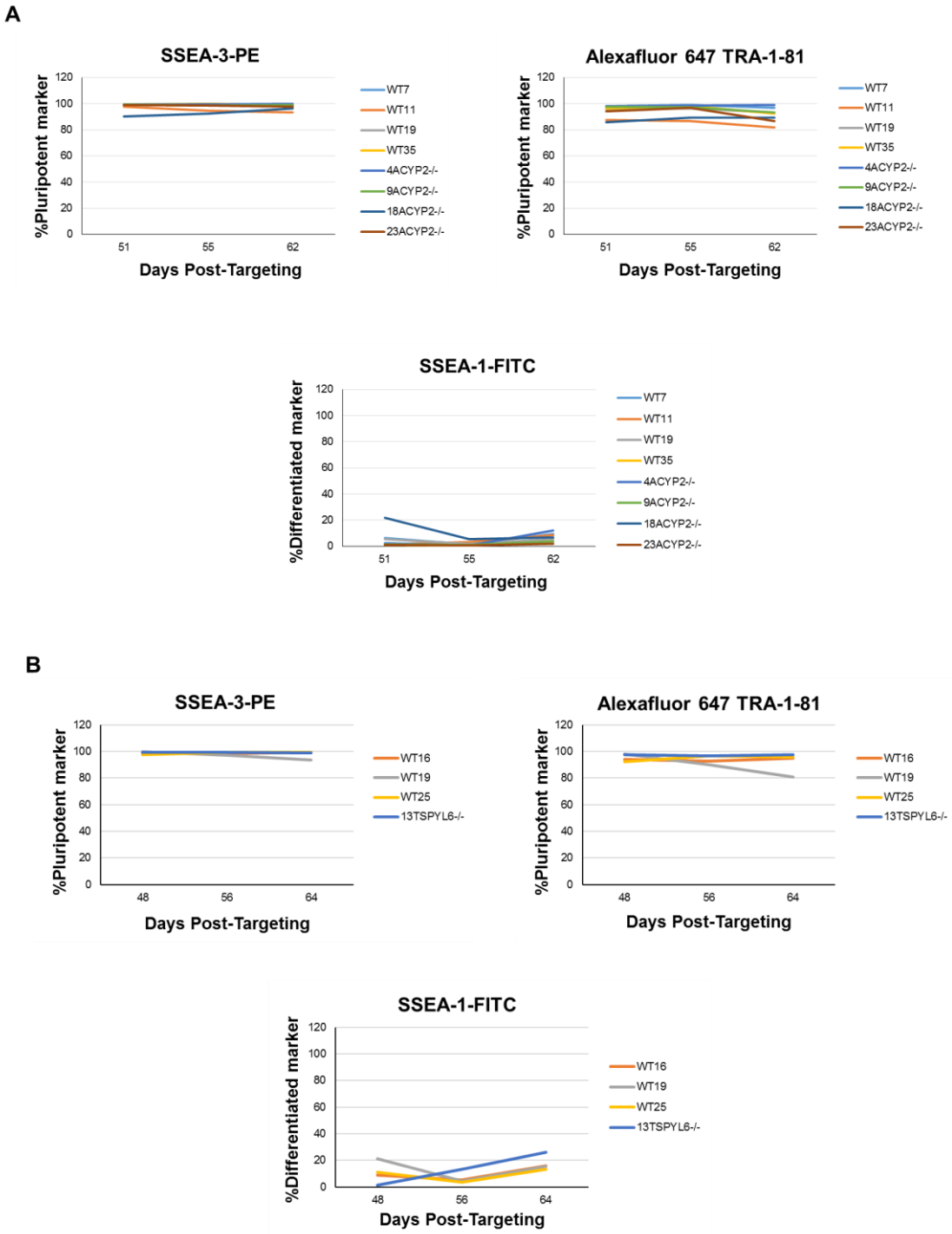


Figure 5.14: Characterization of knockout iPSCs during extended culture by flow cytometry. Flow cytometric analysis for *ACYP2* **(A)** and *TSPYL6* **(B)**. The pluripotent surface marker expression levels as detected using SSEA-3 antibody coated with PE fluorescent dye and TRA-1-81 antibody coated with Alexa Fluor® fluorescent dye and differentiation marker SSEA-1 antibody coated with FITC fluorescent dye. The green lines indicate wild-type clones and the blue lines indicate mutant cell lines.

5.3 Discussion

A major post-GWAS challenge involves the identification of the gene that mediates the effects of the disease or trait associated variants. Over the past few years GWAS have identified the genetic loci that associate with mean telomere length. In the largest of these, Codd and colleagues identified seven loci impacting telomere length, of which five contain strong candidate genes because of known roles in telomere biology, whereas there is no obvious candidate at the other two loci (Codd et al. 2013). In this chapter, the influence of two genes at the *ACYP2* telomere length associated locus, *ACYP2* and *TSPYL6,* on telomere length was investigated using CRISPR/Cas9 mediated gene knockout. Despite relatively efficient targeting of both genes (55.3% and 35% alleles mutated for *ACYP2* and *TSPYL6* respectively) definite homozygous knockout hiPSCs were not generated. For *ACYP2* 8 lines contained indels for both alleles, however, all contained at least one in-frame deletion. For *TSPYL6* the single homozygous line also contained an in-frame deletion for one allele. The effect of the mutations on telomere length was investigated in the same way as described in chapter 4; however, no change in telomere lengths was noted for either the *TSPYL6* or the *ACYP2* homozygote targeted cell lines.

There are several reasons why mutation of *ACYP2* and *TSPYL6* had no effect on telomere length. Firstly, the effects might be too subtle to detect over the time-frame investigated and might only become apparent with even longer culture or with homozygous knockout. Secondly, hiPSCs might not be a suitable cell line to investigate the requirement of these genes on telomere length. Although telomerase is required to maintain telomere length in hiPSCs, other mechanisms might only be active in certain cell-types or detectable in the absence of telomerase expression. This could be investigated by performing similar experiments in differentiated hiPSCs or generating *ACYP2* and *TSPYL6* mutants on a telomerase null background. Finally, it is possible that another more distal gene is affected by the telomere length associated variants, however, during the course of this investigation the results of two other studies suggest *TSPYL6* is a strong candidate for the causal gene.

At the start of this study there was no evidence to support functional effects of either the lead telomere length associated variant, or any of the proxy variants in high linkage

disequilibrium, on *ACYP2* or *TSPYL6*. Recently, the Genotype-Tissue Expression (GTEx) project, which aims to link genetic variants to changes in gene expression (Lonsdale et al. 2013) identified an association between rs62139254, which is in perfect linkage disequilibrium with the lead telomere length associated variant and expression of *TSPYL6* in testis. Interestingly, the effect allele is associated with higher levels of *TSPYL6* expression, suggesting that increased levels of TSPYL6 would cause telomere shortening. GTEx also reports much higher *TSPYL6* expression in testis than in any other tissue, which implies that any effect *TSPYL6* has on telomere length will be a result of its expression in this tissue. Interestingly, telomere length is longer in the germ cells of older men and this is associated with longer mean leukocyte telomere length in their offspring (Kimura et al. 2008). In addition, a high-throughput proteomic screen identified TSPYL6 as an interactor of MAD2L2 (MAD2 mitotic arrest deficient-like 2 (yeast) (Huttlin et al 2015), which has recently been reported to regulate DNA damage at telomeres (Boersma et al. 2015). Functional studies of TSPYL6 should be performed to confirm these findings and to determine its role in telomere length maintenance.

It is unclear why definite homozygous knockout lines were not produced, especially as the overall targeting efficiency appeared promising. There is a possibility that homozygous gene loss is detrimental and led to cell death. Alternatively, the tendency towards in-frame deletions might be a consequence of the DNA sequence in the vicinity of the target sites of the sgRNAs. Further experimentation is required to determine the underlying cause and this should include targeting with additional sgRNAs or different design strategies to exclude local sequence-specific effects.

In conclusion, CRISPR/Cas9 mediated genome editing was used to generate hiPSCs deficient for the candidate telomere length associated genes *ACYP2* and *TSPYL6*. No effect on telomere length was found following mutation of either gene. Future experiments should include other functional assays and a detailed analysis of protein function.

Chapter 6

Discussion

6.1 Discussion

In 2013, the largest GWAS for telomere length reported to date identified seven loci associated with decreased telomere length (Codd et al. 2013). Five of the loci identified contain strong candidate genes with known roles in telomere length maintenance (*TERC, TERT, OBFC1, RTEL1,* and *NAF1*), but the mechanism underlying the association is not known. The other two loci, *ACYP2* and *ZNF208*, do not contain genes with a known role in telomere biology and this is the first study that has attempted to identify the functional gene. The overall objective of this study was to utilise the CRISPR/Cas9 genome editing system to determine the causal gene at one of these loci. CRISPR/Cas9 was first reported shortly before this project began. Since the start of this study, there have been more than 200 studies involving CRISPR/Cas9. Many of these also involved the use of iPSCs and include correction of mutations, for example and disease modelling (Xie et al. 2014, Bassuk et al. 2016, Freiermuth et al. 2017, Bassett et al. 2017). For example, the human haemoglobin beta (*HBB*) gene which causes B-thalassemia was corrected in patient derived iPSCs using (Xie et al. 2014) and an RPGR point mutation, which causes X-linked retinitis pigmentosa, was converted to a wild-type allele in hiPSCs (Bassuk et al. 2016). CRISPR/Cas9 has also been used to investigate GWAS loci (Pashos et al. 2017, Gupta et al. 2017). Pashos *et al*. (2017) used CRISPR/Cas9 to investigate SNPs associated with blood lipids, while Gupta *et al*. (2017) targeted a variant associated with coronary artery disease and to demonstrate that the causal gene was more than 500kb from a disease-asscoaited locus. For this study, CRISPR/Cas9 mediated genome editing was established and optimised **(chapter 3)** and used to target hiPSCs, which maintain their telomeres, and knockout genes related to telomere length. First, CRISPR/Cas9 was used to generate telomerase deficient hiPSCs by knocking out *TERT*, which encodes the enzymatic component of telomerase. Targeting efficiency was relatively high with mutation detected in 31.4% of alleles. Analysis of these cells revealed telomeres shortened by approximately 30% over 1 month of culture. The hiPSCs retained expression of cell-surface pluripotency markers during this time, but were observed to have increased expression of differentiation markers at the end of the experiment, possibly a result of critical telomere shortening. Next, the same approach was used for the investigation of two genes at the *ACYP2* locus; *TSPYL6* and *ACYP2*. Despite achieving targeting efficiencies of 35% and 55.3% no definite knockout cell line was

produced for either gene. A large number of mutant *ACYP2* lines was generated, however, a substantial number of the targeted alleles comprised in-frame deletions, which might not have a damaging effect on protein function. Only one homozygous mutant *TSPYL6* line was created and although the reason for this is unclear, it could potentially be a result of *TSPYL6* being detrimental to hiPSCs. As no difference in telomere length was detected for either the *ACYP2* or *TSPYL6* targeted lines it is not possible to conclude which gene is responsible for mediating the genetic association with telomere length. However, during the course of this research, additional evidence has come to light that identifies *TSPYL6* as a strong candidate gene at this locus and further investigation is warranted.

Around 10 years ago, the first studies to report the creation of hiPSCs and miPSCs from embryonic and adult somatic cells were reported (Takahashi, Yamanaka 2006, Takahashi et al. 2007). The cells are reprogrammed by ectopic expression of pluripotency genes (*Oct4*, *Sox2*, *Klf4* and *c-Myc*) and gain the properties of ESCs, including the capacity for long-term proliferation through expression of telomerase, which is silent in most somatic cells. The reactivation of telomerase activity causes elongation of telomeres during reprogramming and telomerase activity and telomere length can therefore be considered a stem cell marker (Marion et al. 2009). In this study, the CRISPR/Cas9 genome editing system was used to target *TERT* in hiPSCs. TERT is a reverse transcriptase and provides the enzymatic activity of telomerase. This is the first study to target a component of telomerase in hiPSCs, however, there have been previous analyses of the effects of telomerase deficiency in both human and mouse stem cells. Two studies investigated telomerase deficiency in miPSCs generated from *Terc* and *Tert* knockout mice (Wang et al. 2012, Kinoshita et al. 2014). Wang *et al*. (2012) observed a reduction in telomere length of 20-50% in *Terc* null cells over 27 passages. Both of these studies also reported that reprogramming efficiency and pluripotency was reduced when using parental cells from later generation knockout mice, which have shorter telomeres. Although human and mouse telomeres are largely the same the length of mouse telomeres is approximately 5-10 longer than those of humans and means that loss of telomerase might be better tolerated by mouse stem cells. Knockdown of *TERT* in human ESCs has previously been reported to cause reduced proliferation, increased apoptosis and increased differentiation Yang *et al*. (2008). In our CRISPR/Cas9 targeted cells the effect of telomerase deficiency on telomere length was measured by qRT-PCR over more than

60 days of hiPSC culture and homozygous mutant cells showed telomere shortening of around 1% per day, which is a similar rate of attrition to that reported by Wang *et al.* (2012). A trend towards shorter telomeres was also detected in heterozygous mutant cells. Human iPSCs have also been generated from individuals with mutations in various components of telomerase. Batista *et al.* (2011) showed that *DKC1*-mutant iPSCs derived from the somatic cells of dyskeratosis congentia patients lost their self-renewal capacity after prolonged culture. Notably, one of our *TERT*$^{-/-}$ cell lines was observed to have reduced proliferation at the later stages of the extended culture and an increase in expression of differentiation markers, which was possibly due to critical shortening in telomere length. This result requires further investigation including the longer-term culture of other knockout clones, longer-term culture and additional phenotypic analysis in order to identify whether the effects are consistent.

Following the analysis of the effects of telomerase deficiency on telomere shortening in our *TERT*$^{-/-}$ hiPSCs, the novel telomere length associated genes *ACYP2* and *TSPYL6* were investigated. It was expected that targeting these genes and subsequent measurement of telomere length in knockout cell lines would explicitly show whether either of the genes is required for telomere length maintenance. Therefore, CRISPR/ Cas9 was used to target both genes and the effect on telomere length was investigated. Unfortunately, frameshift mutations were not generated in both alleles of the homozygous cell lines for either gene, despite the efficient targeting of 55.3% and 35% for *ACYP2* and *TSPYL6* respectively. Telomere length was measured over approximately 60 days of culture, but no change in length was detected. There are a number of reasons why homozygous frameshift mutations were not generated including homozygous null mutations being detrimental to the growth of iPSCs. Further attempts at targeting these genes should be performed and this should include the use of additional sgRNAs to exclude the lack of homozygous mutants being the result of the sgRNAs used in this study. In addition, heterozygous cell lines should be investigated further in order to determine whether any effect on telomere length can be detected. The phenotype of heterozygous cells would be expected to be more subtle than that for homozygous cell lines, therefore, a larger number of mutant lines should be studied over a longer time-frame.

Recently, data from the GTEx project, that links genetic variants to changes in gene expression variants, became available suggesting *TSPYL6* is probably the causal gene.

The lead telomere length associated SNP is associated with increased expression of *TSPYL6* in germ cells, suggesting that more TSPYL6 causes shorter telomeres. In our mutant *TSPYL6* iPSCs we would therefore expect telomere lengthening during extended culture and although the telomere length was longer in these cells it was within the range seen with wild-type iPSCs and we did not detect a change in length with time. Genome editing could be used to investigate whether overexpression of *TSPYL6* effects on telomere length. Another study in our group involved the knock-in of an inducible promoter upstream of a gene of interest and allows precise control over gene expression (Gong et al. unpublished data) and similar experiments would allow long-term overexpression of *TSPYL6* from its endogenous locus.

Furthermore, in a large-scale interaction study, TSPYL6 was identified as in interactor of MAD2L2 (Huttlin et al 2015) a protein recently reported to mediate the DNA damage response at telomeres (Boersma *et al*. 2015). Utilizing MEFs carrying a mutation in the shelterin component Trf2, which causes NHEJ dependent ligation of telomeres resulting in chromosome fusion and cell death, Boersma *et al*. (2015) performed knockdown of genes implicated in the DNA damage response and screened for cell survival. Mad2l2 was identified as the top hit and found to be a regulator of the DDR that promotes NHEJ at telomeres and DSB repair at genome sites. It is not known if TSPYL6 is involved in the same process, but this obviously demands functional investigation. The mechanism underlying the telomere length associated is unknown for all of the loci and

This was the first study by this group to include both iPSC culture and the CRISPR/Cas9 system and a major objective of the project was to establish these techniques. The GM23720*B hiPSC line from the Apparently Healthy Collection, NIGMS Human Genetic Cell Repository, Coriell was chosen for this project as it is reported to be genetically normal and had previously been cultured in feeder-free media making downstream genome editing experiments more straightforward. The hiPSCs were grown on Matrigel™-coated plates in mTesR®1 and were successfully cultured, appearing healthy and maintaining a stable rate of proliferation. Quality control measures included flow cytometric analysis of pluripotency and differentiation markers. During initial culture of the hiPSCs differentiated cells were observed and expression of differentiation markers was around 20%. This was improved to around 5% by changing the method of cell passage to incorporate the enzyme free reagent ReLeSR™, rather than the enzyme

Dispase II. Optimization of hiPSC culture in terms of genome editing was also performed using a method by Peters *et al*. (2013) as a guide with some alterations. An important influence on the efficiency of genome editing is the delivery of the CRISPR/Cas9 components into the target nucleus and therefore several different transfection techniques were tested. Lipid based transfection methods proved relatively unsuccessful, with less than 30% efficiency achieved for all conditions tested. Electroporation was more effective with efficiencies greater than 80% reached using a NEPA21 electroporator. In this study, targeting efficiency ranged from 23.3% for lipid based transfection to 81.3% for electroporation, while relatively high, could be improved through alteration of the methods. Fluorescent cell sorting or some other form of selection would allow the cells expressing the highest levels of the genome editing machinery to be enriched thereby increasing efficiency. In addition, since the start of this project several alternatives to *S.pyogenes* Cas9 have become available and some of these have been reported to improve efficiency (Hou et al. 2013, Zetsche et al. 2015, Kleinstiver et al. 2016). Other CRISPR/Cas9 delivery formats have also become commercially available and these include viral delivery methods, which should increase delivery efficiency, and Cas9 mRNA or recombinant protein instead of plasmids that removes the need for transcription or translation by the target cell.

6.1.1 Limitations

There are several areas where this study could be improved. First, a major limitation was the failure to generate *ACYP2* and *TSPYL6* cells with the desired genotypes. It is not clear why homozygous knockout lines were not produced and additional targeting of these genes with additional sgRNAs will be needed in order to determine the cause. This study was one of the first CRISPR/Cas9 targeting experiments by both our group and Horizon Discovery Ltd and involved a simple design involving the targeting of the early coding region of each gene in the hope of generating frameshift causing indel mutations. Improved sgRNA design algorithms and alternative targeting strategies such as those described in sections **1.1.2.1.3.1** and **1.1.2.3.1.1** might be help to overcome this issue. Second, off-target effects were not investigated, however, as this study involved a very specific read-out (telomere length) the likelihood that an off-target effect causing a false-positive result is incredibly low. Third, the phenotypic analysis of the targeted cell lines

was limited to the investigation of the expression of pluripotency markers and differentiation markers and further analysis of relevant phenotypes should include embroid body formation, teratoma formation and the expression of the endogenous pluripotent genes including NANOG, OCT4, GDF3 and REX1. In addition, no karyotyping of the CRISPR/Cas9 targeted cell lines was performed. Full cytogenetic analysis of the targeted and control lines would detect any change in chromosome structure such due to the targeting process and long-term culture or phenotypes caused by aberrant telomere function such as altered chromosome number or chromosome fusions. Finally, a major goal of this study was to definitively show which of the genes at the novel telomere length GWAS loci genes is involved in telomere maintenance and this was not achieved. Further optimization of the approach to include additional and more sensitive phenotypic readouts would increase the possibility of identifying the causal gene. These include alternative measures of telomere length such as Southern blotting, TRF, FISH techniques and qPCR-based techniques as well as measuring telomerase activity by TRAP technique or measuring the *TERC* and TERT at the level of the RNA or protein.

6.1.2 Future work.

There are number of aspects of this work that warrant further investigation. Chief among these is the characterization of the genes at the *ACYP2* telomere length associated locus, particularly *TSPYL6*. This should include further investigation of the cells generated in this study, the generation of additional targeted lines and, as described above, more thorough phenotyping. This could include differentiating targeted cells to fast growing cell types such as lung, gut or sin fibroblasts. This would have the benefit of allowing the gene knockout to be investigated in telomerase deficient cell lines as telomerase exrpression should be silenced upon differentiation and the higher-rate of cell division might make telomere related phenotypes more apparent. These experiments should be performed in combination with mechanistic studies to identify the function of the candidate proteins. Further analysis of the *TERT*$^{-/-}$ cells should also be considered and would include a full investigation of the consequences of telomere shortening on iPSC senescence and phenotype.

The screen employed in this study was able to detect a decrease in telomere length in telomerase deficient cells, but not in the *ACYP2* or *TSPYL6* mutant cells. Despite this it could still be a worthwhile approach for the investigation of other telomere length associated SNPs and genes, but with some modifications. The current screen is time-consuming and relatively expensive, but could be optimized. Newer genome editing reagents and improvements to the protocol would allow higher efficiency targeting and increase the number of cells available for study. A high enough efficiency might also mean that there is no need to produce clonal lines. This could also be achieved through the use of an iPSC line modified to constitutively express Cas9. Analysis of an increased number of mutant cells would then allow the detection of changes in telomere length over a shorter time-frame. In addition, other assays could be included to identify the effects of other telomere associated processes. For example, knockout of MAD2L2 might not effect telomere length, but it would affect the DNA damage response. Therefore, the effects of DNA damage response in cells carrying candidate knockout genes could be assayed by measuring the level of chromosomal abnormalities and cell death.

6.1.3 Conclusions

GWAS have provided valuable information regarding the identification of genetic loci that associate with disease and traits, but does not provide information as to which gene at the locus is causal. In this study, the CRISPR/Cas9 genome editing system was utilized in an effort to determine which genes at telomere GWAS loci are involved in telomere length attrition. Both known and novel telomere length associated genes were targeted using CRISPR/Cas9 in hiPSCs, which express telomerase and maintain telomere length. Human iPSCs culture and the techniques required for CRISPR/Cas9 genome editing was optimized and relatively high targeting efficiency was achieved. Telomerase-deficient iPSC lines with homozygous null mutation in *TERT* were successfully generated and after long-term culture a significant reduction in telomere length was observed. In addition, the pluripotency of the mutant cells was not affected by decreasing telomere length, as shown by expression of pluripotency markers, while expression of a differentiation marker increased substantially at the end of long-term culture. iPSCs carrying mutations in *ACYP2* and *TSPYL6*, which are uncharacterized candidates for telomere maintenance,

were also generated, however no impact on telomere length was observed in either case and further investigation is required to identify the causal gene at this locus.

Supplementary

Number of *TERT* clones	Type of Mutation	Predicted genotype	
		Allele 1	**Allele 2**
1	Wild type		
2	Homozygous	1 bp deletion	1 bp insertion
3	Homozygous	[B]Genotype 1 bp deletion	[B]Genotype 1 bp deletion
4	Heterozygous	1 bp insertion	Wild type
5	Homozygous	9 bp deletion	Unknown[A]
6	Wild type		
7	Wild type		
8	Wild type		
9	Heterozygous	1 bp insertion	Wild type
10	Wild type		

11	Homozygous	1 bp insertion	Unknown[A]
12	Wild type		
13	Wild type		
14	Heterozygous	Wild type	1 bp deletion
15	Wild type		
16	Wild type		
17	Wild type		
18	Wild type		
19	Heterozygous	1 bp deletion	Wild type
20	Wild type		
21	Wild type		
22	Wild type		
23	Wild type		
24	Wild type		
25	Homozygous	[B]Genotype	[B]Genotype

		1 bp insertion	1 bp insertion
26	Heterozygous	10 bp deletion	Unknown[A]
27	Wild type		
28	Wild type		
29	Wild type		
30	Heterozygous	Wild type	1 bp insertion
31	Homozygous	Unknown[A]	1 bp insertion
32	Wild type		
33	Homozygous	1 bp deletion	Unknown[A]
34	Wild type		
36	Homozygous	4 bp deletion	1 bp insertion

Table 1: Number and type of mutation in *TERT* in targeted iPSC clones.

Number of *ACYP2* clones	Type of Mutation	Predicted genotype	
		Allele 1	**Allele 2**
1	Heterozygous	Wild type	1 bp insertion
2	Heterozygous	1 bp insertion and 1bp deletion	Wild type
4	Homozygous	[B]Genotype 1 bp insertion	[B]Genotype 9 bp deletion
6	Homozygous	9 bp deletion	1 bp insertion
7	Wild type		
8	Homozygous	6 bp deletion	Unknown[A]
9	Homozygous	[B]Genotype 1 bp insertion	[B]Genotype 7 bp deletion and 16 insertion
10	Wild type		
11	Wild type		

12	Homozygous	Unknown[A]	Unknown[A]
15	Wild type		
16	Wild type		
17	Homozygous	Unknown[A]	1 bp insertion
18	Homozygous	[B]Genotype 1 bp insertion	[B]Genotype 9 bp deletion
19	Wild type		
20	Heterozygous	Wild type	1 bp insertion
21	Homozygous	Unknown[A]	1 bp insertion
22	Homozygous	Unknown[A]	3 bp insertion
23	Wild type		
24	Homozygous	9 bp deletion	1 bp insertion
25	Heterozygous	9 bp deletion	Wild type
27	Wild type		

28	Heterozygous	Unknown[A]	Wild type
30	Heterozygous	Unknown[A]	Wild type
33	Homozygous	Unknown[A]	I bp insertion
34	Homozygous	9 bp deletion	1 bp insertion
35	Wild type		
36	Heterozygous	9 bp deletion	Wild type

Table 2: Number and type of mutation in *ACYP2* in targeted iPSC clones.

Number of *TSPYL6* clones	**Type of Mutation**	**Predicted genotype**	
		Allele 1	**Allele 2**
1	Heterozygous	Unknown[A]	Wild type
2	Heterozygous	1 bp deletion	Wild type
3	Heterozygous	1 bp deletion	Wild type
4	Heterozygous	1 bp deletion	Wild type

6	Heterozygous	Wild type	Unknown[A]
7	Heterozygous	1 pb deletion	Wild type
10	Wild type		
11	Heterozygous	1 bp deletion	Wild type
12	Heterozygous	1 bp deletion	Wild type
13	Homozygous	[B]Genotype 6 bp deletion	[B]Genotype 18 bp deletion and 1 bp deletion
14	Heterozygous	1 bp deletion	Wild type
15	Heterozygous	Wild type	1 bp insertion
16	Wild type		
17	Heterozygous	1 pb deletion	Wild type
18	Heterozygous	1 bp deletion	Wild type
19	Wild type		
20	Heterozygous	1 bp deletion	Wild type
21	Heterozygous	1 bp deletion	Wild type
22	Wild type		

23	Heterozygous	1 bp deletion	Wild type
24	Wild type		
25	Wild type		
26	Wild type		
27	Heterozygous	1 bp deletion	Wild type
28	Heterozygous	Wild type	3 pb insertion
31	Wild type		
32	Heterozygous	1 bp deletion	Wild type
33	Wild type		
35	Wild type		
36	Heterozygous	1 pb deletion	Wild type

Table 3: Number and type of mutation in *TSPYL6* in targeted iPSC clones.

Genotype of each allele of all of the CRISPR/Cas9 targeted clones as predicted using the TIDE sequence deconvolution tool.

[A] Unknown mutant allele. TIDE does not predict indels greater than 10 bp.

[B] Genotype was found to be incorrect following cloning and Sanger sequencing.

Permission

ELSEVIER LICENSE
TERMS AND CONDITIONS

Aug 06, 2017

This Agreement between Miss. Noor Shamkhi -- Noor Shamkhi ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	4163180113481
License date	Aug 06, 2017
Licensed Content Publisher	Elsevier
Licensed Content Publication	DNA Repair
Licensed Content Title	Give me a break: How telomeres suppress the DNA damage response
Licensed Content Author	Eros Lazzerini Denchi
Licensed Content Date	Sep 2, 2009
Licensed Content Volume	8
Licensed Content Issue	9
Licensed Content Pages	9
Start Page	1118
End Page	1126
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Original figure numbers	Figure 2
Title of your thesis/dissertation	generation of knockout human iPSCs to investigate genes associated with telomere length
Expected completion date	Sep 2017

Bibliography

- Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilić, J., Pekarik, V. and Tiscornia, G. (2008) 'Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes', *Nature biotechnology,* 26(11), pp. 1276-1284.
- Adaikalakoteswari, A., Balasubramanyam, M. and Mohan, V. (2005) 'Telomere shortening occurs in Asian Indian Type 2 diabetic patients', *Diabetic Medicine,* 22(9), pp. 1151-1156.
- Adewumi, O., Aflatoonian, B., Ahrlund-Richter, L., Amit, M., Andrews, P.W., Beighton, G., Bello, P.A., Benvenisty, N., Berry, L.S. and Bevan, S. (2007) 'Characterization of human embryonic stem cell lines by the International Stem Cell Initiative', *Nature biotechnology,* 25(7), pp. 803-816.
- Al-Attar, S., Westra, E.R., van der Oost, J. and Brouns, S.J. (2011) 'Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes', *Biological chemistry,* 392(4), pp. 277-289.
- Allshire, R.C., Dempster, M. and Hastie, N.D. (1989) 'Human telomeres contain at least three types of G–rich repeat distributed non-randomly', *Nucleic acids research,* 17(12), pp. 4611-4627.
- Allsopp, R. (2012) 'Telomere length and iPSC re-programming: survival of the longest', *Cell research,* 22(4), pp. 614.
- Alter, B.P., Baerlocher, G.M., Savage, S.A., Chanock, S.J., Weksler, B.B., Willner, J.P., Peters, J.A., Giri, N. and Lansdorp, P.M. (2007) 'Very short telomere length by flow fluorescence in situ hybridization identifies patients with dyskeratosis congenita', *Blood,* 110(5), pp. 1439-1447.
- Amit, M. and Itskovitz-Eldor, J. (2012) 'Morphology of human embryonic and induced pluripotent stem cell colonies cultured with feeders', in Anonymous *Atlas of Human Pluripotent Stem Cells.* Springer, pp. 15-39.
- Amit, M., Margulets, V., Segev, H., Shariki, K., Laevsky, I., Coleman, R. and Itskovitz-Eldor, J. (2003) 'Human feeder layers for human embryonic stem cells', *Biology of reproduction,* 68(6), pp. 2150-2156.

- Andersson, A.F. and Banfield, J.F. (2008) 'Virus population dynamics and acquired virus resistance in natural microbial communities', *Science (New York, N.Y.),* 320(5879), pp. 1047-1050.
- Andrew, T., Aviv, A., Falchi, M., Surdulescu, G.L., Gardner, J.P., Lu, X., Kimura, M., Kato, B.S., Valdes, A.M. and Spector, T.D. (2006) 'Mapping genetic loci that determine leukocyte telomere length in a large sample of unselected female sibling pairs', *The American Journal of Human Genetics,* 78(3), pp. 480-486.
- Aoi, T., Yae, K., Nakagawa, M., Ichisaka, T., Okita, K., Takahashi, K., Chiba, T. and Yamanaka, S. (2008) 'Generation of pluripotent stem cells from adult mouse liver and stomach cells', *Science (New York, N.Y.),* 321(5889), pp. 699-702.
- Aouizerat, B.E., Vittinghoff, E., Musone, S.L., Pawlikowska, L., Kwok, P., Olgin, J.E. and Tseng, Z.H. (2011) 'GWAS for discovery and replication of genetic loci associated with sudden cardiac arrest in patients with coronary artery disease', *BMC cardiovascular disorders,* 11(1), pp. 1.
- Armanios, M.Y., Chen, J.J., Cogan, J.D., Alder, J.K., Ingersoll, R.G., Markin, C., Lawson, W.E., Xie, M., Vulto, I. and Phillips III, J.A. (2007) 'Telomerase mutations in families with idiopathic pulmonary fibrosis', *New England Journal of Medicine,* 356(13), pp. 1317-1326.
- Artandi, S.E., Chang, S., Lee, S. and Alson, S. (2000) 'Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice', *Nature,* 406(6796), pp. 641.
- Aubert, G., Hills, M. and Lansdorp, P.M. (2012) 'Telomere length measurement—Caveats and a critical assessment of the available technologies and tools', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* 730(1), pp. 59-67.
- Aubert, G., Hills, M. and Lansdorp, P.M. (2012) 'Telomere length measurement—Caveats and a critical assessment of the available technologies and tools', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* 730(1), pp. 59-67.
- Aviv, A., Hunt, S.C., Lin, J., Cao, X., Kimura, M. and Blackburn, E. (2011) 'Impartial comparative analysis of measurement of leukocyte telomere

length/DNA content by Southern blots and qPCR', *Nucleic acids research,* 39(20), pp. e134-e134.

- Aviv, H., Yusuf Khan, M., Skurnick, J., Okuda, K., Kimura, M., Gardner, J., Priolo, L. and Aviv, A. (2001) 'Age dependent aneuploidy and telomere length of the human vascular endothelium', *Atherosclerosis,* 159(2), pp. 281-287.
- Ballew, B.J., Yeager, M., Jacobs, K., Giri, N., Boland, J., Burdett, L., Alter, B.P. and Savage, S.A. (2013) 'Germline mutations of regulator of telomere elongation helicase 1, RTEL1, in dyskeratosis congenita', Human genetics, 132(4), pp. 473-480.
- Ban, H., Nishishita, N., Fusaki, N., Tabata, T., Saeki, K., Shikamura, M., Takada, N., Inoue, M., Hasegawa, M., Kawamata, S. and Nishikawa, S. (2011) 'Efficient generation of transgene-free human induced pluripotent stem cells (iPSCs) by temperature-sensitive Sendai virus vectors', *Proceedings of the National Academy of Sciences of the United States of America,* 108(34), pp. 14234-14239.
- Banik, S.S., Guo, C., Smith, A.C., Margolis, S.S., Richardson, D.A., Tirado, C.A. and Counter, C.M. (2002) 'C-terminal regions of the human telomerase catalytic subunit essential for in vivo enzyme activity', *Molecular and cellular biology,* 22(17), pp. 6234-6246.
- Barber, L.J., Youds, J.L., Ward, J.D., McIlwraith, M.J., O'Neil, N.J., Petalcorin, M.I., Martin, J.S., Collis, S.J., Cantor, S.B. and Auclair, M. (2008) 'RTEL1 maintains genomic stability by suppressing homologous recombination', *Cell,* 135(2), pp. 261-271.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A. and Horvath, P. (2007) 'CRISPR provides acquired resistance against viruses in prokaryotes', *Science (New York, N.Y.),* 315(5819), pp. 1709-1712.
- Barrett, J.C., Hansoul, S., Nicolae, D.L., Cho, J.H., Duerr, R.H., Rioux, J.D., Brant, S.R., Silverberg, M.S., Taylor, K.D. and Barmada, M.M. (2008) 'Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease', *Nature genetics,* 40(8), pp. 955-962.
- Bassett, A.R. (2017) 'Editing the genome of hiPSC with CRISPR/Cas9: disease models', *Mammalian Genome,* , pp. 1-17.

- Bassett, A.R., Tibbit, C., Ponting, C.P. and Liu, J. (2013) 'Highly Efficient Targeted Mutagenesis of< i> Drosophila</i> with the CRISPR/Cas9 System', *Cell reports,* 4(1), pp. 220-228.
- Bassuk, A.G., Zheng, A., Li, Y., Tsang, S.H. and Mahajan, V.B. (2016) 'Precision Medicine: Genetic Repair of Retinitis Pigmentosa in Patient-Derived Stem Cells', *Scientific reports,* 6, pp. 19969.
- Batista, L.F., Pech, M.F., Zhong, F.L., Nguyen, H.N., Xie, K.T., Zaug, A.J., Crary, S.M., Choi, J., Sebastiano, V. and Cherry, A. (2011) 'Telomere shortening and loss of self-renewal in dyskeratosis congenita induced pluripotent stem cells', *Nature,* 474(7351), pp. 399-402.
- Bauman, J., Wiegant, J., Borst, P. and Van Duijn, P. (1980) 'A new method for fluorescence microscopical localization of specific DNA sequences by in situ hybridization of fluorochrome-labelled RNA', *Experimental cell research,* 128(2), pp. 485-490.
- Beerli, R.R. and Barbas, C.F. (2002) 'Engineering polydactyl zinc-finger transcription factors', *Nature biotechnology,* 20(2), pp. 135-141.
- Benetos, A., Gardner, J.P., Zureik, M., Labat, C., Xiaobin, L., Adamopoulos, C., Temmar, M., Bean, K.E., Thomas, F. and Aviv, A. (2004) 'Short telomeres are associated with increased carotid atherosclerosis in hypertensive subjects', *Hypertension,* 43(2), pp. 182-185.
- Bhutani, K., Nazor, K.L., Williams, R., Tran, H., Dai, H., Dzakula, Z., Cho, E.H., Pang, A.W., Rao, M., Cao, H., Schork, N.J. and Loring, J.F. (2016) 'Whole-genome mutational burden analysis of three pluripotency induction methods', *Nature communications,* 7, pp. 10536.
- Bibikova, M., Beumer, K., Trautman, J.K. and Carroll, D. (2003) 'Enhancing gene targeting with designed zinc finger nucleases', *Science (New York, N.Y.),* 300(5620), pp. 764.
- Bischoff, C., Graakjaer, J., Petersen, H.C., Hjelmborg, J.v.B., Vaupel, J.W., Bohr, V., Koelvraa, S. and Christensen, K. (2005) 'The heritability of telomere length among the elderly and oldest-old', *Twin Research and Human Genetics,* 8(05), pp. 433-439.

- Blasco, M.A., Lee, H., Hande, M.P., Samper, E., Lansdorp, P.M., DePinho, R.A. and Greider, C.W. (1997) 'Telomere shortening and tumor formation by mouse cells lacking telomerase RNA', *Cell,* 91(1), pp. 25-34.
- Blelloch, R., Venere, M., Yen, J. and Ramalho-Santos, M. (2007) 'Generation of induced pluripotent stem cells in the absence of drug selection', *Cell stem cell,* 1(3), pp. 245-247.
- Boch, J. and Bonas, U. (2010) 'Xanthomonas AvrBs3 family-type III effectors: discovery and function', *Annual Review of Phytopathology,* 48, pp. 419-436.
- Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A. and Bonas, U. (2009) 'Breaking the code of DNA binding specificity of TAL-type III effectors', *Science (New York, N.Y.),* 326(5959), pp. 1509-1512.
- Bodnar, A.G., Ouellette, M., Frolkis, M., Holt, S.E., Chiu, C.P., Morin, G.B., Harley, C.B., Shay, J.W., Lichtsteiner, S. and Wright, W.E. (1998) 'Extension of life-span by introduction of telomerase into normal human cells', *Science (New York, N.Y.),* 279(5349), pp. 349-352.
- Boersma, V., Moatti, N., Segura-Bayona, S., Peuscher, M.H., Van der Torre, J., Wevers, B.A., Orthwein, A., Durocher, D. and Jacobs, J.J. (2015) 'MAD2L2 controls DNA repair at telomeres and DNA breaks by inhibiting 5 [prime] end resection', *Nature,* 521(7553), pp. 537-540.
- Bojesen, S.E., Pooley, K.A., Johnatty, S.E., Beesley, J., Michailidou, K., Tyrer, J.P., Edwards, S.L., Pickett, H.A., Shen, H.C. and Smart, C.E. (2013) 'Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer', *Nature genetics,* 45(4), pp. 371-384.
- Bosnakovski, D., Islam, S., Nandez, R., Zaidman, N., Struck, M., Dandapat, A. and Kyba, M. (2012) 'Maintenance of Human iPS Cells in a Feeder-free Culture System', *BD Biosciences,* .
- Boulting, G.L., Kiskinis, E., Croft, G.F., Amoroso, M.W., Oakley, D.H., Wainger, B.J., Williams, D.J., Kahler, D.J., Yamaki, M. and Davidow, L. (2011) 'A functionally characterized test set of human induced pluripotent stem cells', *Nature biotechnology,* 29(3), pp. 279-286.
- Brambrink, T., Foreman, R., Welstead, G.G., Lengner, C.J., Wernig, M., Suh, H. and Jaenisch, R. (2008) 'Sequential expression of pluripotency markers

during direct reprogramming of mouse somatic cells', *Cell stem cell,* 2(2), pp. 151-159.

- Brennand, K.J., Simone, A., Jou, J., Gelboin-Burkhart, C., Tran, N., Sangar, S., Li, Y., Mu, Y., Chen, G. and Yu, D. (2011) 'Modelling schizophrenia using human induced pluripotent stem cells', *Nature,* 473(7346), pp. 221-225.
- Brinkman, E.K., Chen, T., Amendola, M. and van Steensel, B. (2014) 'Easy quantitative assessment of genome editing by sequence trace decomposition', *Nucleic acids research,* 42(22), pp. e168.
- Broer, L., Codd, V., Nyholt, D.R., Deelen, J., Mangino, M., Willemsen, G., Albrecht, E., Amin, N., Beekman, M., de Geus, E. J., Henders, A., Nelson, C. P., Steves, C. J., Wright, M. J., de Craen, A. J., Isaacs, A., Matthews, M., Moayyeri, A., Montgomery, G. W., Oostra, B. A., Vink, J. M., Spector, T. D., Slagboom, P. E., Martin, N. G., Samani, N. J., van Duijn, C. M., & Boomsma, D. I. 2013, "Meta-analysis of telomere length in 19713 subjects reveals high heritability, stronger maternal inheritance and a paternal age effect", European journal of human genetics, doi:10.1038/ejhg.2012.303, pp. 1-6.7
- Brouilette, S.W., Whittaker, A., Stevens, S.E., van der Harst, P., Goodall, A.H. and Samani, N.J. (2008) 'Telomere length is shorter in healthy offspring of subjects with coronary artery disease: support for the telomere hypothesis', *Heart (British Cardiac Society),* 94(4), pp. 422-425.
- Brouwer, M., Zhou, H. and Kasri, N.N. (2016) 'Choices for induction of pluripotency: recent developments in human induced pluripotent stem cell reprogramming strategies', *Stem Cell Reviews and Reports,* 12(1), pp. 54-72.
- Bryan, T.M., Englezou, A., Dalla-Pozza, L., Dunham, M.A. and Reddel, R.R. (1997) 'Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines', *Nature medicine,* 3(11), pp. 1271-1274.
- Bryan, T.M., Englezou, A., Gupta, J., Bacchetti, S. and Reddel, R.R. (1995) 'Telomere elongation in immortal human cells without detectable telomerase activity', *The EMBO journal,* 14(17), pp. 4240-4248.
- Bush, W.S. and Moore, J.H. (2012) 'Genome-wide association studies', *PLoS Comput Biol,* 8(12), pp. e1002822.

- Caisander, G., Park, H., Frej, K., Lindqvist, J., Bergh, C., Lundin, K. and Hanson, C. (2006) 'Chromosomal integrity maintained in five human embryonic stem cell lines after prolonged in vitro culture', *Chromosome Research,* 14(2), pp. 131-137.
- Calado, R.T. & Dumitriu, B. (2013) 'Telomere dynamics in mice and humans', *Seminars in hematology*Elsevier, pp. 165.
- Campisi, J. (2001) 'Cellular senescence as a tumor-suppressor mechanism', *Trends in cell biology,* 11(11), pp. S27-S31.
- Campisi, J., Kim, S., Lim, C. and Rubio, M. (2001) 'Cellular senescence, cancer and aging: the telomere connection', *Experimental gerontology,* 36(10), pp. 1619-1637.
- Canela, A., Vera, E., Klatt, P. and Blasco, M.A. (2007) 'High-throughput telomere length quantification by FISH and its application to human population studies', *Proceedings of the National Academy of Sciences of the United States of America,* 104(13), pp. 5300-5305.
- Carbery, I.D., Ji, D., Harrington, A., Brown, V., Weinstein, E.J., Liaw, L. and Cui, X. (2010) 'Targeted genome modification in mice using zinc-finger nucleases', *Genetics,* 186(2), pp. 451-459.
- Carlson, D.F., Fahrenkrug, S.C. and Hackett, P.B. (2012) 'Targeting DNA with fingers and TALENs', *Molecular therapy.Nucleic acids,* 1(1), pp. e3.
- Carlson, S.G., Eng, E., Kim, E.G., Perlman, E.J., Copeland, T.D. and Ballermann, B.J. (1998) 'Expression of SET, an inhibitor of protein phosphatase 2A, in renal development and Wilms' tumor', *Journal of the American Society of Nephrology : JASN,* 9(10), pp. 1873-1880.
- Carroll, D. (2008) 'Progress and prospects: zinc-finger nucleases as gene therapy agents', *Gene therapy,* 15(22), pp. 1463-1468.
- Carroll, K.A. and Ly, H. (2009) 'Telomere dysfunction in human diseases: the long and short of it!', *International journal of clinical and experimental pathology,* 2(6), pp. 528-543.
- Cassidy, A., De Vivo, I., Liu, Y., Han, J., Prescott, J., Hunter, D.J. and Rimm, E.B. (2010) 'Associations between diet, lifestyle factors, and telomere length in women', *The American Journal of Clinical Nutrition,* 91(5), pp. 1273-1280.

- Casteel, D.E., Zhuang, S., Zeng, Y., Perrino, F.W., Boss, G.R., Goulian, M. and Pilz, R.B. (2009) 'A DNA polymerase-{alpha}{middle dot}primase cofactor with homology to replication protein A-32 regulates DNA replication in mammalian cells', *The Journal of biological chemistry,* 284(9), pp. 5807-5818.
- Cathomen, T. and Joung, J.K. (2008) 'Zinc-finger nucleases: the next generation emerges', *Molecular Therapy,* 16(7), pp. 1200-1207.
- Cawthon, R.M. (2002) 'Telomere measurement by quantitative PCR', *Nucleic acids research,* 30(10), pp. e47.
- Cawthon, R.M. (2009) 'Telomere length measurement by a novel monochrome multiplex quantitative PCR method', *Nucleic acids research,* 37(3), pp. e21-e21.
- Cawthon, R.M. (2009) 'Telomere length measurement by a novel monochrome multiplex quantitative PCR method', *Nucleic acids research,* 37(3), pp. e21-e21.
- Chan, E.M., Ratanasirintrawoot, S., Park, I., Manos, P.D., Loh, Y., Huo, H., Miller, J.D., Hartung, O., Rho, J. and Ince, T.A. (2009) 'Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells', *Nature biotechnology,* 27(11), pp. 1033-1037.
- Chang, E. and Harley, C.B. (1995) 'Telomere length and replicative aging in human vascular tissues', *Proceedings of the National Academy of Sciences of the United States of America,* 92(24), pp. 11190-11194.
- Chatterjee, P., Cheung, Y. and Liew, C. (2011) 'Transfecting and nucleofecting human induced pluripotent stem cells', *JoVE (Journal of Visualized Experiments),* (56), pp. e3110-e3110.
- Chen, G., Gulbranson, D.R., Hou, Z., Bolin, J.M., Ruotti, V., Probasco, M.D., Smuga-Otto, K., Howden, S.E., Diol, N.R. and Propson, N.E. (2011) 'Chemically defined conditions for human iPSC derivation and culture', *Nature methods,* 8(5), pp. 424-429.
- Cheng, L., Hammond, H., Ye, Z., Zhan, X. and Dravid, G. (2003) 'Human adult marrow cells support prolonged expansion of human embryonic stem cells in culture', *Stem cells,* 21(2), pp. 131-142.
- Chiang, Y.J., Hemann, M.T., Hathcock, K.S., Tessarollo, L., Feigenbaum, L., Hahn, W.C. and Hodes, R.J. (2004) 'Expression of telomerase RNA template,

but not telomerase reverse transcriptase, is limiting for telomere length maintenance in vivo', *Molecular and cellular biology,* 24(16), pp. 7024-7031.

- Chiarugi, P., Degl'Innocenti, D., Raugei, G., Fiaschi, T. and Ramponi, G. (1997) 'Differential migration of acylphosphatase isoenzymes from cytoplasm to nucleus during apoptotic cell death', *Biochemical and biophysical research communications,* 231(3), pp. 717-721.
- Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiuwu, O., Richter, L. and Zhang, J. (2009) 'Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures', *Cell stem cell,* 5(1), pp. 111-123.
- Cho, S.W., Kim, S., Kim, J.M. and Kim, J. (2013) 'Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease', *Nature biotechnology,* 31(3), pp. 230-232.
- Cho, S.W., Kim, S., Kim, J.M. and Kim, J. (2013) 'Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease', *Nature biotechnology,* 31(3), pp. 230-232.
- Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S. and Kim, J.S. (2014) 'Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases', *Genome research,* 24(1), pp. 132-141.
- Codd, V., Mangino, M., van der Harst, P., Braund, P.S., Kaiser, M., Beveridge, A.J., Rafelt, S., Moore, J., Nelson, C. and Soranzo, N. (2010) 'Common variants near TERC are associated with mean telomere length', *Nature genetics,* 42(3), pp. 197-199.
- Codd, V., Nelson, C.P., Albrecht, E., Mangino, M., Deelen, J., Buxton, J.L., Hottenga, J.J., Fischer, K., Esko, T. and Surakka, I. (2013) 'Identification of seven loci affecting mean telomere length and their association with disease', *Nature genetics,* 45(4), pp. 422-427.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A. and Zhang, F. (2013) 'Multiplex genome engineering using CRISPR/Cas systems', *Science (New York, N.Y.),* 339(6121), pp. 819-823.

- Cowan, C.A., Atienza, J., Melton, D.A. and Eggan, K. (2005) 'Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells', *Science (New York, N.Y.),* 309(5739), pp. 1369-1373.
- Cunningham, J.M., Johnson, R.A., Litzelman, K., Skinner, H.G., Seo, S., Engelman, C.D., Vanderboom, R.J., Kimmel, G.W., Gangnon, R.E., Riegert-Johnson, D.L., Baron, J.A., Potter, J.D., Haile, R., Buchanan, D.D., Jenkins, M.A., Rider, D.N., Thibodeau, S.N., Petersen, G.M. and Boardman, L.A. (2013) 'Telomere length varies by DNA extraction method: implications for epidemiologic research', *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology,* 22(11), pp. 2047-2054.
- Cyranoski, D. (2014) 'Japanese woman is first recipient of next-generation stem cells', *Nature,* 12.
- Daniali, L., Benetos, A., Susser, E., Kark, J.D., Labat, C., Kimura, M., Desai, K.K., Granick, M. and Aviv, A. (2013) 'Telomeres shorten at equivalent rates in somatic tissues of adults', *Nature communications,* 4, pp. 1597.
- Davis, R.L., Weintraub, H. and Lassar, A.B. (1987) 'Expression of a single transfected cDNA converts fibroblasts to myoblasts', *Cell,* 51(6), pp. 987-1000.
- de Lange, T. (2005) 'Shelterin: the protein complex that shapes and safeguards human telomeres', *Genes & development,* 19(18), pp. 2100-2110.
- De Meyer, T., Rietzschel, E.R., De Buyzere, M.L., De Bacquer, D., Van Criekinge, W., De Backer, G.G., Gillebert, T.C., Van Oostveldt, P., Bekaert, S. and Asklepios investigators (2007) 'Paternal age at birth is an important determinant of offspring telomere length', *Human molecular genetics,* 16(24), pp. 3097-3102.
- Denchi, E.L. (2009) 'Give me a break: how telomeres suppress the DNA damage response', *DNA repair,* 8(9), pp. 1118-1126.
- Deveau, H., Garneau, J.E. and Moineau, S. (2010) 'CRISPR/Cas system and its role in phage-bacteria interactions', *Annual Review of Microbiology,* 64, pp. 475-493.
- Devine, M.J., Ryten, M., Vodicka, P., Thomson, A.J., Burdon, T., Houlden, H., Cavaleri, F., Nagano, M., Drummond, N.J. and Taanman, J. (2011)

'Parkinson's disease induced pluripotent stem cells with triplication of the α-synuclein locus', *Nature communications,* 2, pp. 440.

- di Fagagna, F.d., Reaper, P.M., Clay-Farrace, L., Fiegler, H., Carr, P., von Zglinicki, T., Saretzki, G., Carter, N.P. and Jackson, S.P. (2003) 'A DNA damage checkpoint response in telomere-initiated senescence', *Nature,* 426(6963), pp. 194-198.
- DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J. and Church, G.M. (2013) 'Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems', *Nucleic acids research,* 41(7), pp. 4336-4343.
- Dimos, J.T., Rodolfa, K.T., Niakan, K.K., Weisenthal, L.M., Mitsumoto, H., Chung, W., Croft, G.F., Saphier, G., Leibel, R., Goland, R., Wichterle, H., Henderson, C.E. and Eggan, K. (2008) 'Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons', *Science (New York, N.Y.),* 321(5893), pp. 1218-1221.
- Dimri, G.P., Lee, X., Basile, G., Acosta, M., Scott, G., Roskelley, C., Medrano, E.E., Linskens, M., Rubelj, I. and Pereira-Smith, O. (1995) 'A biomarker that identifies senescent human cells in culture and in aging skin in vivo', *Proceedings of the National Academy of Sciences of the United States of America,* 92(20), pp. 9363-9367.
- Ding, C. and Cantor, C.R. (2004) 'Quantitative analysis of nucleic acids-the last few years of progress', *BMB Reports,* 37(1), pp. 1-10.
- Ding, H., Schertzer, M., Wu, X., Gertsenstein, M., Selig, S., Kammori, M., Pourvali, R., Poon, S., Vulto, I. and Chavez, E. (2004) 'Regulation of Murine Telomere Length by< i> Rtel</i>: An Essential Gene Encoding a Helicase-like Protein', *Cell,* 117(7), pp. 873-886.
- Ding, Q., Lee, Y., Schaefer, E.A., Peters, D.T., Veres, A., Kim, K., Kuperwasser, N., Motola, D.L., Meissner, T.B. and Hendriks, W.T. (2013) 'A TALEN genome-editing system for generating human stem cell-based disease models', *Cell Stem Cell,* 12(2), pp. 238-251.
- Ding, Q., Regan, S.N., Xia, Y., Oostrom, L.A., Cowan, C.A. and Musunuru, K. (2013) 'Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs', *Cell stem cell,* 12(4), pp. 393-394.

- Dityateva, G., Hammond, M., Thiel, C., Ruonala, M.O., Delling, M., Siebenkotten, G., Nix, M. and Dityatev, A. (2003) 'Rapid and efficient electroporation-based gene transfer into primary dissociated neurons', *Journal of neuroscience methods,* 130(1), pp. 65-73.
- Dokal, I. (2000) 'Dyskeratosis congenita in all its forms', *British journal of haematology,* 110(4), pp. 768-779.
- Doksani, Y., Wu, J.Y., de Lange, T. and Zhuang, X. (2013) 'Super-resolution fluorescence imaging of telomeres reveals TRF2-dependent T-loop formation', *Cell,* 155(2), pp. 345-356.
- Doyon, Y., McCammon, J.M., Miller, J.C., Faraji, F., Ngo, C., Katibah, G.E., Amora, R., Hocking, T.D., Zhang, L. and Rebar, E.J. (2008) 'Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases', *Nature biotechnology,* 26(6), pp. 702-708.
- Draper, J.S., Pigott, C., Thomson, J.A. and Andrews, P.W. (2002) 'Surface antigens of human embryonic stem cells: changes upon differentiation in culture', *Journal of anatomy,* 200(3), pp. 249-258.
- Draper, J.S., Smith, K., Gokhale, P., Moore, H.D., Maltby, E., Johnson, J., Meisner, L., Zwaka, T.P., Thomson, J.A. and Andrews, P.W. (2004) 'Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells', *Nature biotechnology,* 22(1), pp. 53-54.
- Dunham, M.A., Neumann, A.A., Fasching, C.L. and Reddel, R.R. (2000) 'Telomere maintenance by recombination in human cells', *Nature genetics,* 26(4), pp. 447-450.
- Egan, E.D. and Collins, K. (2012) 'An enhanced H/ACA RNP assembly mechanism for human telomerase RNA', *Molecular and cellular biology,* 32(13), pp. 2428-2439.
- Eichler, E.E., Hoffman, S.M., Adamson, A.A., Gordon, L.A., McCready, P., Lamerdin, J.E. and Mohrenweiser, H.W. (1998) 'Complex beta-satellite repeat structures and the expansion of the zinc finger gene cluster in 19p12', *Genome research,* 8(8), pp. 791-808.
- Eminli, S., Utikal, J., Arnold, K., Jaenisch, R. and Hochedlinger, K. (2008) 'Reprogramming of neural progenitor cells into induced pluripotent stem cells

in the absence of exogenous Sox2 expression', *Stem cells,* 26(10), pp. 2467-2474.

- Epel, E.S., Blackburn, E.H., Lin, J., Dhabhar, F.S., Adler, N.E., Morrow, J.D. and Cawthon, R.M. (2004) 'Accelerated telomere shortening in response to life stress', *Proceedings of the National Academy of Sciences of the United States of America*, 101(49), pp. 17312-17315.
- Epping, M., Lunardi, A., Nachmani, D., Castillo-Martin, M., Thin, T., Cordon-Cardo, C. and Pandolfi, P. (2015) 'TSPYL2 is an essential component of the REST/NRSF transcriptional complex for TGFβ signaling activation', *Cell Death & Differentiation,* 22(8), pp. 1353-1362.
- Epping, M.T., Meijer, L.A., Krijgsman, O., Bos, J.L., Pandolfi, P.P. and Bernards, R. (2011) 'TSPYL5 suppresses p53 levels and function by physical interaction with USP7', *Nature cell biology,* 13(1), pp. 102-108.
- Evans, M.J. and Kaufman, M.H. (1981) 'Establishment in culture of pluripotential cells from mouse embryos', *Nature,* 292(5819), pp. 154-156.
- Fiaschi, T., Raugei, G., Marzocchini, R., Chiarugi, P., Cirri, P. and Ramponi, G. (1995) 'Cloning and expression of the cDNA coding for the erythrocyte isoenzyme of human acylphosphatase', *FEBS letters,* 367(2), pp. 145-148.
- Fingerlin, T.E., Murphy, E., Zhang, W., Peljto, A.L., Brown, K.K., Steele, M.P., Loyd, J.E., Cosgrove, G.P., Lynch, D. and Groshong, S. (2013) 'Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis', *Nature genetics,* 45(6), pp. 613-620.
- Flisikowska, T., Thorey, I.S., Offner, S., Ros, F., Lifke, V., Zeitler, B., Rottmann, O., Vincent, A., Zhang, L. and Jenkins, S. (2011) 'Efficient immunoglobulin gene disruption and targeted replacement in rabbit using zinc finger nucleases', *PloS one,* 6(6), pp. e21045.
- Flores, I., Benetti, R. and Blasco, M.A. (2006) 'Telomerase regulation and stem cell behaviour', *Current opinion in cell biology,* 18(3), pp. 254-260.
- Freiermuth, J.L., Powell-Castilla, I.J. and Gallicano, G.I. (2017) 'Towards a CRISPR picture: Use of CRISPR/Cas9 to model diseases in human stem cells in vitro', *Journal of cellular biochemistry,* .

- Friedland, A.E., Tzur, Y.B., Esvelt, K.M., Colaiácovo, M.P., Church, G.M. and Calarco, J.A. (2013) 'Heritable genome editing in C. elegans via a CRISPR-Cas9 system', *Nature methods,* 10(8), pp. 741-743.
- Fu, D. and Collins, K. (2007) 'Purification of human telomerase complexes identifies factors involved in telomerase biogenesis and telomere length regulation', *Molecular cell,* 28(5), pp. 773-785.
- Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K. and Sander, J.D. (2013) 'High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells', *Nature biotechnology,* 31(9), pp. 822-826.
- Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M. and Joung, J.K. (2014) 'Improving CRISPR-Cas nuclease specificity using truncated guide RNAs', *Nature biotechnology,* 32(3), pp. 279-284.
- Fujita, J., Itabashi, Y., Seki, T., Tohyama, S., Tamura, Y., Sano, M. and Fukuda, K. (2012) 'Myocardial cell sheet therapy and cardiac function', *American journal of physiology.Heart and circulatory physiology,* 303(10), pp. H1169-82.
- Fusaki, N., Ban, H., Nishiyama, A., Saeki, K. and Hasegawa, M. (2009) 'Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome', *Proceedings of the Japan Academy, Series B,* 85(8), pp. 348-362.
- Gabourdes, M., Bourgine, V., Mathis, G., Bazin, H. and Alpha-Bazin, B. (2004) 'A homogeneous time-resolved fluorescence detection of telomerase activity', *Analytical Biochemistry,* 333(1), pp. 105-113.
- Garneau, J.E., Dupuis, M., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H. and Moineau, S. (2010) 'The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA', *Nature,* 468(7320), pp. 67-71.
- Genbacev, O., Krtolica, A., Zdravkovic, T., Brunette, E., Powell, S., Nath, A., Caceres, E., McMaster, M., McDonagh, S. and Li, Y. (2005) 'Serum-free derivation of human embryonic stem cell lines on human placental fibroblast feeders', *Fertility and sterility,* 83(5), pp. 1517-1529.

- Geurts, A.M., Cost, G.J., Freyvert, Y., Zeitler, B., Miller, J.C., Choi, V.M., Jenkins, S.S., Wood, A., Cui, X., Meng, X., Vincent, A., Lam, S., Michalkiewicz, M., Schilling, R., Foeckler, J., Kalloway, S., Weiler, H., Menoret, S., Anegon, I., Davis, G.D., Zhang, L., Rebar, E.J., Gregory, P.D., Urnov, F.D., Jacob, H.J. and Buelow, R. (2009) 'Knockout rats via embryo microinjection of zinc-finger nucleases', *Science (New York, N.Y.),* 325(5939), pp. 433.
- Goldman, F.D., Aubert, G., Klingelhutz, A.J., Hills, M., Cooper, S.R., Hamilton, W.S., Schlueter, A.J., Lambie, K., Eaves, C.J. and Lansdorp, P.M. (2008) 'Characterization of primitive hematopoietic cells from patients with dyskeratosis congenita', *Blood,* 111(9), pp. 4523-4531.
- Gomez, D.E., Armando, R.G., Farina, H.G., Menna, P.L., Cerrudo, C.S., Ghiringhelli, P.D. and Alonso, D.F. (2012) 'Telomere structure and telomerase in health and disease (review)', *International journal of oncology,* 41(5), pp. 1561-1569.
- Gong P, Nafabadaji Ghaderi M, Kaiser MA, Nath M, Schofield C, Samani NJ, Webb TR, unpublished data
- González, F., Zhu, Z., Shi, Z., Lelli, K., Verma, N., Li, Q.V. and Huangfu, D. (2014) 'An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells', *Cell stem cell,* 15(2), pp. 215-226.
- González, F., Zhu, Z., Shi, Z., Lelli, K., Verma, N., Li, Q.V. and Huangfu, D. (2014) 'An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells', *Cell stem cell,* 15(2), pp. 215-226.
- Gramatges, M.M., Telli, M.L., Balise, R. and Ford, J.M. (2010) 'Longer relative telomere length in blood from women with sporadic and familial breast cancer compared with healthy controls', *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology,* 19(2), pp. 605-613.
- Grandela, C. and Wolvetang, E. (2007) 'hESC adaptation, selection and stability', *Stem Cell Reviews and Reports,* 3(3), pp. 183-191.
- Greider, C.W. and Blackburn, E.H. (1985) 'Identification of a specific telomere terminal transferase activity in Tetrahymena extracts', *Cell,* 43(2), pp. 405-413.

- Greider, C.W. and Blackburn, E.H. (1987) 'The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity', *Cell,* 51(6), pp. 887-898.
- Griffith, J.D., Comeau, L., Rosenfield, S., Stansel, R.M., Bianchi, A., Moss, H. and De Lange, T. (1999) 'Mammalian telomeres end in a large duplex loop', *Cell,* 97(4), pp. 503-514.
- Grundy, S.M., Pasternak, R., Greenland, P., Smith, S. and Fuster, V. (1999) 'Assessment of cardiovascular risk by use of multiple-risk-factor assessment equationsa statement for healthcare professionals from the American Heart Association and the American College of Cardiology', *Journal of the American College of Cardiology,* 34(4), pp. 1348-1359.
- Guilinger, J.P., Thompson, D.B. and Liu, D.R. (2014) 'Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification', *Nature biotechnology,* 32(6), pp. 577-582.
- Günes, C. and Rudolph, K.L. (2013) 'The role of telomeres in stem cells and cancer', *Cell,* 152(3), pp. 390-393.
- Guo, L., Abrams, R.M., Babiarz, J.E., Cohen, J.D., Kameoka, S., Sanders, M.J., Chiao, E. and Kolaja, K.L. (2011) 'Estimating the risk of drug-induced proarrhythmia using human induced pluripotent stem cell–derived cardiomyocytes' *Toxicological Sciences*, 123(1), pp.281-289.
- Gupta, R.M., Hadaya, J., Trehan, A., Zekavat, S.M., Roselli, C., Klarin, D., Emdin, C.A., Hilvering, C.R., Bianchi, V. and Mueller, C. (2017) 'A Genetic Variant Associated with Five Vascular Diseases Is a Distal Regulator of Endothelin-1 Gene Expression', *Cell*, 170(3), pp. 522-533. e15.
- Gurdon, J.B. (1962) 'The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles', *Journal of embryology and experimental morphology,* 10, pp. 622-640.
- Gutierrez-Aranda, I., Ramos-Mejia, V., Bueno, C., Munoz-Lopez, M., Real, P., Mácia, A., Sanchez, L., Ligero, G., Garcia-Parez, J.L. and Menendez, P. (2010) 'Human induced pluripotent stem cells develop teratoma more efficiently and faster than human embryonic stem cells regardless the site of injection', *Stem cells,* 28(9), pp. 1568-1570.

- Hamm, A., Krott, N., Breibach, I., Blindt, R. and Bosserhoff, A.K. (2002) 'Efficient transfection method for primary cells', *Tissue engineering,* 8(2), pp. 235-245.
- Han, J., Qureshi, A.A., Prescott, J., Guo, Q., Ye, L., Hunter, D.J. and De Vivo, I. (2008) 'A prospective study of telomere length and the risk of skin cancer', *Journal of Investigative Dermatology,* 129(2), pp. 415-421.
- Händel, E., Alwin, S. and Cathomen, T. (2009) 'Expanding or restricting the target site repertoire of zinc-finger nucleases: the inter-domain linker as a major determinant of target site selectivity', *Molecular Therapy,* 17(1), pp. 104-111.
- Hanna, J., Markoulaki, S., Schorderet, P., Carey, B.W., Beard, C., Wernig, M., Creyghton, M.P., Steine, E.J., Cassady, J.P. and Foreman, R. (2008) 'Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency', *Cell,* 133(2), pp. 250-264.
- Hanna, J., Wernig, M., Markoulaki, S., Sun, C.W., Meissner, A., Cassady, J.P., Beard, C., Brambrink, T., Wu, L.C., Townes, T.M. and Jaenisch, R. (2007) 'Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin', *Science (New York, N.Y.),* 318(5858), pp. 1920-1923.
- Hao, L., Armanios, M., Strong, M.A., Karim, B., Feldser, D.M., Huso, D. and Greider, C.W. (2005) 'Short telomeres, even in the presence of telomerase, limit tissue renewal capacity', *Cell,* 123(6), pp. 1121-1131.
- Harley, C.B., Futcher, A.B. and Greider, C.W. (1990) 'Telomeres shorten during ageing of human fibroblasts', *Nature,* 345(6274), pp. 458.
- Harrington, L., Zhou, W., McPhail, T., Oulton, R., Yeung, D.S., Mar, V., Bass, M.B. and Robinson, M.O. (1997) 'Human telomerase contains evolutionarily conserved catalytic and structural subunits', *Genes & development,* 11(23), pp. 3109-3115.
- Hauschild, J., Petersen, B., Santiago, Y., Queisser, A.L., Carnwath, J.W., Lucas-Hahn, A., Zhang, L., Meng, X., Gregory, P.D., Schwinzer, R., Cost, G.J. and Niemann, H. (2011) 'Efficient generation of a biallelic knockout in pigs using zinc-finger nucleases', *Proceedings of the National Academy of Sciences of the United States of America,* 108(29), pp. 12013-12017.

- Hayflick, L. and Moorhead, P.S. (1961) 'The serial cultivation of human diploid cell strains', *Experimental cell research,* 25(3), pp. 585-621.
- Hegele, R.A. (2010) 'Genome-wide association studies of plasma lipids: have we reached the limit?', *Arteriosclerosis, Thrombosis, and Vascular Biology,* 30(11), pp. 2084-2086.
- Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S. and Manolio, T.A. (2009) 'Potential etiologic and functional implications of genome-wide association loci for human diseases and traits', *Proceedings of the National Academy of Sciences of the United States of America,* 106(23), pp. 9362-9367.
- Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKelver, R.C., Katibah, G.E., Amora, R., Boydston, E.A. and Zeitler, B. (2009) 'Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases', *Nature biotechnology,* 27(9), pp. 851-857.
- Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G.J., Zhang, L., Santiago, Y. and Miller, J.C. (2011) 'Genetic engineering of human pluripotent cells using TALE nucleases', *Nature biotechnology,* 29(8), pp. 731-734.
- Holditch, S.J., Terzic, A. and Ikeda, Y. (2014) 'Concise Review: Pluripotent Stem Cell-Based Regenerative Applications for Failing β-Cell Function', *Stem cells translational medicine,* 3(5), pp. 653-661.
- Horii, T., Morita, S., Kimura, M., Kobayashi, R., Tamura, D., Takahashi, R., Kimura, H., Suetake, I., Ohata, H. and Okamoto, K. (2013) 'Genome engineering of mammalian haploid embryonic stem cells using the Cas9/RNA system', *PeerJ,* 1, pp. e230.
- Hou, P., Li, Y., Zhang, X., Liu, C., Guan, J., Li, H., Zhao, T., Ye, J., Yang, W., Liu, K., Ge, J., Xu, J., Zhang, Q., Zhao, Y. and Deng, H. (2013) 'Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds', *Science (New York, N.Y.),* 341(6146), pp. 651-654.
- Hovatta, O., Mikkola, M., Gertow, K., Stromberg, A.M., Inzunza, J., Hreinsson, J., Rozell, B., Blennow, E., Andang, M. and Ahrlund-Richter, L. (2003) 'A culture system using human foreskin fibroblasts as feeder cells allows production of human embryonic stem cells', *Human reproduction (Oxford, England),* 18(7), pp. 1404-1409.

- Hsiao, L., Carr, C., Chang, K., Lin, S. and Clarke, K. (2013) 'Stem cell-based therapy for ischemic heart disease', *Cell transplantation,* 22(4), pp. 663-675.
- Hu, K., Yu, J., Suknuntha, K., Tian, S., Montgomery, K., Choi, K.D., Stewart, R., Thomson, J.A. and Slukvin, I.I. (2011) 'Efficient generation of transgene-free induced pluripotent stem cells from normal and neoplastic bone marrow and cord blood mononuclear cells', *Blood,* 117(14), pp. e109-19.
- Huang, J., Wang, F., Okuka, M., Liu, N., Ji, G., Ye, X., Zuo, B., Li, M., Liang, P. and William, W.G. (2011) 'Association of telomere length with authentic pluripotency of ES/iPS cells', *Cell research,* 21(5), pp. 779-792.
- Huang, X., Wang, Y., Yan, W., Smith, C., Ye, Z., Wang, J., Gao, Y., Mendelsohn, L. and Cheng, L. (2015) 'Production of Gene-Corrected Adult Beta Globin Protein in Human Erythrocytes Differentiated from Patient iPSCs After Genome Editing of the Sickle Point Mutation', *Stem cells,* 33(5), pp. 1470-1479.
- Huertas, P. (2010) 'DNA resection in eukaryotes: deciding how to fix the break', *Nature structural & molecular biology,* 17(1), pp. 11-16.
- Hultdin, M., Grönlund, E., Norrback, K., Eriksson-Lindström, E., Roos, G. and Just, T. (1998) 'Telomere analysis by fluorescence in situ hybridization and flow cytometry', *Nucleic acids research,* 26(16), pp. 3651-3656.
- Hunt, S.C., Chen, W., Gardner, J.P., Kimura, M., Srinivasan, S.R., Eckfeldt, J.H., Berenson, G.S. and Aviv, A. (2008) 'Leukocyte telomeres are longer in African Americans than in whites: the national heart, lung, and blood institute family heart study and the Bogalusa heart study', *Aging cell,* 7(4), pp. 451-458.
- Huttlin, E.L., Ting, L., Bruckner, R.J., Gebreab, F., Gygi, M.P., Szpyt, J., Tam, S., Zarraga, G., Colby, G. and Baltier, K. (2015) 'The BioPlex network: a systematic exploration of the human interactome', *Cell,* 162(2), pp. 425-440.
- Huzen, J., van Veldhuisen, D.J., van Gilst, W.H. and van der Harst, P. (2008) 'Telomeres and biological ageing in cardiovascular disease', *Nederlands tijdschrift voor geneeskunde,* 152(22), pp. 1265-1270.
- Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.J. and Joung, J.K. (2013) 'Efficient genome editing in

zebrafish using a CRISPR-Cas system', *Nature biotechnology,* 31(3), pp. 227-229.

- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. and Nakata, A. (1987) 'Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product', *Journal of Bacteriology,* 169(12), pp. 5429-5433.
- Israel, M.A., Yuan, S.H., Bardy, C., Reyna, S.M., Mu, Y., Herrera, C., Hefferan, M.P., Van Gorp, S., Nazor, K.L. and Boscolo, F.S. (2012) 'Probing sporadic and familial Alzheimer/'s disease using induced pluripotent stem cells', *Nature,* 482(7384), pp. 216-220.
- Jacobs, S.A., Podell, E.R. and Cech, T.R. (2006) 'Crystal structure of the essential N-terminal domain of telomerase reverse transcriptase', *Nature structural & molecular biology,* 13(3), pp. 218-225.
- Ji, J., Ng, S.H., Sharma, V., Neculai, D., Hussein, S., Sam, M., Trinh, Q., Church, G.M., Mcpherson, J.D. and Nagy, A. (2012) 'Elevated coding mutation rate during the reprogramming of human somatic cells into induced pluripotent stem cells', *Stem cells,* 30(3), pp. 435-440.
- Jia, F., Wilson, K.D., Sun, N., Gupta, D.M., Huang, M., Li, Z., Panetta, N.J., Chen, Z.Y., Robbins, R.C., Kay, M.A., Longaker, M.T. and Wu, J.C. (2010) 'A nonviral minicircle vector for deriving human iPS cells', *Nature methods,* 7(3), pp. 197-199.
- Jiang, W., Bikard, D., Cox, D., Zhang, F. and Marraffini, L.A. (2013) 'RNA-guided editing of bacterial genomes using CRISPR-Cas systems', *Nature biotechnology,* 31(3), pp. 233-239.
- Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B. and Weeks, D.P. (2013) 'Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice', *Nucleic acids research,* 41(20), pp. e188.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012) 'A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity', *Science (New York, N.Y.),* 337(6096), pp. 816-821.
- Jinek, M., East, A., Cheng, A., Lin, S., Ma, E. and Doudna, J. (2013) 'RNA-programmed genome editing in human cells', *eLife,* 2, pp. e00471.

- Kaji, K., Norrby, K., Paca, A., Mileikovsky, M., Mohseni, P. and Woltjen, K. (2009) '17-P022 Virus-free induction of pluripotency and subsequentexcision of reprogramming factors', *Mechanisms of development,* 126, pp. S277.
- Kallioniemi, A., Kallioniemi, O., Sudar, D., Rutovitz, D., Gray, J.W., Waldman, F. and Pinkel, D. (1992) 'Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors', *Science,* 258(5083), pp. 818-822.
- Keller, R.B., Gagne, K.E., Usmani, G.N., Asdourian, G.K., Williams, D.A., Hofmann, I. and Agarwal, S. (2012) 'CTC1 Mutations in a patient with dyskeratosis congenita', Pediatric blood & cancer, 59(2), pp. 311-314.
- Kiechl, S. and Willeit, J. (1999) 'The natural course of atherosclerosis. Part II: vascular remodeling. Bruneck Study Group', *Arteriosclerosis, Thrombosis, and Vascular Biology,* 19(6), pp. 1491-1498.
- Kilian, A., Bowtell, D.D., Abud, H.E., Hime, G.R., Venter, D.J., Keese, P.K., Duncan, E.L., Reddel, R.R. and Jefferson, R.A. (1997) 'Isolation of a candidate human telomerase catalytic subunit gene, which reveals complex splicing patterns in different cell types', *Human molecular genetics,* 6(12), pp. 2011-2019.
- Kim, D., Bae, S., Park, J., Kim, E., Kim, S., Yu, H.R., Hwang, J., Kim, J. and Kim, J. (2015) 'Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells', *Nature methods,* 12(3), pp. 237-243.
- Kim, D., Kim, C.H., Moon, J.I., Chung, Y.G., Chang, M.Y., Han, B.S., Ko, S., Yang, E., Cha, K.Y., Lanza, R. and Kim, K.S. (2009) 'Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins', *Cell stem cell,* 4(6), pp. 472-476.
- Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C.B., West, M.D., Ho, P.L., Coviello, G.M., Wright, W.E., Weinrich, S.L. and Shay, J.W. (1994) 'Specific association of human telomerase activity with immortal cells and cancer', *Science (New York, N.Y.),* 266(5193), pp. 2011-2015.
- Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C.B., West, M.D., Ho, P.L., Coviello, G.M., Wright, W.E., Weinrich, S.L. and Shay, J.W. (1994) 'Specific association of human telomerase activity with immortal cells and cancer', *Science,* , pp. 2011-2015.

- Kim, S., Kim, D., Cho, S.W., Kim, J. and Kim, J.S. (2014) 'Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins', *Genome research,* 24(6), pp. 1012-1019.
- Kimura, M., Cherkas, L.F., Kato, B.S., Demissie, S., Hjelmborg, J.B., Brimacombe, M., Cupples, A., Hunkin, J.L., Gardner, J.P. and Lu, X. (2008) 'Offspring's leukocyte telomere length, paternal age, and telomere elongation in sperm', *PLoS Genet,* 4(2), pp. e37.
- Kimura, M., Stone, R.C., Hunt, S.C., Skurnick, J., Lu, X., Cao, X., Harley, C.B. and Aviv, A. (2010) 'Measurement of telomere length by the Southern blot analysis of terminal restriction fragment lengths', *Nature protocols,* 5(9), pp. 1596.
- Kinoshita, T., Nagamatsu, G., Saito, S., Takubo, K., Horimoto, K. and Suda, T. (2014) 'Telomerase reverse transcriptase has an extratelomeric function in somatic cell reprogramming', *The Journal of biological chemistry,* 289(22), pp. 15776-15787.
- Kleinman, H.K. & Martin, G.R. (2005) 'Matrigel: basement membrane matrix with biological activity', *Seminars in cancer biology*Elsevier, pp. 378.
- Kleinman, H.K., McGarvey, M.L., Liotta, L.A., Robey, P.G., Tryggvason, K. and Martin, G.R. (1982) 'Isolation and characterization of type IV procollagen, laminin, and heparan sulfate proteoglycan from the EHS sarcoma', *Biochemistry,* 21(24), pp. 6188-6193.
- Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z. and Joung, J.K. (2016) 'High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects', *Nature,* 529(7587), pp. 490-495.
- Kong, D., Jin, Y., Yin, Y., Mi, H. and Shen, H. (2007) 'Real-time PCR detection of telomerase activity using specific molecular beacon probes', *Analytical and bioanalytical chemistry,* 388(3), pp. 699-709.
- Koppelstaetter, C., Jennings, P., Hochegger, K., Perco, P., Ischia, R., Karkoszka, H. and Mayer, G. (2005) 'Effect of tissue fixatives on telomere length determination by quantitative PCR', *Mechanisms of ageing and development,* 126(12), pp. 1331-1333.
- Krejci, K. and Koch, J. (1998) 'Improved detection and comparative sizing of human chromosomal telomeres in situ', *Chromosoma,* 107(3), pp. 198-203.

- Kyo, S., Takakura, M., Kanaya, T., Zhuo, W., Fujimoto, K., Nishio, Y., Orimo, A. and Inoue, M. (1999) 'Estrogen activates telomerase', *Cancer research,* 59(23), pp. 5917-5921.
- Lan, Q., Cawthon, R., Shen, M., Weinstein, S.J., Virtamo, J., Lim, U., Hosgood, H.D.,3rd, Albanes, D. and Rothman, N. (2009) 'A prospective study of telomere length measured by monochrome multiplex quantitative PCR and risk of non-Hodgkin lymphoma', *Clinical cancer research : an official journal of the American Association for Cancer Research,* 15(23), pp. 7429-7433.
- Lansdorp, P.M., Verwoerd, N.P., Van De Rijke, Frans M, Dragowska, V., Little, M., Dirks, R.W., Raap, A.K. and Tanke, H.J. (1996) 'Heterogeneity in telomere length of human chromosomes', *Human molecular genetics,* 5(5), pp. 685-691.
- Latrick, C.M. and Cech, T.R. (2010) 'POT1-TPP1 enhances telomerase processivity by slowing primer dissociation and aiding translocation', *The EMBO journal,* 29(5), pp. 924-933.
- Le Guen, T., Jullien, L., Touzot, F., Schertzer, M., Gaillard, L., Perderiset, M., Carpentier, W., Nitschke, P., Picard, C., Couillault, G., Soulier, J., Fischer, A., Callebaut, I., Jabado, N., Londono-Vallejo, A., de Villartay, J.P. and Revy, P. (2013) 'Human RTEL1 deficiency causes Hoyeraal-Hreidarsson syndrome with short telomeres and genome instability', Human molecular genetics, 22(16), pp. 3239-3249.
- Lenz, P., Bacot, S.M., Frazier-Jessen, M.R. and Feldman, G.M. (2003) 'Nucleoporation of dendritic cells: efficient gene transfer by electroporation into human monocyte-derived dendritic cells', *FEBS letters,* 538(1), pp. 149-154.
- Levy, D., Neuhausen, S.L., Hunt, S.C., Kimura, M., Hwang, S.J., Chen, W., Bis, J.C., Fitzpatrick, A.L., Smith, E., Johnson, A.D., Gardner, J.P., Srinivasan, S.R., Schork, N., Rotter, J.I., Herbig, U., Psaty, B.M., Sastrasinh, M., Murray, S.S., Vasan, R.S., Province, M.A., Glazer, N.L., Lu, X., Cao, X., Kronmal, R., Mangino, M., Soranzo, N., Spector, T.D., Berenson, G.S. and Aviv, A. (2010) 'Genome-wide association identifies OBFC1 as a locus involved in human leukocyte telomere biology', *Proceedings of the National Academy of Sciences of the United States of America,* 107(20), pp. 9293-9298.

- Li, H.L., Fujimoto, N., Sasakawa, N., Shirai, S., Ohkame, T., Sakuma, T., Tanaka, M., Amano, N., Watanabe, A. and Sakurai, H. (2015) 'Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9', *Stem cell reports,* 4(1), pp. 143-154.
- Li, J., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M. and Sheen, J. (2013) 'Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9', *Nature biotechnology,* 31(8), pp. 688-691.
- Liang, Y., Zhang, R., Zhang, S., Ji, G., Shi, P., Yang, T., Liu, F., Feng, J., Li, C. and Guo, D. (2016) 'Association of ACYP2 and TSPYL6 Genetic Polymorphisms with Risk of Ischemic Stroke in Han Chinese Population', *Molecular neurobiology,* , pp. 1-8.
- Lieber, M.R. (1999) 'The biochemistry and biological significance of nonhomologous DNA end joining: an essential repair process in multicellular eukaryotes', *Genes to Cells,* 4(2), pp. 77-85.
- Lingner, J., Hughes, T.R., Shevchenko, A., Mann, M., Lundblad, V. and Cech, T.R. (1997) 'Reverse transcriptase motifs in the catalytic subunit of telomerase', *Science (New York, N.Y.),* 276(5312), pp. 561-567.
- Liu, J., Yang, Y., Zhang, H., Zhao, S., Liu, H., Ge, N., Yang, H., Xing, J. and Chen, Z. (2011) 'Longer leukocyte telomere length predicts increased risk of hepatitis b virus-related hepatocellular carcinoma', *Cancer,* 117(18), pp. 4247-4256.
- Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F. and Young, N. (2013) 'The genotype-tissue expression (GTEx) project', *Nature genetics,* 45(6), pp. 580-585.
- Lue, N.F. (2004) 'Adding to the ends: what makes telomerase processive and how important is it?', *Bioessays,* 26(9), pp. 955-962.
- Ma, H., Zhou, Z., Wei, S., Liu, Z., Pooley, K.A., Dunning, A.M., Svenson, U., Roos, G., Hosgood III, H.D. and Shen, M. (2011) 'Shortened telomere length is associated with increased risk of cancer: a meta-analysis', *PloS one,* 6(6), pp. e20466.

- Maherali, N. and Hochedlinger, K. (2008) 'Guidelines and techniques for the generation of induced pluripotent stem cells', *Cell stem cell,* 3(6), pp. 595-605.
- Makarova, K.S., Aravind, L., Wolf, Y.I. and Koonin, E.V. (2011) 'Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems', *Biol Direct,* 6(1), pp. 38.
- Mali, P., Chou, B., Yen, J., Ye, Z., Zou, J., Dowey, S., Brodsky, R.A., Ohm, J.E., Yu, W. and Baylin, S.B. (2010) 'Butyrate greatly enhances derivation of human induced pluripotent stem cells by promoting epigenetic remodeling and the expression of pluripotency-associated genes', *Stem cells,* 28(4), pp. 713-720.
- Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M. (2013) 'RNA-guided human genome engineering via Cas9', *Science (New York, N.Y.),* 339(6121), pp. 823-826.
- Malik, N. and Rao, M.S. (2013) 'A review of the methods for human iPSC derivation', *Pluripotent Stem Cells: Methods and Protocols,* , pp. 23-33.
- Mangino, M., Richards, J.B., Soranzo, N., Zhai, G., Aviv, A., Valdes, A.M., Samani, N.J., Deloukas, P. and Spector, T.D. (2009) 'A genome-wide association study identifies a novel locus on chromosome 18q12.2 influencing white cell telomere length', *Journal of medical genetics,* 46(7), pp. 451-454.
- Marion, R.M., Strati, K., Li, H., Tejera, A., Schoeftner, S., Ortega, S., Serrano, M. and Blasco, M.A. (2009) 'Telomeres acquire embryonic stem cell characteristics in induced pluripotent stem cells', *Cell stem cell,* 4(2), pp. 141-154.
- Martí, M., Mulero, L., Pardo, C., Morera, C., Carrió, M., Laricchia-Robbio, L., Esteban, C.R. and Belmonte, J.C.I. (2013) 'Characterization of pluripotent stem cells', *Nature protocols,* 8(2), pp. 223-253.
- Martin, G.R. (1981) 'Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells', *Proceedings of the National Academy of Sciences of the United States of America,* 78(12), pp. 7634-7638.
- Mashimo, T., Takizawa, A., Voigt, B., Yoshimi, K., Hiai, H., Kuramoto, T. and Serikawa, T. (2010) 'Generation of knockout rats with X-linked severe

combined immunodeficiency (X-SCID) using zinc-finger nucleases', *PloS one,* 5(1), pp. e8870.

- Mathew, R., Jia, W., Sharma, A., Zhao, Y., Clarke, L.E., Cheng, X., Wang, H., Salli, U., Vrana, K.E., Robertson, G.P., Zhu, J. and Wang, S. (2010) 'Robust activation of the human but not mouse telomerase gene during the induction of pluripotency', *FASEB journal : official publication of the Federation of American Societies for Experimental Biology,* 24(8), pp. 2702-2715.
- Mayshar, Y., Ben-David, U., Lavon, N., Biancotti, J., Yakir, B., Clark, A.T., Plath, K., Lowry, W.E. and Benvenisty, N. (2010) 'Identification and classification of chromosomal aberrations in human induced pluripotent stem cells', *Cell stem cell,* 7(4), pp. 521-531.
- McGrath, M., Wong, J.Y., Michaud, D., Hunter, D.J. and De Vivo, I. (2007) 'Telomere length, cigarette smoking, and bladder cancer risk in men and women', *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology,* 16(4), pp. 815-819.
- Meisner, L.F. and Johnson, J.A. (2008) 'Protocols for cytogenetic studies of human embryonic stem cells', *Methods,* 45(2), pp. 133-141.
- Meng, X., Noyes, M.B., Zhu, L.J., Lawson, N.D. and Wolfe, S.A. (2008) 'Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases', *Nature biotechnology,* 26(6), pp. 695-701.
- Meyerson, M., Counter, C.M., Eaton, E.N., Ellisen, L.W., Steiner, P., Caddle, S.D., Ziaugra, L., Beijersbergen, R.L., Davidoff, M.J. and Liu, Q. (1997) '< i> hEST2</i>, the Putative Human Telomerase Catalytic Subunit Gene, Is Up-Regulated in Tumor Cells and during Immortalization', *Cell,* 90(4), pp. 785-795.
- Meyne, J., Ratliff, R.L. and Moyzis, R.K. (1989) 'Conservation of the human telomere sequence (TTAGGG)n among vertebrates', *Proceedings of the National Academy of Sciences of the United States of America,* 86(18), pp. 7049-7053.
- Miller, J., McLachlan, A.D. and Klug, A. (1985) 'Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes', *The EMBO journal,* 4(6), pp. 1609-1614.

- Mitalipova, M.M., Rao, R.R., Hoyer, D.M., Johnson, J.A., Meisner, L.F., Jones, K.L., Dalton, S. and Stice, S.L. (2005) 'Preserving the genetic integrity of human embryonic stem cells', *Nature biotechnology,* 23(1), pp. 19-20.
- Miyake, Y., Nakamura, M., Nabetani, A., Shimamura, S., Tamura, M., Yonehara, S., Saito, M. and Ishikawa, F. (2009) 'RPA-like mammalian Ctc1-Stn1-Ten1 complex binds to single-stranded DNA and protects telomeres independently of the Pot1 pathway', *Molecular cell,* 36(2), pp. 193-206.
- Miyoshi, N., Ishii, H., Nagano, H., Haraguchi, N., Dewi, D.L., Kano, Y., Nishikawa, S., Tanemura, M., Mimori, K. and Tanaka, F. (2011) 'Reprogramming of mouse and human cells to pluripotency using mature microRNAs', *Cell stem cell,* 8(6), pp. 633-638.
- Modesti, A., Raugei, G., Taddei, N., Marzocchini, R., Vecchi, M., Camici, G., Manao, G. and Ramponi, G. (1993) 'Chemical synthesis and expression of a gene coding for human muscle acylphosphatase', *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression,* 1216(3), pp. 369-374.
- Moehle, E.A., Rock, J.M., Lee, Y.L., Jouvenot, Y., DeKelver, R.C., Gregory, P.D., Urnov, F.D. and Holmes, M.C. (2007) 'Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases', *Proceedings of the National Academy of Sciences of the United States of America,* 104(9), pp. 3055-3060.
- Montpetit, A.J., Alhareeri, A.A., Montpetit, M., Starkweather, A.R., Elmore, L.W., Filler, K., Mohanraj, L., Burton, C.W., Menzies, V.S., Lyon, D.E. and Jackson-Cook, C.K. (2014) 'Telomere length: a review of methods for measurement', *Nursing research,* 63(4), pp. 289-299.
- Morin, G.B. (1989) 'The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats', *Cell,* 59(3), pp. 521-529.

- Morin, G.B. (1989) 'The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats', *Cell,* 59(3), pp. 521-529.

- Morla, M., Busquets, X., Pons, J., Sauleda, J., MacNee, W. and Agusti, A.G. (2006) 'Telomere shortening in smokers with and without COPD', *The European respiratory journal,* 27(3), pp. 525-528.

- Moscou, M.J. and Bogdanove, A.J. (2009) 'A simple cipher governs DNA recognition by TAL effectors', *Science (New York, N.Y.),* 326(5959), pp. 1501.
- Musunuru, K. and Kathiresan, S. (2010) 'Genetics of coronary artery disease', *Annual review of genomics and human genetics,* 11, pp. 91-108.
- Nakamura, M. and Okano, H. (2013) 'Cell transplantation therapies for spinal cord injury focusing on induced pluripotent stem cells', *Cell research,* 23(1), pp. 70-80.
- Nakaoka, H., Nishiyama, A., Saito, M. and Ishikawa, F. (2012) 'Xenopus laevis Ctc1-Stn1-Ten1 (xCST) protein complex is involved in priming DNA synthesis on single-stranded DNA template in Xenopus egg extract', *The Journal of biological chemistry,* 287(1), pp. 619-627.
- Nakayama, T., Fish, M.B., Fisher, M., Oomen-Hajagos, J., Thomsen, G.H. and Grainger, R.M. (2013) 'Simple and efficient CRISPR/Cas9-mediated targeted mutagenesis in Xenopus tropicalis', *genesis,* 51(12), pp. 835-843.
- Narsinh, K.H., Jia, F., Robbins, R.C., Kay, M.A., Longaker, M.T. and Wu, J.C. (2011) 'Generation of adult human induced pluripotent stem cells using nonviral minicircle DNA vectors', *Nature protocols,* 6(1), pp. 78-88.
- Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J.D. and Kamoun, S. (2013) 'Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease', *Nature biotechnology,* 31(8), pp. 691-693.
- Nelson, N.D. and Bertuch, A.A. (2012) 'Dyskeratosis congenita as a disorder of telomere maintenance', Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 730(1), pp. 43-51.
- Obokata, H., Wakayama, T., Sasai, Y., Kojima, K., Vacanti, M.P., Niwa, H., Yamato, M. and Vacanti, C.A. (2014) 'Stimulus-triggered fate conversion of somatic cells into pluripotency', *Nature,* 505(7485), pp. 641.
- O'Callaghan, N.J. and Fenech, M. (2011) 'A quantitative PCR method for measuring absolute telomere length', *Biological Procedures Online,* 13(1), pp. 3.
- O'Connor, M.S., Safari, A., Xin, H., Liu, D. and Songyang, Z. (2006) 'A critical role for TPP1 and TIN2 interaction in high-order telomeric complex assembly', *Proceedings of the National Academy of Sciences of the United States of America,* 103(32), pp. 11874-11879.

- Oeseburg, H., de Boer, R.A., van Gilst, W.H. and van der Harst, P. (2010) 'Telomere biology in healthy aging and disease', *Pflügers Archiv-European Journal of Physiology,* 459(2), pp. 259-268.
- Oeseburg, H., Iusuf, D., van der Harst, P., van Gilst, W.H., Henning, R.H. and Roks, A.J. (2009) 'Bradykinin protects against oxidative stress-induced endothelial cell senescence', *Hypertension,* 53(2), pp. 417-422.
- Okita, K., Hong, H., Takahashi, K. and Yamanaka, S. (2010) 'Generation of mouse-induced pluripotent stem cells with plasmid vectors', *Nature protocols,* 5(3), pp. 418-428.
- Okita, K., Ichisaka, T. and Yamanaka, S. (2007) 'Generation of germline-competent induced pluripotent stem cells', *Nature,* 448(7151), pp. 313-317.
- Okuda, K., Bardeguez, A., Gardner, J.P., Rodriguez, P., Ganesh, V., Kimura, M., Skurnick, J., Awad, G. and Aviv, A. (2002) 'Telomere length in the newborn', *Pediatric research,* 52(3), pp. 377-381.
- Okuda, K., Khan, M.Y., Skurnick, J., Kimura, M., Aviv, H. and Aviv, A. (2000) 'Telomere attrition of the human abdominal aorta: relationships with age and atherosclerosis', *Atherosclerosis,* 152(2), pp. 391-398.
- Olovnikov, A.M. (1971) 'Principle of marginotomy in template synthesis of polynucleotides', *Doklady Akademii nauk SSSR,* 201(6), pp. 1496-1499.
- Opresko, P.L. and Shay, J.W. (2016) 'Telomere-Associated Aging Disorders', *Ageing research reviews*.33, pp. 52-66.
- Orkin, R.W., Gehron, P., McGoodwin, E.B., Martin, G.R., Valentine, T. and Swarm, R. (1977) 'A murine tumor producing a matrix of basement membrane', *The Journal of experimental medicine,* 145(1), pp. 204-220.
- Ouellette, M.M., Liao, M., Herbert, B.S., Johnson, M., Holt, S.E., Liss, H.S., Shay, J.W. and Wright, W.E. (2000) 'Subsenescent telomere lengths in fibroblasts immortalized by limiting amounts of telomerase', *The Journal of biological chemistry,* 275(14), pp. 10072-10076.
- Ouellette, M.M., McDaniel, L.D., Wright, W.E., Shay, J.W. and Schultz, R.A. (2000) 'The establishment of telomerase-immortalized cell lines representing human chromosome instability syndromes', *Human molecular genetics,* 9(3), pp. 403-411.

- Pakzad, M., Totonchi, M., Taei, A., Seifinejad, A., Hassani, S.N. and Baharvand, H. (2010) 'Presence of a ROCK inhibitor in extracellular matrix supports more undifferentiated growth of feeder-free human embryonic and induced pluripotent stem cells upon passaging', *Stem Cell Reviews and Reports,* 6(1), pp. 96-107.
- Palm, W. and de Lange, T. (2008) 'How shelterin protects mammalian telomeres', *Annual Review of Genetics,* 42, pp. 301-334.
- Park, I., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M.W., Cowan, C., Hochedlinger, K. and Daley, G.Q. (2008) 'Disease-specific induced pluripotent stem cells', *Cell,* 134(5), pp. 877-886.
- Park, I., Zhao, R., West, J.A., Yabuuchi, A., Huo, H., Ince, T.A., Lerou, P.H., Lensch, M.W. and Daley, G.Q. (2008) 'Reprogramming of human somatic cells to pluripotency with defined factors', *Nature,* 451(7175), pp. 141-146.
- Pattanayak, V., Ramirez, C.L., Joung, J.K. and Liu, D.R. (2011) 'Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection', *Nature methods,* 8(9), pp. 765-770.
- Perez-Pinera, P., Ousterout, D.G. and Gersbach, C.A. (2012) 'Advances in targeted genome editing', *Current opinion in chemical biology,* 16(3), pp. 268-277.
- Peters, D.T., Cowan, C.A. and Musunuru, K. (2013) 'Genome editing in human pluripotent stem cells', .
- Pieler, T. and Bellefroid, E. (1994) 'Perspectives on zinc finger protein function and evolution-an update', *Molecular biology reports,* 20(1), pp. 1-8.
- Pinkel, D., Landegent, J., Collins, C., Fuscoe, J., Segraves, R., Lucas, J. and Gray, J. (1988) 'Fluorescence in situ hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4', *Proceedings of the National Academy of Sciences of the United States of America,* 85(23), pp. 9138-9142.
- Podlevsky, J.D. and Chen, J.J. (2012) 'It all comes together at the ends: telomerase structure, function, and biogenesis', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* 730(1), pp. 3-11.

- Radecke, S., Radecke, F., Cathomen, T. and Schwarz, K. (2010) 'Zinc-finger nuclease-induced gene repair with oligodeoxynucleotides: wanted and unwanted target locus modifications', *Molecular Therapy,* 18(4), pp. 743-753.
- Ramirez, R.D., Morales, C.P., Herbert, B.S., Rohde, J.M., Passons, C., Shay, J.W. and Wright, W.E. (2001) 'Putative telomere-independent mechanisms of replicative aging reflect inadequate growth conditions', *Genes & development,* 15(4), pp. 398-403.
- Ramsden, C.M., Powner, M.B., Carr, A.J., Smart, M.J., da Cruz, L. and Coffey, P.J. (2013) 'Stem cells in retinal regeneration: past, present and future', *Development (Cambridge, England),* 140(12), pp. 2576-2585.
- Ran, F.A., Hsu, P.D., Lin, C., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S. and Zhang, Y. (2013a) 'Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity', *Cell,* 154(6), pp. 1380-1389.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A. and Zhang, F. (2013b) 'Genome engineering using the CRISPR-Cas9 system', *Nature protocols,* 8(11), pp. 2281-2308.
- Rauscher, B., Heigwer, F., Breinig, M., Winter, J. and Boutros, M. (2017) 'GenomeCRISPR-a database for high-throughput CRISPR/Cas9 screens', *Nucleic acids research,* 45(D1), pp. D679-D686.
- Reubinoff, B.E., Pera, M.F., Fong, C., Trounson, A. and Bongso, A. (2000) 'Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro', *Nature biotechnology,* 18(4), pp. 399-404.
- Richards, M., Fong, C., Chan, W., Wong, P. and Bongso, A. (2002) 'Human feeders support prolonged undifferentiated growth of human inner cell masses and embryonic stem cells', *Nature biotechnology,* 20(9), pp. 933-936.
- Rode, L., Nordestgaard, B.G. and Bojesen, S.E. (2016) 'Long telomeres and cancer risk among 95 568 individuals from the general population', *International journal of epidemiology,* 45(5), pp. 1634-1643.
- Rodriguez, P., Munroe, D., Prawitt, D., Chu, L.L., Bric, E., Kim, J., Reid, L.H., Davies, C., Nakagama, H. and Loebbert, R. (1997) 'Functional characterization of human nucleosome assembly protein-2 (NAP1L4) suggests a role as a histone chaperone', *Genomics,* 44(3), pp. 253-265.

- Rog, O., Smolikov, S., Krauskopf, A. and Kupiec, M. (2005) 'The yeast VPS genes affect telomere length regulation', *Current genetics,* 47(1), pp. 18-28.
- Rojas, S.V., Kensah, G., Rotaermel, A., Baraki, H., Kutschka, I., Zweigerdt, R., Martin, U., Haverich, A., Gruh, I. and Martens, A. (2017) 'Transplantation of purified iPSC-derived cardiomyocytes in myocardial infarction', *PloS one,* 12(5), pp. e0173222.
- Ross, C.A. and Akimov, S.S. (2014) 'Human-induced pluripotent stem cells: potential for neurodegenerative diseases', *Human molecular genetics,* 23(R1), pp. R17-R26.
- Ruiz, S., Gore, A., Li, Z., Panopoulos, A.D., Montserrat, N., Fung, H., Giorgetti, A., Bilic, J., Batchelder, E.M. and Zaehres, H. (2013) 'Analysis of protein-coding mutations in hiPSCs and their possible role during somatic cell reprogramming', *Nature communications,* 4, pp. 1382.
- Sachamitr, P., Hackett, S. and Fairchild, P.J. (2014) 'Induced pluripotent stem cells: challenges and opportunities for cancer immunotherapy', *Frontiers in immunology,* 5, pp. 176.
- Pashos, E.E., Park, Y., Wang, X., Raghavan, A., Yang, W., Abbey, D., Peters, D.T., Arbelaez, J., Hernandez, M. and Kuperwasser, N. (2017) 'Large, diverse population cohorts of hiPSCs and derived hepatocyte-like cells reveal functional genetic variation at blood lipid-associated loci', *Cell Stem Cell*, 20(4), pp. 558-570. e10.
- Samani, N.J., Boultby, R., Butler, R., Thompson, J.R. and Goodall, A.H. (2001) 'Telomere shortening in atherosclerosis', *The Lancet,* 358(9280), pp. 472-473.
- Samassekou, O., Gadji, M., Drouin, R. and Yan, J. (2010) 'Sizing the ends: normal length of human telomeres', *Annals of Anatomy-Anatomischer Anzeiger,* 192(5), pp. 284-291.
- Sander, J.D. and Joung, J.K. (2014) 'CRISPR-Cas systems for editing, regulating and targeting genomes', *Nature biotechnology,* .
- Sandin, S. and Rhodes, D. (2014) 'Telomerase structure', *Current opinion in structural biology,* 25, pp. 104-110.
- Schneuwly, S. and Klemenz, R. (1987) 'Redesigning the body plan of Drosophila by ectopic expression', *Nature,* 325, pp. 26.

- Schwarz, B.A., Bar-Nur, O., Silva, J.C. and Hochedlinger, K. (2014) 'Nanog is dispensable for the generation of induced pluripotent stem cells', *Current Biology,* 24(3), pp. 347-350.
- Seki, T., Yuasa, S., Oda, M., Egashira, T., Yae, K., Kusumoto, D., Nakata, H., Tohyama, S., Hashimoto, H., Kodaira, M., Okada, Y., Seimiya, H., Fusaki, N., Hasegawa, M. and Fukuda, K. (2010) 'Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells', *Cell stem cell,* 7(1), pp. 11-14.
- Serrano, A.L. and Andres, V. (2004) 'Telomeres and cardiovascular disease: does size matter?', *Circulation research,* 94(5), pp. 575-584.
- Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K., Liu, J., Xi, J.J. and Qiu, J. (2013) 'Targeted genome modification of crop plants using a CRISPR-Cas system', *Nature biotechnology,* 31(8), pp. 686-688.
- Shay, J. and Bacchetti, S. (1997) 'A survey of telomerase activity in human cancer', *European journal of cancer,* 33(5), pp. 787-791.
- Shay, J.W. (2010) 'Telomerase as a target for cancer therapeutics', in Anonymous *Gene-Based Therapies for Cancer.* Springer, pp. 231-249.
- Shay, J.W. and Wright, W.E. (1996) 'Telomerase activity in human cancer', *Current opinion in oncology,* 8(1), pp. 66-71.
- Shelton, D.N., Chang, E., Whittier, P.S., Choi, D. and Funk, W.D. (1999) 'Microarray analysis of replicative senescence', *Current biology,* 9(17), pp. 939-945.
- Shen, J., Gammon, M.D., Terry, M.B., Wang, Q., Bradshaw, P., Teitelbaum, S.L., Neugut, A.I. and Santella, R.M. (2009) 'Telomere length, oxidative damage, antioxidants and breast cancer risk', *International Journal of Cancer,* 124(7), pp. 1637-1643.
- Shen, M., Cawthon, R., Rothman, N., Weinstein, S.J., Virtamo, J., Hosgood III, H.D., Hu, W., Lim, U., Albanes, D. and Lan, Q. (2011) 'A prospective study of telomere length measured by monochrome multiplex quantitative PCR and risk of lung cancer', *Lung Cancer,* 73(2), pp. 133-137.
- Simón, C., Escobedo, C., Valbuena, D., Genbacev, O., Galan, A., Krtolica, A., Asensi, A., Sánchez, E., Esplugues, J. and Fisher, S. (2005) 'First derivation in Spain of human embryonic stem cell lines: use of long-term cryopreserved

embryos and animal-free conditions', *Fertility and sterility,* 83(1), pp. 246-249.

- Singh, U., Quintanilla, R.H., Grecian, S., Gee, K.R., Rao, M.S. and Lakshmipathy, U. (2012) 'Novel live alkaline phosphatase substrate for identification of pluripotent stem cells', *Stem Cell Reviews and Reports,* 8(3), pp. 1021-1029.
- Slagboom, P.E., Droog, S. and Boomsma, D.I. (1994) 'Genetic determination of telomere size in humans: a twin study of three age groups', *American Journal of Human Genetics,* 55(5), pp. 876-882.
- Smith, C., Gore, A., Yan, W., Abalde-Atristain, L., Li, Z., He, C., Wang, Y., Brodsky, R.A., Zhang, K., Cheng, L. and Ye, Z. (2014) 'Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs', *Cell stem cell,* 15(1), pp. 12-13.
- Smith, E.N., Bloss, C.S., Badner, J.A., Barrett, T., Belmonte, P.L., Berrettini, W., Byerley, W., Coryell, W., Craig, D. and Edenberg, H.J. (2009) 'Genome-wide association study of bipolar disorder in European American and African American individuals', *Molecular psychiatry,* 14(8), pp. 755-763.
- Soldner, F., Laganière, J., Cheng, A.W., Hockemeyer, D., Gao, Q., Alagappan, R., Khurana, V., Golbe, L.I., Myers, R.H. and Lindquist, S. (2011) 'Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations', *Cell,* 146(2), pp. 318-331.
- Somers, A., Jean, J., Sommer, C.A., Omari, A., Ford, C.C., Mills, J.A., Ying, L., Sommer, A.G., Jean, J.M. and Smith, B.W. (2010) 'Generation of transgene-free lung disease-specific human induced pluripotent stem cells using a single excisable lentiviral stem cell cassette', *Stem cells,* 28(10), pp. 1728-1740.
- Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochedlinger, K., Kotton, D.N. and Mostoslavsky, G. (2009) 'Induced pluripotent stem cell generation using a single lentiviral stem cell cassette', *Stem cells,* 27(3), pp. 543-549.
- Song, Z., Von Figura, G., Liu, Y., Kraus, J.M., Torrice, C., Dillon, P., Rudolph-Watabe, M., Ju, Z., Kestler, H.A. and Sanoff, H. (2010) 'Lifestyle impacts on the aging-associated expression of biomarkers of DNA damage and telomere dysfunction in human blood', *Aging cell,* 9(4), pp. 607-615.

- Stadtfeld, M., Brennand, K. and Hochedlinger, K. (2008) 'Reprogramming of pancreatic β cells into induced pluripotent stem cells', *Current Biology,* 18(12), pp. 890-894.
- Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G. and Hochedlinger, K. (2008) 'Induced pluripotent stem cells generated without viral integration', *Science (New York, N.Y.),* 322(5903), pp. 945-949.
- Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G. and Hochedlinger, K. (2008) 'Induced pluripotent stem cells generated without viral integration', *Science (New York, N.Y.),* 322(5903), pp. 945-949.
- Stefani, M. and Ramponi, G. (1995) 'Acylphosphate phosphohydrolases', *Life Chem.Rep,* 12, pp. 271-301.
- Stewart, S.A. and Weinberg, R.A. (2006) 'Telomeres: cancer to human aging', *Annu.Rev.Cell Dev.Biol.,* 22, pp. 531-557.
- Sugiura, M., Kasama, Y., Araki, R., Hoki, Y., Sunayama, M., Uda, M., Nakamura, M., Ando, S. and Abe, M. (2014) 'Induced pluripotent stem cell generation-associated point mutations arise during the initial stages of the conversion of these cells', *Stem cell reports,* 2(1), pp. 52-63.
- Suhr, S.T., Chang, E.A., Rodriguez, R.M., Wang, K., Ross, P.J., Beyhan, Z., Murthy, S. and Cibelli, J.B. (2009) 'Telomere dynamics in human cells reprogrammed to pluripotency', *PloS one,* 4(12), pp. e8124.
- Sun, N. and Zhao, H. (2013) 'Transcription activator-like effector nucleases (TALENs): A highly efficient and versatile tool for genome editing', *Biotechnology and bioengineering,* 110(7), pp. 1811-1821.
- Sun, N., Abil, Z. and Zhao, H. (2012) 'Recent advances in targeted genome engineering in mammalian systems', *Biotechnology journal,* 7(9), pp. 1074-1087.
- Surovtseva, Y.V., Churikov, D., Boltz, K.A., Song, X., Lamb, J.C., Warrington, R., Leehy, K., Heacock, M., Price, C.M. and Shippen, D.E. (2009) 'Conserved telomere maintenance component 1 interacts with STN1 and maintains chromosome ends in higher eukaryotes', *Molecular cell,* 36(2), pp. 207-218.
- Svenson, U., Ljungberg, B. and Roos, G. (2009) 'Telomere length in peripheral blood predicts survival in clear cell renal cell carcinoma', *Cancer research,* 69(7), pp. 2896-2901.

- Svenson, U., Nordfjall, K., Stegmayr, B., Manjer, J., Nilsson, P., Tavelin, B., Henriksson, R., Lenner, P. and Roos, G. (2008) 'Breast cancer survival is associated with telomere length in peripheral blood cells', *Cancer research,* 68(10), pp. 3618-3623.
- Szilard, R.K. and Durocher, D. (2006) 'Telomere protection: an act of God', *Current biology,* 16(14), pp. R544-R546.
- Tada, M., Takahama, Y., Abe, K., Nakatsuji, N. and Tada, T. (2001) 'Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells', *Current Biology,* 11(19), pp. 1553-1558.
- Takahashi, K. and Yamanaka, S. (2006) 'Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors', *Cell,* 126(4), pp. 663-676.
- 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(5), pp. 861-872.
- Takubo, K., Izumiyama-Shimomura, N., Honma, N., Sawabe, M., Arai, T., Kato, M., Oshimura, M. and Nakamura, K. (2002) 'Telomere lengths are characteristic in each human individual', *Experimental gerontology,* 37(4), pp. 523-531.
- Tao, K.P., Fong, S.W., Lu, Z., Ching, Y.P., Chan, K.W. and Chan, S.Y. (2011) 'TSPYL2 is important for G1 checkpoint maintenance upon DNA damage', *PLoS One,* 6(6), pp. e21602.
- Tecirlioglu, R.T., Nguyen, L., Koh, K., Trounson, A.O. and Michalska, A.E. (2010) 'Derivation and maintenance of human embryonic stem cell line on human adult skin fibroblast feeder cells in serum replacement medium', *In Vitro Cellular & Developmental Biology-Animal,* 46(3-4), pp. 231-235.
- Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S. and Jones, J.M. (1998) 'Embryonic stem cell lines derived from human blastocysts', *Science (New York, N.Y.),* 282(5391), pp. 1145-1147.
- Tong, M., Lv, Z., Liu, L., Zhu, H., Zheng, Q.Y., Zhao, X.Y., Li, W., Wu, Y.B., Zhang, H.J., Wu, H.J., Li, Z.K., Zeng, F., Wang, L., Wang, X.J., Sha, J.H. and Zhou, Q. (2011) 'Mice generated from tetraploid complementation competent

iPS cells show similar developmental features as those from ES cells but are prone to tumorigenesis', *Cell research,* 21(11), pp. 1634-1637.

- Tsai, S.Q., Wyvekens, N., Khayter, C., Foden, J.A., Thapar, V., Reyon, D., Goodwin, M.J., Aryee, M.J. and Joung, J.K. (2014) 'Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing', *Nature biotechnology,* 32(6), pp. 569-576.
- Tsakiri, K.D., Cronkhite, J.T., Kuan, P.J., Xing, C., Raghu, G., Weissler, J.C., Rosenblatt, R.L., Shay, J.W. and Garcia, C.K. (2007) 'Adult-onset pulmonary fibrosis caused by mutations in telomerase', *Proceedings of the National Academy of Sciences of the United States of America,* 104(18), pp. 7552-7557.
- Tsang, K.H., Lai, S.K., Li, Q., Yung, W.H., Liu, H., Mak, P.H.S., Ng, C.C.P., McAlonan, G., Chan, Y.S. and Chan, S.Y. (2014) 'The nucleosome assembly protein TSPYL2 regulates the expression of NMDA receptor subunits GluN2A and GluN2B', *Scientific reports,* 4, pp. 3654.
- Unryn, B.M., Cook, L.S. and Riabowol, K.T. (2005) 'Paternal age is positively linked to telomere length of children', *Aging cell,* 4(2), pp. 97-101.
- Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S. and Gregory, P.D. (2010) 'Genome editing with engineered zinc finger nucleases', *Nature Reviews Genetics,* 11(9), pp. 636-646.
- Vasa-Nicotera, M., Brouilette, S., Mangino, M., Thompson, J.R., Braund, P., Clemitson, J., Mason, A., Bodycote, C.L., Raleigh, S.M. and Louis, E. (2005) 'Mapping of a major locus that determines telomere length in humans', *The American Journal of Human Genetics,* 76(1), pp. 147-151.
- Vera, E. and Blasco, M.A. (2012) 'Beyond average: potential for measurement of short telomeres', *Aging,* 4(6), pp. 379-392.
- Veres, A., Gosis, B.S., Ding, Q., Collins, R., Ragavendran, A., Brand, H., Erdin, S., Cowan, C.A., Talkowski, M.E. and Musunuru, K. (2014) 'Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing', *Cell stem cell,* 15(1), pp. 27-30.
- Villa-Diaz, L.G., Garcia-Perez, J.L. and Krebsbach, P.H. (2010) 'Enhanced transfection efficiency of human embryonic stem cells by the incorporation of

DNA liposomes in extracellular matrix', *Stem cells and development,* 19(12), pp. 1949-1957.

- Vinagre, J., Almeida, A., Pópulo, H., Batista, R., Lyra, J., Pinto, V., Coelho, R., Celestino, R., Prazeres, H. and Lima, L. (2013) 'Frequency of TERT promoter mutations in human cancers', *Nature communications,* 4.
- Vinci, G., Brauner, R., Tar, A., Rouba, H., Sheth, J., Sheth, F., Ravel, C., McElreavey, K. and Bashamboo, A. (2009) 'Mutations in the TSPYL1 gene associated with 46, XY disorder of sex development and male infertility', *Fertility and sterility,* 92(4), pp. 1347-1350.
- Von Zglinicki, T. (2002) 'Oxidative stress shortens telomeres', *Trends in biochemical sciences,* 27(7), pp. 339-344.
- Vulliamy, T., Beswick, R., Kirwan, M., Marrone, A., Digweed, M., Walne, A. and Dokal, I. (2008) 'Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita', Proceedings of the National Academy of Sciences of the United States of America, 105(23), pp. 8073-8078.
- Wakayama, T., Perry, A.C., Zuccotti, M., Johnson, K.R. and Yanagimachi, R. (1998) 'Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei', *Nature,* 394(6691), pp. 369.
- Walne, A.J. and Dokal, I. (2009) 'Advances in the understanding of dyskeratosis congenita', *British journal of haematology,* 145(2), pp. 164-172.
- Walne, A.J., Bhagat, T., Kirwan, M., Gitiaux, C., Desguerre, I., Leonard, N., Nogales, E., Vulliamy, T. and Dokal, I.S. (2013b) 'Mutations in the telomere capping complex in bone marrow failure and related syndromes', Haematologica, 98(3), pp. 334-338.
- Walne, A.J., Vulliamy, T., Kirwan, M., Plagnol, V. and Dokal, I. (2013a) 'Constitutional mutations in RTEL1 cause severe dyskeratosis congenita', The American Journal of Human Genetics, 92(3), pp. 448-453.
- Walne, A.J., Vulliamy, T., Marrone, A., Beswick, R., Kirwan, M., Masunari, Y., Al-Qurashi, F.H., Aljurf, M. and Dokal, I. (2007) 'Genetic heterogeneity in autosomal recessive dyskeratosis congenita with one subtype due to mutations in the telomerase-associated protein NOP10', Human molecular genetics, 16(13), pp. 1619-1629.

- Wang, C. and Meier, U.T. (2004) 'Architecture and assembly of mammalian H/ACA small nucleolar and telomerase ribonucleoproteins', *The EMBO journal,* 23(8), pp. 1857-1867.
- Wang, D.G., Fan, J.B., Siao, C.J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N., Winchester, E., Spencer, J., Kruglyak, L., Stein, L., Hsie, L., Topaloglou, T., Hubbell, E., Robinson, E., Mittmann, M., Morris, M.S., Shen, N., Kilburn, D., Rioux, J., Nusbaum, C., Rozen, S., Hudson, T.J., Lipshutz, R., Chee, M. and Lander, E.S. (1998) 'Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome', *Science (New York, N.Y.),* 280(5366), pp. 1077-1082.
- Wang, F., Podell, E.R., Zaug, A.J., Yang, Y., Baciu, P., Cech, T.R. and Lei, M. (2007) 'The POT1–TPP1 telomere complex is a telomerase processivity factor', *Nature,* 445(7127), pp. 506-510.
- Wang, F., Yin, Y., Ye, X., Liu, K., Zhu, H., Wang, L., Chiourea, M., Okuka, M., Ji, G. and Dan, J. (2012) 'Molecular insights into the heterogeneity of telomere reprogramming in induced pluripotent stem cells', *Cell research,* 22(4), pp. 757-768.
- Wang, H., Naghavi, M., Allen, C., Barber, R.M., Bhutta, Z.A., Carter, A., Casey, D.C., Charlson, F.J., Chen, A.Z. and Coates, M.M. (2016) 'Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015', *The Lancet,* 388(10053), pp. 1459-1544.
- Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F. and Jaenisch, R. (2013) 'One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering', *Cell,* 153(4), pp. 910-918.
- Wang, L., Ma, F., Tang, B. and Zhang, C. (2017) 'Sensing telomerase: From in vitro detection to in vivo imaging', *Chemical Science,* 8(4), pp. 2495-2502.
- Wang, P., Lin, M., Pedrosa, E., Hrabovsky, A., Zhang, Z., Guo, W., Lachman, H.M. and Zheng, D. (2015) 'CRISPR/Cas9-mediated heterozygous knockout of the autism gene CHD8 and characterization of its transcriptional networks in neurodevelopment', *Molecular autism,* 6(1), pp. 55.
- Wang, X., Wang, Y., Wu, X., Wang, J., Wang, Y., Qiu, Z., Chang, T., Huang, H., Lin, R. and Yee, J. (2015) 'Unbiased detection of off-target cleavage by

CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors', *Nature biotechnology,* 33(2), pp. 175-178.

- Warren, L., Manos, P.D., Ahfeldt, T., Loh, Y., Li, H., Lau, F., Ebina, W., Mandal, P.K., Smith, Z.D. and Meissner, A. (2010) 'Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA', *Cell stem cell,* 7(5), pp. 618-630.
- Werner, C., Furster, T., Widmann, T., Poss, J., Roggia, C., Hanhoun, M., Scharhag, J., Buchner, N., Meyer, T., Kindermann, W., Haendeler, J., Bohm, M. and Laufs, U. (2009) 'Physical exercise prevents cellular senescence in circulating leukocytes and in the vessel wall', *Circulation,* 120(24), pp. 2438-2447.
- Whyte, J.J., Zhao, J., Wells, K.D., Samuel, M.S., Whitworth, K.M., Walters, E.M., Laughlin, M.H. and Prather, R.S. (2011) 'Gene targeting with zinc finger nucleases to produce cloned eGFP knockout pigs', *Molecular reproduction and development,* 78(1), pp. 2-2.
- Wiedenheft, B., Sternberg, S.H. and Doudna, J.A. (2012) 'RNA-guided genetic silencing systems in bacteria and archaea', *Nature,* 482(7385), pp. 331-338.
- Wilmut, I., Schnieke, A., McWhir, J., Kind, A. and Campbell, K. (2007) 'Viable offspring derived from fetal and adult mammalian cells', *Cloning and stem cells,* 9(1), pp. 3-7.
- Wilson, W.R., Herbert, K.E., Mistry, Y., Stevens, S.E., Patel, H.R., Hastings, R.A., Thompson, M.M. and Williams, B. (2008) 'Blood leucocyte telomere DNA content predicts vascular telomere DNA content in humans with and without vascular disease', *European heart journal,* 29(21), pp. 2689-2694.
- Winkler, T., Hong, S.G., Decker, J.E., Morgan, M.J., Wu, C., Hughes, W.M., Yang, Y., Wangsa, D., Padilla-Nash, H.M. and Ried, T. (2013) 'Defective telomere elongation and hematopoiesis from telomerase-mutant aplastic anemia iPSCs', *The Journal of clinical investigation,* 123(5), pp. 1952-1963.
- Wolfe, S.A., Nekludova, L. and Pabo, C.O. (2000) 'DNA recognition by Cys2His2 zinc finger proteins', *Annual Review of Biophysics and Biomolecular Structure,* 29(1), pp. 183-212.
- Woltjen, K., Michael, I.P., Mohseni, P., Desai, R., Mileikovsky, M., Hämäläinen, R., Cowling, R., Wang, W., Liu, P. and Gertsenstein, M. (2009) 'piggyBac

transposition reprograms fibroblasts to induced pluripotent stem cells', *Nature,* 458(7239), pp. 766-770.

- Wright, W.E., Piatyszek, M.A., Rainey, W.E., Byrd, W. and Shay, J.W. (1996) 'Telomerase activity in human germline and embryonic tissues and cells', *Developmental genetics,* 18(2), pp. 173-179.
- Wright, W.E., Tesmer, V.M., Huffman, K.E., Levene, S.D. and Shay, J.W. (1997) 'Normal human chromosomes have long G-rich telomeric overhangs at one end', *Genes & development,* 11(21), pp. 2801-2809.
- Wu, X., Amos, C.I., Zhu, Y., Zhao, H., Grossman, B.H., Shay, J.W., Luo, S., Hong, W.K. and Spitz, M.R. (2003) 'Telomere dysfunction: a potential cancer predisposition factor', *Journal of the National Cancer Institute,* 95(16), pp. 1211-1218.
- Xiao, A., Wang, Z., Hu, Y., Wu, Y., Luo, Z., Yang, Z., Zu, Y., Li, W., Huang, P., Tong, X., Zhu, Z., Lin, S. and Zhang, B. (2013) 'Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish', *Nucleic acids research,* 41(14), pp. e141.
- Xiao, Y., Dane, K.Y., Uzawa, T., Csordas, A., Qian, J., Soh, H.T., Daugherty, P.S., Lagally, E.T., Heeger, A.J. and Plaxco, K.W. (2010) 'Detection of telomerase activity in high concentration of cell lysates using primer-modified gold nanoparticles', *Journal of the American Chemical Society,* 132(43), pp. 15299-15307.
- Xie, F., Ye, L., Chang, J.C., Beyer, A.I., Wang, J., Muench, M.O. and Kan, Y.W. (2014) 'Seamless gene correction of beta-thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac', *Genome research,* 24(9), pp. 1526-1533.
- Xie, K. and Yang, Y. (2013) 'RNA-guided genome editing in plants using a CRISPR-Cas system', *Molecular plant,* 6(6), pp. 1975-1983.
- Xu, C. (2006) 'Characterization and evaluation of human embryonic stem cells', *Methods in enzymology,* 420, pp. 18-37.
- Yagi, T., Ito, D., Okada, Y., Akamatsu, W., Nihei, Y., Yoshizaki, T., Yamanaka, S., Okano, H. and Suzuki, N. (2011) 'Modeling familial Alzheimer's disease with induced pluripotent stem cells', *Human molecular genetics,* 20(23), pp. 4530-4539.

- Yahata, N., Asai, M., Kitaoka, S., Takahashi, K., Asaka, I. and Okazawa, H. (2011) 'Anti-Ab Drug Screening Platform Using Human iPS Cell-Derived Neurons for the Treatment',
- Yamanaka, S. (2012) 'Induced pluripotent stem cells: past, present, and future', *Cell stem cell,* 10(6), pp. 678-684.
- Yang, C., Przyborski, S., Cooke, M.J., Zhang, X., Stewart, R., Anyfantis, G., Atkinson, S.P., Saretzki, G., Armstrong, L. and Lako, M. (2008) 'A key role for telomerase reverse transcriptase unit in modulating human embryonic stem cell proliferation, cell cycle dynamics, and in vitro differentiation', *Stem cells,* 26(4), pp. 850-863.
- Yang, D., Xu, J., Zhu, T., Fan, J., Lai, L., Zhang, J. and Chen, Y.E. (2014) 'Effective gene targeting in rabbits using RNA-guided Cas9 nucleases', *Journal of molecular cell biology,* 6(1), pp. 97-99.
- Yang, J., Chang, E., Cherry, A.M., Bangs, C.D., Oei, Y., Bodnar, A., Bronstein, A., Chiu, C.P. and Herron, G.S. (1999) 'Human endothelial cell life extension by telomerase expression', *The Journal of biological chemistry,* 274(37), pp. 26141-26148.
- Yehezkel, S., Rebibo-Sabbah, A., Segev, Y., Tzukerman, M., Shaked, R., Huber, I., Gepstein, L., Skorecki, K. and Selig, S. (2011) 'Reprogramming of telomeric regions during the generation of human induced pluripotent stem cells and subsequent differentiation into fibroblast-like derivatives', *Epigenetics: official journal of the DNA Methylation Society,* 6(1), pp. 63-75.
- Yoshihara, M., Hayashizaki, Y. and Murakawa, Y. (2017) 'Genomic instability of iPSCs: challenges towards their clinical applications', *Stem Cell Reviews and Reports,* 13(1), pp. 7-16.
- Yu, C., Liu, Y., Ma, T., Liu, K., Xu, S., Zhang, Y., Liu, H., La Russa, M., Xie, M. and Ding, S. (2015) 'Small molecules enhance CRISPR genome editing in pluripotent stem cells', *Cell stem cell,* 16(2), pp. 142-147.
- Yu, J., Hu, K., Smuga-Otto, K., Tian, S., Stewart, R., Slukvin, I.I. and Thomson, J.A. (2009) 'Human induced pluripotent stem cells free of vector and transgene sequences', *Science (New York, N.Y.),* 324(5928), pp. 797-801.

- 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', *Science (New York, N.Y.),* 318(5858), pp. 1917-1920.
- 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', *Science (New York, N.Y.),* 318(5858), pp. 1917-1920.
- Yumlu, S., Stumm, J., Bashir, S., Dreyer, A., Lisowski, P., Danner, E. and Kühn, R. (2017) 'Gene editing and clonal isolation of human induced pluripotent stem cells using CRISPR/Cas9', *Methods,* .
- Yunis, J.J. (1976) 'High resolution of human chromosomes', *Science (New York, N.Y.),* 191(4233), pp. 1268-1270.
- Zalzman, M., Falco, G., Sharova, L.V., Nishiyama, A., Thomas, M., Lee, S., Stagg, C.A., Hoang, H.G., Yang, H. and Indig, F.E. (2010) 'Zscan4 regulates telomere elongation and genomic stability in ES cells', *Nature,* 464(7290), pp. 858-863.
- Zeggini, E., Weedon, M.N., Lindgren, C.M., Frayling, T.M., Elliott, K.S., Lango, H., Timpson, N.J., Perry, J., Rayner, N.W. and Freathy, R.M. (2007) 'Wellcome Trust Case Control Consortium (WTCCC), McCarthy MI, Hattersley AT: Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes', *Science,* 316(5829), pp. 1336-1341.
- Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J. and Regev, A. (2015) 'Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system', *Cell,* 163(3), pp. 759-771.
- Zhang, X., Lou, X. and Xia, F. (2017) 'Advances in the detection of telomerase activity using isothermal amplification', *Theranostics,* 7(7), pp. 1847.
- Zhou, H., Wu, S., Joo, J.Y., Zhu, S., Han, D.W., Lin, T., Trauger, S., Bien, G., Yao, S., Zhu, Y., Siuzdak, G., Scholer, H.R., Duan, L. and Ding, S. (2009) 'Generation of induced pluripotent stem cells using recombinant proteins', *Cell stem cell,* 4(5), pp. 381-384.

- Zhou, W. and Freed, C.R. (2009) 'Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells', *Stem cells,* 27(11), pp. 2667-2674.
- Zhou, X. and Xing, D. (2012) 'Assays for human telomerase activity: progress and prospects', *Chemical Society Reviews,* 41(13), pp. 4643-4656.
- Zou, J., Maeder, M.L., Mali, P., Pruett-Miller, S.M., Thibodeau-Beganny, S., Chou, B., Chen, G., Ye, Z., Park, I. and Daley, G.Q. (2009) 'Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells', *Cell Stem Cell,* 5(1), pp. 97-110.
- Zvereva, M., Shcherbakova, D. and Dontsova, O. (2010) 'Telomerase: structure, functions, and activity regulation', *Biochemistry (Moscow),* 75(13), pp. 1563-1583.