Developing a platform for the study of multifactorial diseases

by

Stian Eide Ingebrigtsen

Thesis
for the degree of Master in Molecular Biosciences
Main field
Neurobiology

Program for Molecular Biology
Department of Biosciences
Faculty of Mathematics and Natural Sciences

UNIVERSITY OF OSLO

November, 2017

-

Developing a platform for the study of multifactorial diseases
http://www.duo.uio.no/
Printed: Reprosentralen, University of Oslo

Acknowledgements

The work presented in the thesis at hand was carried out in the Hafting-Fyhn laboratory in the department of physiology, which is organized under the Institute of Biosciences at the University of Oslo.

I would like to thank for a splendid opportunity to take my masters degree in this group; with practically no leashes I have done what I thought right for the project. For better or worse it has led me to learn a lot, and to be self-dependent. I will want to thank my supervisor Marianne Fyhn for enthusiasm, and sharing with me personal things as well as discussing science. Thanks also to Torkel Hafting for some nice running sessions at Vollsløkka. And thanks to Andy Edwards for being the intelligent and humble person he is, whom I have come to admire very much!

As the everyday contact is devoted to the "Fyhn-family", I must also thank my sisters at the office Mina Martine Frey and Tovy Dinh for keeping up with me and staying together throughout this. Especially thanks to Tovy for grammar-checking my work in the extremly bussy period towards the deadline. It has been good times being part of the Fyhn-family, thanks to all of you guys.

A degree in science is not simply acquired by the help from supervisors. The importance of having peers working through syllabus together, but with an interest not bound by that syllabus has taught me to appreciate aspects of life far beyond this. Some knowledge cannot be learnt by adding more lessons, it needs to be grown and delicately tended to by different people. Those people have provided friendships I'll be grateful for till the end. Hallvard Heiberg, Magnus Osnes, Åsta Dala, Ida Schjerven, Harald Dahl, Lisa Lunde, Per Christian Olsen (aka Pax), thank you!

The achievment of this kind of work, for me, also depends on how my life plays out in other aspects. Playing music with my good friends can relieve stress any day, thank you Kyrre Grøtan and Hallvard Heiberg!

And then of course family, what would we do without? Always supportive and more than eager to hear what is happening in the field of science and in my life in general, leading to valuable discussions and bonding now that the everyday contact is no longer a matter of course. Thank you beautiful people, Mads, Line, Mamma and Pappa!

And then – Benedikte – the centerpiece of my world; having you with me in all my ups and downs is doing so much more than I can manage to describe here. Patiently waiting for the night-and-day-work to pass so that we, finally, can spend some normal time together again. That being more important now than ever, as life grows in your belly and we become three! I Love you.

Contents

List of abbreviations

ANOVA Analysis of Variance

AP Action Potential

ATCC American type culture collection

Cas9 CRISPR-associated Protein 9

CDKN2D Cyclin Dependent Kinase Inhibitor 2D

cDNA complimentary DNA

CGM Complete Growth Medium

CRISPR Clustered Regularly Interspaced Short Palindromic Repeats

crRNA CRISPR RNA

Ct Cycle threshold

DMSO Dimethyl Sulphoxide

dNTP Dideoxynucleotide Triphosphate

DSB Double Stranded Brakes

EDTA EthyleneDiamineTetraacetic Acid

GAPDH Glyceraldehyde 3-Phosphate Dehydrogenase

GFP Green Fluorescent Protein

GWAS Genome Wide Association Study

HDR Homology Directed Repair

HPRT Hypoxanthine-Guanine Phosphoribosyltransferase

InDels Insertions/Deletions

iPSC induced Pluripotent Stem Cells

LD Linkage Disequilibrium

MFD MultiFactorial Disease

NaOAC Sodium Acetate

NCBI National Center for Biotechnology Information

NEB New England Biolab

NHEJ Non-Homologues End Joining

NT Nucleotides

PAM Protospacer Adjacent Motif

PBS Phosphate-Buffered Saline

PCR Polymerase Chain Reaction

qRT PCR quantitative Reverse Transcription Polymerase Chain Reaction

RGEN RNA-Guided Endonucleases

RNP Ribonucleo Protein

RT Reverse Transctription

SCZ Schizophrenia

SDSA Synthesis Dependent Strand Annealing

sgRNA single guide RNA

SNP Single Nucleotide Polymorphism

T7E1 T7 Endonuclease 1

TALEN Transcription Activatior-Like Effector Nucleases

tracrRNA trans-activating CRISPR RNA

WT Wild Type

ZFN Zinc Finger Nucleases

1 Introduction

Brain related diseases affect up to 1/3 of the population and include neurological disorders, mental illnesses and brain injuries including tumors and stroke. Mental disorders or mental illness, affecting more than 500 mill people world wide (WHO ref, 2014), are mental patterns causing reduced ability to function in life, and may be devastating for the individual and represent huge societal challenges. The underlying mechanisms of mental illness is at large poorly understood resulting in poor treatment often with severe side effects. It was early recognized that mental disorders such as schizophrenia (SCZ) had high heritability [1, 2], but traditional genetic linkage studies proved hard to reproduce these complex diseases compared to single gene disorders. However, the immense development in high-throughput sequencing after the mapping of the human genome has opened novel opportunities to identify combinations of gene variants that together may contribute to the phenotype of polygenic diseases such as SCZ.

Sequencing data has been pooled together in huge studies. In these genome wide association studies (GWAS) the data is used to see if one can find differences between groups of people with and without disease with regards to their genetic composition, focusing on single nucleotides known to vary amongst people. A recent large GWAS, examining 36,989 cases and 113,075 controls, makes connections between different psychiatric diseases, and 108 genetic loci [3]. Even tough complex and polygenic, this study shows that some gene variants recur more often than others, in that they are aberrant in patients diagnosed with SCZ. The contributions of the identified gene variants on the development and behavioral traits of schizophrenia remain unknown but represent novel opportunities to explore and understand disease mechanisms and also identify novel drug targets and improved treatments. Recent breakthroughs with patient-derived induced pluripotent stem cells (iPSC) and the efficient and targeted genome editing of CRISPR/Cas9 in combination with optical and electrophysiological recordings and computational modeling introduce novel ways to investigate disease mechanisms but such an endeavor remains to be explored.

In order to establish such a work flow, each element needs to be developed and optimized. The current work explored the initial steps in this process and first followed differentiating iPSCs to evaluate their maturation stage and then pilot the use of genome editing of CRISPR/Cas9 with the aim to possibly manipulate the function of an ion channel postulated to be a contributor for impaired neuron function in schizophrenia.

1.1 Mental Disorders

A mental disorder is a condition of abnormal patterns of behavior, self-perception, feelings, and thinking [4].

The WHO, 2000-2012 Global health report estimates that 7 % of the Disability-adjusted life years are caused by mental and behavioral disorders [5] making mental illnesses world wide a grand societal challenge both for the individual and their dependents, but also for society as such [6].

The most common mental disorders are mainly subdivided into schizophrenia, bipolar disorder, major depressive disorder, autism spectrum disorder, and dementia [4]. If possible, disease are categorized by means of etiology, but causative information remains largely elusive.The diagnose is made by a mental health professional based on dialogue, history and observations of symptoms as there are no known biomarkers [7]. The lack of objective measures makes it challenging to make precise diagnoses and each disorder is likely composed of several subcategories with shared symptoms but perhaps caused by different mechanisms. Few therapeutic targets

have been found, leading to similar treatment of a potentially wide variety of diseases [8, 9]. This limitation was noted by the Schizophrenia Working Group of the Psychiatric Genomics Consortium [3]:

> «All available antipsychotic drugs are thought to exert their main therapeutic effects through blockade of the type 2 dopaminergic receptor, but since the discovery of this mechanism over 60 years ago, no new antipsychotic drug of proven efficacy has been developed based on other target molecules [3].»

Treatments not only often fail to work, but may have severe unwanted side effects [10, 11] that may further reduce function. With greater knowledge about the genetic composition and its contribution to mental illness, novel drug targets can be discovered and better understanding on how existing drugs work can be obtained.

1.1.1 Schizophrenia

Schizophrenia is a mental disorder affecting about 21 million people worldwide [4]. It is a disease that progresses with the distortion in thinking patterns, hallucination, delusions, abnormal motor behavior and negative symptoms such as apathy and lethargy. These are all symptoms emphasized in the DSM-5, and observed in this disorder, and a minimum of two symptoms qualifies for disease (DSM-5) [7]. Life time risk of getting the disease is 1%, and the heritability is estimated to be about 80%, in various studies of twins with different upbringing [2].

Onset of disease is late adolescence to early adulthood, which differs from other neurodevelopmental disorders that usually have an earlier onset [12]. This makes for an interesting detail as a lot of important developmental processes have already occurred in these patients, without obvious disruption of function. The genetics are complex and the diagnostic tools are symptom-based, leading to drug treatments that are unspecific and unreliable for a large fraction of the treated patients [9, 8].

The GWA studies may be the start of getting around this problem. These studies are based on the complete mapping of the human genome [13]. When conducting this in many individuals, it became apparent that many locations in the genome varies with regards to nucleotides in that specific position in so called single nucleotide polymorphisms (SNPs). The known SNPs can be screened in a tagging assay to find an individuals SNP variant [14]. When comparing a large group of healthy subjects with patient groups, one can study associations between a particular basepair at particular locations, and examine the correlation with specific diseases. In order to obtain enough statistical power, such studies require large sample sizes [3] as the statistical power of multiple tests done on the same dataset must compensate for the amount of tests conducted [15]. Since potentially every SNP tested is treated as a test, only very high correlations would yield significant results (false negatives) unless population sizes are big [15, 3]. This is accommodated by the ease with which acquiring of SNP assay data improves with sequencing technology [16]. Another challenge is the occasional co-inheritance of SNPs in so called haplotypes. This challenge, termed linkage disequilibrium (LD), are rising a potential risk for the statistical detection of several SNPs where only one is rooted in the causal link with the disease. Increasing the likelihood of false positives, unless a limit for the proximity between detected SNPs is set.

The work by Ripke and co-workers [3] collected data from a vast number of GWA studies in order to obtain the statistical power to identify loci with low disease correlation. Novel candidate genes for SCZ where identified associated with different functions [3]. One category of gene variants are involved with the immune system such as the immune relevant MHC protein while a group of gene SNPs are associated with genes specific to the central nervous system. Of these, some are important for neuronal development supporting the notion that schizophrenia may have a developmental component and others are important for neurotransmission.

1.2 Induced Pluripotent Stem Cells

The establishment of the conditions required for cells to grow outside their natural habitat, has been an important milestone for studying the human being [17]. Since experiments on humans have possible ethical challenges, the ability of growing cells in a dish has been the only possible way to get to study human cells, and not just some model animal assumed to be alike. Researchers have long seen that most cells stop dividing if grown in a dish, leading to a big breakthrough when cancer cells first were successfully grown outside the body [17]. These immortal cells had overcome senescence, the state where a cell no longer divides, and lives on. Many human cancer cell lines are very stable and have turned out suitable for a range of fields in molecular biology and cell biology.

Embryonic stem cells' ability to divide indefinitely and maintain pluripotency led researchers to investigate certain properties of these cells from the inner part of a blastocyst. These cells have the cancer cells ability to be maintained in a dish, and they constitute the representative genotype of the individual from whom they were obtained. This is a long sought solution to the problem of using model genomes that carries large abnormalities compared to the natural system. Cancer cell lines often carries mutations and aneuploidic karyotypes. It should be mentioned though, that also in iPSCs aberrant genomes have been implicated [18].

Another interesting feature are important for the advent of iPSCs, namely the property of becoming any type of cell [19]. Even egg cells have been made from iPSCs [20]. The DNA does not change during cell division, but the cloning technique of Briggs and King (1952), the SCNT(somatic cell nuclear transfer) was preliminary in the understanding of this being the case, and that other reversible factors exist to change differentiated cells back to pluripotency [19].

In 2006, S. Yamanaka published the finding that four transcription factors were enough to enable reprogramming of mouse fibroblasts into pluripotent stem cells. For this finding he was awarded the Nobel Prize in Medicine or Physiology together with John B. Gurdon in 2012 [21]. The achievement was reproduced with human cells the year after [21, 22, 23]. Many techniques have later developed to improve the reprogramming, with the goal of doing it in the least invasive manner, termed footprint-free iPSC derivation [24].

These findings allow for reprogramming of any person's cells, to be stably grown in a dish. The huge advantage is to get the exact patient genotype available for studying. This paves the road for personalized medicine [25]. When perfect genome-matching cells are obtainable for any one person, drug efficacy can be better assessed before clinical trials as well as reducing the costs of initiating clinical trials [26]. The genereric value is higher when the test is conducted on cells obtained from different individuals with variable genotype

In addition, it turns out these cells are quite robust, allowing for manipulation of their genome [27]. So rather than the common way of studying mutations (by inducing them in a cell line), patient derived genotyped cells of various diseases causing mutations, may be investigated. The disease-causing-SNPs can be studied in rectification experiments where correction by genome editing may be conducted, instead of induced [28, 29].

Challenges with these cells are the difficulty with which the pluripotency is induced making the efficiency of reprogramming relatively low [30, 26]. The cells require high maintenance to avoid differentiation, and laborious routines for controlling the proper reprogramming. Additionally, these cells are immature, meaning that when cells are being differentiated into a certain cell type their state of maturation can be difficult to verify. One may question if it is reasonable to compare a neuron that has been differentiated for years inside a human brain, with a neuron differentiated in a dish for 10 weeks.

Although the iPSCs would be the preferred choice, poor accessability, and the relatively difficult cultivation procedure made us choose the neurobloastoma cell line SH-SY5Y for the

initial establishment of the genome editing protocol in the current study. The neuroblastoma cell line SH-SY5Y is a human cell line from bone marrow that has measurable L-type calcium currents [31], making it suitable for this project.

1.3 Deciding the functional state of a neuronal cell

What is a mature neuron? When differentiating cells in a dish, one can not rely solely on the visuals obtained from a light microscope. Cells, even in the same dish, forms a continuum of maturational states that can be categorized[32, 33]. Here, the electrophysiological measures of sodium currents are mainly used for the classification. A threshold of -10mV is set, and action potential (AP) types is defined by the ability of a cell to elicit an AP overshooting that threshold. Furthermore, the types are divided by how many APs they can elicit and the highest class of neurons(AP type 5) can sustain at least a 10 Hz firing frequency above the threshold [32].

Mapping the potential in-a-dish-differences is important before proceeding with experiments working to find between-group differences - what are disease-specific-variations and what are in-culture-differences. A thorough work by Bardy and co-workers (2016) aimed to correlate specific gene markers (a high throughput method), to the electrophysiological state of a cell, acquired by patch clamping (low throughput method) [32]. The work elucidated nine up-regulated genes correlating with the highest AP type, 5. If these genes are upregulated in highly functional neurons they could work as a marker for when differentiated neurons are mature. The genes and their function are listed in table 1.

Gene	Function
CDKN2D	A cyclin dependent kinase inhibitor (CDKN) 2D is a cell cycle regulator for the progression of G1 phase. The protein has fluctuating expression with the progression of cell cycle stages. The concentration difference between these CDK inhibitors, and CDK's are controlling the fate of the cycle stage [34].
CKMT1B	A mitochondrial Creatine Kinase responsible for the phosphorylation of creatine, a cytosolic carrier of phosphate, that regenerates ATP (adenosine triphosphate) from ADP (adenosine diphosphate). It seems to have a buffering role for the concentration of ATP in fluctuations of energetic demands [35].
MTSS1	Metastasis Suppressor-1 is a gene implicated in cytoskeletal regulation, and the differentiation of neuronal cells. This tumor suppressor gene is found to be down-regulated in different metastasised cancers [36].
RGS9	Regulator of G-protein signaling 9 is a GTPase activating protein (GAP*), important for the regulation of the inactivation of the GPCR subunit, Gα, which is bound to GTP (guanosine triphosphate). This inactivation converting GTP to Gα – GDP, is implicated in transducin recovery of photoreceptors [37]. A brain variant RGS9-2 also exist [38].
TRAPPC6B	Trafficking protein particle complex 6B is a subunit of a protein found in vesicle transport. The TRAPP complex is found to work as a guanosine exchange factor (GEF*) [39], and possibly facilitate ER - golgi transport, it has been ascribed a possible tethering function [40]. Another TRAPP subunit, TRAPPC4, has been shown to bind syndecan-2, a factor implicated in dendritic spine formation [41].
TUBB4A	Tubuline β 4A is a gene encoding the microtubule (MT) subunit β-tubulin, which assembles in to the MT filament together with the α-subunit. Mutations in the gene is correlated with hypomyelination and atrophy of the basal ganglia and cerebellum (H-ABC) [42].
SCN9A	The sodium channel 9A encodes a voltage sensitive sodium channel (Nav1.7), A channel implicated in nociception, with loss of function leading to loss of pain sensitivity [43], and a gain of function associated with increased pain sensation [44, 45].
PCLO	The piccolo presynaptic cytomatrix protein is important for the active zone of the presynaptic cell. It has a suggested role in clustering of synaptic vesicles of hippocampal cells in culture [46]. SNPs in this gene are associated with MDD (major depressive disorder) [47], and expression analysis are indicative of up-regulation of this gene in Bipolar Disorder (BD) patients [48], these are observations made based on the aforementioned GWA studies [3].
GDAP1L1	Ganglioside induced differentiation associated protein 1-like protein 1 resembles a glutathion-S-transferase (GST), and is linked to the neuropathy of the PNS (peripheral nervous system), Charcot-Marie-Tooth disease [49]. As its name implies the protein is shown to be upregulated in N2A neurons differentiated by the use of the gangloside GD3 synthase.

*GEF's and GAP's serve opposite functions in the cell

Table 1: List of significant genes from Bardy et al 2016 [32] and their functions.

1.3.1 Quantitative Reverse Transcription Polymerase Chain Reaction

Quantitative(q) reverse transctription (RT) polymerase chain reaction (PCR), or qRT PCR has revolutionized the data acquisition of expression data. This technique, which is based on reverse transcribing the RNA content of a sample and then using the sensitivity of PCR to produce amplicons of specific genes, has long been the gold standard for studying specific genes' expressional state in a cell culture or whole tissue [50]. The PCR technique is implementing a fluorescent probe which binds the double stranded amplicon of the PCR. In that way, a camera can detect the fluorescence signal from a PCR reaction **[51]**. For each cycle of the PCR, the intensity of the fluorescence is measured and can be plotted against the cycle number. At a specific cycle the amplification in a successful PCR run reaches a threshold value, cycle threshold (C_t), where fluorescence exceeds background fluorescence. The amplification by PCR is potentially doubling the amount of fragments for each cycle in the early cycles before depletion of the primers, or competing of the polymeraseses occurs. Thus yielding a predictable growth phase that can be backtracked to the amount of initial RNA if compared to a standard curve generated from a set of samples with known concentrations. This method is termed absolute quantification, and assumes an equal efficiency of amplification between those samples [52].

The exact transcript number is a value which is not necessarily important. Often is the comparison of samples more relevant, and a relative quantification can be done. The compensation of possibly comparing samples with widely different amounts of cells or RNA concentrations is then important [53]. A gene assumed to be stably expressed in both treatment groups are working as reference, and can be used to normalize the values obtained from the target gene. When comparing different samples it is common place to make a calibrator sample which may consist of pooled RNA of all the samples, or maybe a different tissue than the one tested [54]. Then a relative fold change of expression compared to a calibrator yields a single value that can be compared between samples.

Multiple methods for calculating the relative difference between samples exist and they all utilize the normalizing reference [53]. The reference could be an internal gene, but it could also be an externally added RNA fragment. The use of internal reference genes have been widely criticized, due to the observed variation in expression levels [55]. They should thus be used with thorough verification of stability, or several genes should be used together to collectively constitute the normalization[55, 56]. A best-fit-reference-gene test can be done using a method that compares the variation of several tested genes[57].

This technique for expression level analysis, together with the knowledge of the nine genes found by Bardy et al. (2016)(previous section) will be used as an initial test to see if these nine genes can work as a quality control of differentiated neurons.

1.4 Genome Editing

The advantages that iPS cells are bringing to the table are flourishing in the scientific literature, the studying of the complete genome of an individual might reveal information which cancerous cell lines could not. With the study of gene expression one might see what small changes of the genome, as the SNPs, might contribute to the cell and a stable cell representative of the cell in our body is essential.

The ability to induce single nucleotide mutations into the genome has been possible for some time with techniques such as transcription activatior-like effector nucleases (TALENs) [58] and zinc finger nucleases (ZFN) [59]. The work load and price has decreased drastically however, with the utilization of RNA-guided endonucleases (RGENs). The existence of protein-DNA binding motifs, so called zinc finger motifs, have been assembled with nuclease proteins (*Fok*1) to bind specific sequences and induce a double stranded break(DSB)[59]. TALENs have been found in

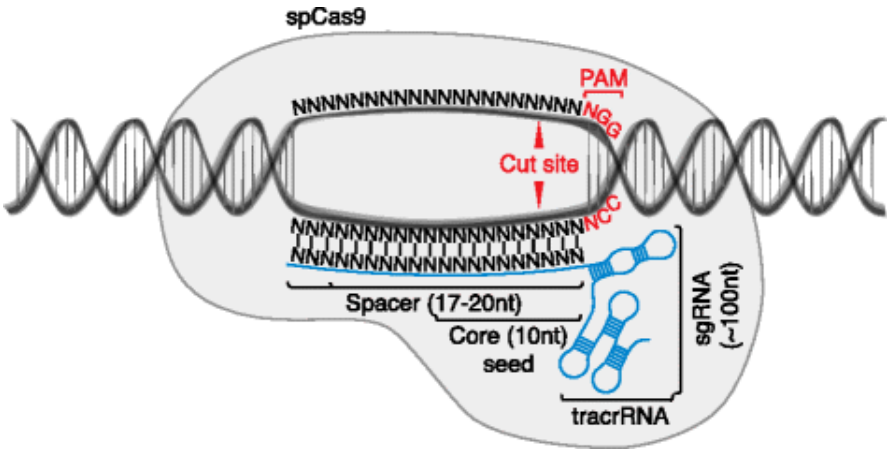


Figure 1: Cartoon of the CRISPR/Cas9 system from Graham and Root's 2015 paper[64].

bacterial cells and have a simpler DNA interacting domain. Two amino acid residues interact with one base of the DNA strand, thus allowing smaller protein changes between complexes that detect different DNA regions compared to the ZFN. These TALENs did not enjoy popular use after their mechanistic unveiling in 2009 [58], just before the RGENs stole the spot light in 2012 [60].

The production of proteins is a time consuming process, and the finding of nucleases interacting with RNA for recognition specificity, to conduct the same cleavage of DNA as these ZFNs and TALENs, has lifted the field to be accessible to any laboratory's inquiries. The specificity of the complementarity found between the bases in an RNA-DNA hybridization, is also not of yet possible to match by protein-DNA interactions.

1.4.1 CRISPR/Cas9

The genome editing technique clustered regularly interspaced short palindromic repeats(CRISPR)/ CRISPR-associated protein 9 (Cas9), is derived from an adaptive immune system of the bacterium *Streptococcus pyogenes* [61]. The immune system is found in most archaea and many bacteria, but differ somewhat in its molecular components [62]. The Cas9 protein from the *S. pyogenes* has proven useful for its combined helicase and its two nuclease domains, in addition to its ability to bind RNA which work as a guide for sequence recognition [63]. Thus a single protein, complexed with an RNA strand, enables the cleavage of almost any given sequence of DNA [60].

The single guide RNA (sgRNA), or the RNA strand, are found as arrays in the bacterium's genome. It occurs as alterations between phage-originating sequences and a regular repeat [65, 66, 67]. The phage-originating sequence, transcribes into CRISPR RNA (crRNA) which are amenable to some change without altering the complex, other than changing the specificity to what is bound, and can thus be made to base-pair with any sequence [60, 68]. The regular repeats translate into a specific trans-activating CRISPR RNA (tracrRNA), that has a binding affinity for both the Cas9 protein and the crRNA [60]. A palindromic sequence spanning the crRNA and the tracrRNA enables a loop formation and base-pairing between the two RNA parts, to form the sgRNA [69, 60]. In the sgRNA the two strands are covalently bound, but in the bacterial system the two strands are cleaved to yield two strands linked only by hydrogen bonding (base

pairing) in a dual-RNA complex [60].

The complex has been shown to require a protospacer adjacent motif (PAM) sequence to be able to initiate Watson-Crick base pairing and helix formation between the sgRNA, and the target DNA [70]. The PAM specific sequence for the spCas9 (*S. pyogenes* variant of Cas9) is a 5' - NGG – 3'. It is found immediately downstream of the sense strand of the target DNA, i.e. the strand that resembles the guide and thus the opposite strand of which the guide binds [60]. When the PAM sequence is identified, the "seed" part of the 20 nucleotides (nt) long recognition sequence on the 5-prime end of the crRNA starts hybridizing with the target DNA, if there is complementarity between the sequences. The "seeding" sequence consist of 10 - 12 of the 3'most nt of this recognition sequence [69].

When the sgRNA binds the Cas9 protein, the CRISPR/Cas9 complex undergoes a large conformational change which puts the HNH nuclease domain in proximity to the sgRNA, and makes space for a potential complementary target strand in the central channel of the protein [69]. This change enables the PAM interacting domain to interact with the GG motif, and initiate the interrogation of the target DNA with the "seeding" sequence, and then the whole 20 nt sequence, essential for producing the double-stranded break (DSB) [69, 60].

Different lengths of the recognition sequence have been tested for efficiency of cleavage. The N_{17}(17 nt) and N_{20}(20 nt) are equally efficient in their nuclease domains, but the RuvC exonuclease activity in the 3' - 5' direction is less efficient in the N_{17} guide, compared to the N_{20} guide [60]. And so the gold standard, for producing double stranded breaks with CRISPR/Cas9, has become using a N_{20} guide.

The CRISPR/Cas9 technique has been successfully used for genome editing in iPS cells [27]. There are many genes that would be interesting to study in iPSCs and experiments done in cancer cell lines might have different outcomes if conducted on these cells. The big unknown is however the interplay of genes and how the SNP variants are changing that intricate system one way or the other. We will adopt the technique of CRISPR/Cas9 to establish the first piece of the method for unveiling the SNPs. By inducing a well studied mutation implicated in a severe cardiac syndrome (Timothy syndrome) in one of the genes (CACNA1C) also found to be highly correlative with schizophrenia in GWA studies [3], we will initiate this work.

1.4.2 The natural repair mechanisms

There are two known pathways utilized by eukaryotic cells for the repair of DSB repair in the DNA; the non-homologues end joining (NHEJ) and the homology directed repair (HDR)[71, 72]. Utilization of CRISPR/Cas9 as a technique depends on these repair mechanisms for successful editing of the genome. The CRISPR/Cas9 only produces a break whereas the manipulation of the repair mechanisms produces the desired alteration in the gene of interest [73].

As the names suggest, the two mechanisms differ in whether or not a homologous template is used in the repair. The type of repair is critical for the outcome of the editing. The **NHEJ** pathway produces a mutation that knocks out the expressed gene often by generating frameshift mutations, whereas the HDR can be manipulated to produce deliberate knock-in mutations [74]. The understanding of how these pathways work can be informative of the outcome of the repair.

The NHEJ pathway utilizes DNA nucleases, DNA polymerases and DNA ligases [75]. The polymerases recruited for NHEJ in mammalian cells, are the pol μ and pol λ, from the polymerase X family members [75, 76]. The pol μ shows template independent polymerase function *in vitro*, with different preferences for some dideoxynucleotide triphosphates (dNTP) [77]. This activity has been suggested to be important for creating micro homology that can aid the NHEJ pathway. Homology of 1-4 bp can enable the XRCC4:DNA ligase IV to exert its function without the other proteins of the pathway [77, 78].

Direct repeats, inverted repeats, insertions and deletions are often seen at NHEJ junctions and the nature of the proteins in this pathway have tendencies to create such lesions. Direct repeats, two sequences that are exact copies and downstream of each other, can be caused by slippage of polymerases, an event where the polymerase falls off the template and continues over an already replicated part [79, 80, 81, 75]. Inverted repeats, two sequences downstream of each other where one is the reverse complement of the other, might be explained by having a strand from the DSB site folding on itself and induce replication, this is seen in DSB of V(D)J recombination sites in vertebrates [82]. Insertions/Deletions (InDels), the resulting loss or gain of bases at a locus, are often generated due to the tendencies of the proteins to add nucleotides to the breakage point. On some occasions nucleotides can be lost (e.g. endonucleolytic activity of DNA-PKcs) but most often polymerases are incorporating nucleotides at the ends [83]. The inadequate performance of the ku NHEJ-pathway ligase (XRCC4 ligase IV), does not help with these varying endpoint changes. The ligase shows the ability to ligate two duplexes with 3' overhangs, despite differences in the length of the overhang and with complete lack of complementarity [77]. Thus follows a wide arsenal of potential results generated by this repair mechanism.

The **HDR** pathways are utilizing an homologues template for the repair of the DNA lesion, and are thus limited to occur when such homologs are present in the nucleus [84, 72]. Many pathways exist but mainly one is working to produce the sought editing result when a homologous template is added to the transfection mix of the genome editing system, namely the synthesis dependent strand annealing (SDSA) [85]. This pathway invades the double stranded DNA, with one strand, and gets extended based on the homolog. This single strand then dissociate completely from the homolog, to associate the two helices resulting from the DSB, now possible due to the complementarity shared among the extended ssDNA and the other helix [86]. Since the mechanism is invading the break site with a single strand on both sides of the break, the donor template provided alongside the CRISPR/Cas reagent during delivery to the cell nucleus is often single stranded. Instead of the dissociation occurring in the natural mechanism of SDSA, this donor template strand is implemented in the repaired DNA.

The Homologous recombination have long been the method of choice for inducing exact changes in the genome, but the rate of it happening is minute. The inducing of DSBs have vastly improved this efficiency, but still the occurrence is low, and good verification methods have been developed. The verification of the occurrence of such an event can be done with slight manipulation of the template. First, if the correct implementation has occurred, the CRISPR/Cas9 must not continue to cut. A shield mutation – a silent mutation that disrupts the PAM recognition site – can be generated to hide or shield the recognized site from the nuclease [87, 88]. Second, implementing a silent mutation that yields a new restriction site, enables the visualization of a cut band on a gel to distinguish the wild type (WT) from the mutant by use of the respective restriction enzyme, in a verification of the editing [89]. Additional manipulations of the donor template, which is not of the silent kind, can then be implemented as the experimenter sees fit.

The observed low efficiency of this technique is a problem that follows suit the difficulty of getting all the needed components into the cell nucleus. If not optimized, these can compromise the experiments.

1.5 Transfection of cells

Lots of techniques have been developed to get molecules into the cell. Molecules like RNA, DNA, proteins, color dyes, fluorescent dyes, or hybrids of these, have been applied in order to cause changes or report events. The complexity and relative difference amongst cell membranes and

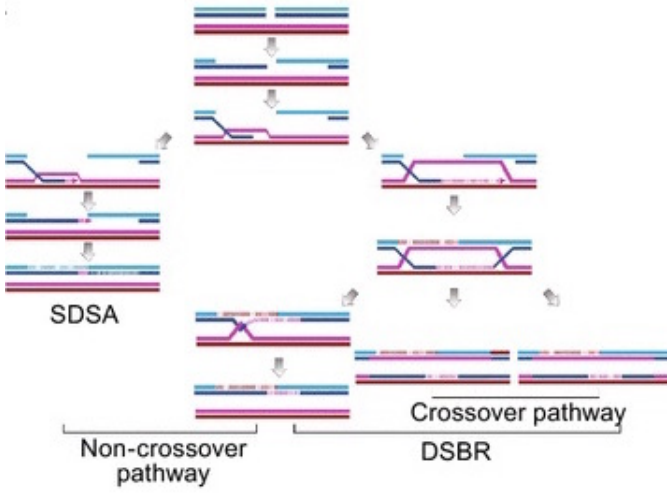


Figure 2: The homologous repair pathways DSBR and SDSA and their possible outcomes, figure from Liang et al 2015[87].

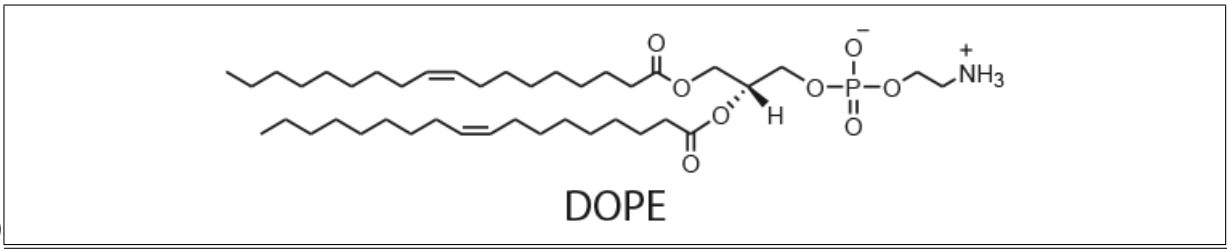
among these molecules makes for a vast variability, and there is no universal techinque working under all circumstances. Lipofectamine, electroporation, and viral transfection are commonly used and discussed briefly in the following section (for review see [90]).

1.5.1 Lipofectamine

The lipid based amines, or more generally referred to as cationic lipids, are binding the negative charge of DNA to shield it of from the negative membrane. The positive charge of the amine groups then assembles into structures with DNA, neutralizing its negative charge, to facilitate the transfer of the DNA to the inside of the cell [91]. The formulation of these compounds are highly proprietary and difficult to find, here is however a description of the composition of LipofectAMINE (available from Life Technologies, Inc., Rockville, Md.): A 3:1 ratio of the dioleoylphosphatidylethanolamine (DOPE) to the 2,3-di-oleyloxy-N-[2(spermine-carboxamido)ethyl]-N,N-dimethyl-l-propan-aminium (DOSP A) compound constitute the product, found in the cited patent [92]. The figure 3 depicts these two compounds.

As small RNA molecules performing silencing of transcribed genes and CRISPR/Cas genome editing have entered the field, these products have been optimized with extra reagents to better suite each compound. The technique constitute a simple and achievable way of transfecting cells, and the promise of successful transfection in iPS cells [27]. The near null efficacy under some circumstances are however leading to alot of uncertainty when using this method for transfection of a cell line and a compound which is never tested together before [93].

a)



b)

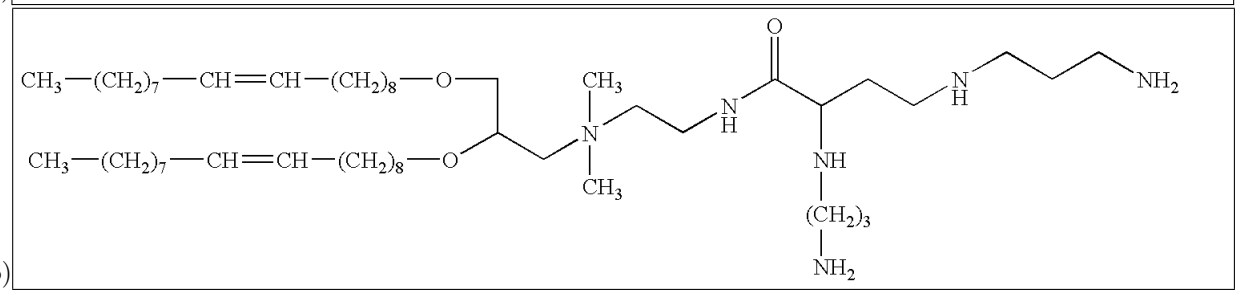


Figure 3: Depicted are the compounds found in LipofectAMINE. A 3:1 formulation of the DOPE (a) and the DOSP A (b) constitutes the LipofectAMINE (available from Life Technologies, Inc., Rockville, Md.). Figure of DOPE was borrowed from the web page, https://www.mirusbio.com/transfectopedia/methods, whereas the upper DOSP A figure was taken from the patent document [92].

1.5.2 Electroporation

The electroporation method is based on shocking the cells in a solution with a brief electric pulse/pulses, the membrane destabilizes and pores allow the flux of solutes in and out of the cell. The flux direction is dependent on the charge of the particles, as can be seen in the experiment conducted by Paganin and co-workers (2011) where negatively charged siRNA only got introduced at the side of the cell facing the cathode [94]. Electroporation seems to represent more stable efficiency between different cell types/lines compared lipofectamine [27, 93].

1.5.3 Viral transfection

Viral transfection is based on the packaging of the system into the genome of a viral particle. Several viral vectors are available and lenti viruses and Adeno-associatiated viruses (AAVs) are most commonly used in neuroscience [95]. The relatively large size of the Cas9-package have raised some problems for the relatively low capacity of the AAVs [96], although this is partially solved by the discovery of additional nuclease proteins with smaller sizes(e.g. Cpf1) [97], or co-delivery of multiple particles containing different parts of the complex [96]. These endonucleases have however not been tested to the same extent as the Cas9, but the Cpf1 has been suggested to be better suited for HDR editing as the nuclease cuts further from the PAM sequence. This limits the effect a NHEJ event has on the recognition sequence, promoting the cutting until an HDR event occurs [97]. The lentiviral capsid has about twice the capacity of the AAV [95], but has the potential risk of infection looming, not because the likelihood is greater, but because the potential outcome is worse [98]. The viral particles can, for a high price, be ordered ready made, but that again is time consuming and not compliant with a flexible system that should be easily manipulated to target new sites.

1.6 Aims of the study

The complexity observed in multifactorial diseases (MFD) have proven a hard nut to crack, but the development in iPS cells enables the study of a full patient genotype. This might improve the field of MFD, and our big goal is to map the severity of SNPs found to be important in GWA studies, and figure out which targets are more efficient for treatment in a given set of mutations.

The goal of this study was first to establish a procedure for how to identify maturation stage of differentiated neurons from patient derived iPSCs. To address this I constructed a qPCR experiment to measure gene expression of genes correlating with a high functional state neuron. Secondly, I sought to establish CRISPR/Cas9 as a tool for the investigation of SNPs found in patients with psychiatric diseases. Addressing this was done by transfecting a cell line exhibiting L-type Calcium channel activity and as a proof of principle trying to induce a DSB in the pore forming alpha unit of the channel.

2 Materials and methods

2.1 Cell Culturing

2.1.1 Resuscitation

The bone marrow neuroblastoma cell line, SH-SY5Y, was purchased from Sigma (England) and were stored at -80 degrees upon receipt. Resuscitation of the frozen cells was done according to the American Type Culture Collection (ATCC ®) recommendations (CRL-2266^{TM}) and aliquots were made for later use. Resuscitation was done by incubating the vial in 37 $^\circ$ C for 1-2 min. The working bench, surfaces, utensils and the vial with cells were pre-wiped with 70% ethanol. Cell content was transferred to a 15 ml sterile centrifuge tube. Thereafter, five ml pre-warmed medium was added to the tube, and centrifuged for five minutes, at 150 times gravity ($\cdot g$). Then dimethyl sulphoxide (DMSO) and medium containing supernatant was removed, and five ml fresh medium was added to the cells. The pipette was then used to re-suspend the sedimented cell pellet. The solution was transferred to a flask (T175), and grown at 37 $^\circ$ C in an incubator, with 5% CO2. The medium was changed after 24 hours to rid the flask of debris from dead cells, and then medium was renewed every 4-7 days.

2.1.2 Cryopreservation

Once the confluency reached 80-90% cells from the resuscitation protocol were acquired, the cells were aliquoted into separate vials of one ml and cooled down gradually at -80 degrees before, cryopreserved at -196 degrees in liquid nitrogen for later use. At this 80-90% confluency, the medium was first aspirated and the cells washed with phosphate-buffered saline (PBS). Then the cells were incubated in 1-2 min at room temperature in preheated 1-2 ml trypsin(0.25%)/ethylenediaminetetraacetic acid (EDTA). Detachment of the cells from the bottom of the flask and separation of cell clusters were quickly verified in a light microscope. A neutralizing step was conducted adding fresh medium, twice the amount of the 1-2 ml trypsin/EDTA.

A cell count was performed to estimate number of vials to make. Approximately $2-4\cdot10^6$ cells per vial (one ml) were considered ideal. The cells were centrifuged at $150\cdot g$ for 5 min, and the supernatant was removed. The cells were resuspended in a $2-4\cdot10^6$ cells/ml freeze medium (see table 2), and loaded in cryoprotective ampoules of one ml. Date, concentration, passage number, and cell line were noted. The ampoules were placed in a styrofoam box with paper inside to ensure slow cooling rate of about 1 $^\circ$ C/min down to -80 $^\circ$ C. Once at -80 $^\circ$ C, the vials were transferred to a liquid nitrogen tank for storage.

2.1.3 Subculturing

Subculturing routines was conducted by preheating medium, trypsin, and PBS to 37 $^\circ$ C. Bottles, surfaces, and equipment were sprayed with 70% ethanol. The medium was aspirated from the flask, and cells were rinsed in PBS $\sim$ 6 or $\sim$ 10 ml (T25 or T75 flasks, respectively). The PBS was aspirated and 0,5 or 1-2 ml 0,25% Trypsin/EDTA (T25 or T75 flasks, respectively) were added for the cells to separate in solution. Incubation for 1-2 min at room temperature, gently stirring and knocking the flask against the edge of the bench were done to best ensure resuspension and separation of cell clumps. The proteolytic activity of trypsin was quenched by adding of medium. Depending on the planned day of experimentation, the amount of cells seeded on to a new flask, and what flask to use were chosen. The AATC culture procedures suggest a mixing of medium containing cells with fresh medium in a 1:20 or 1:50 ratio. I usually chose a higher sub-cultivation ratio for experiments planed just a couple of days a head (1:6). Flasks were labeled with passage

Complete Growth Medium(CGM)	1:1 DMEM:F12 88% FCS(fetal calf serum) 10% Pen/strep 1% Glutamax 1%
Freeze medium	DMSO 5% CGM 95%

Table 2: Table of the different types of medium used, and their contents.

number, date, cell line and name of experimenter, and placed in the incubator at 37 ° C with 95% air and 5% CO2.

HEK293 cells were tested as a verification of the transfection and were subcultured by the same methods as SH-SY5Y cells, only difference being the use of Eagle's Minimum Essential Medium (EMEM, life technologies) instead of the Advanced Dulbecco's Modified Eagle's Medium/F12 (DMEM/F12, life technologies).

2.1.4 Cell counting

Cell counting was performed after trypsin treatment of cells(see subculturing). Afterwards, a 5 µl volume of the medium with suspended and separated cells were pipetted into a 1,5 ml eppendorf tube, and an equal volume of tryphan blue. The solution was placed with a 1-10 µl pipette onto a micro grid which was used as a reference for counting cells under the microscope (Olympus CK2, magnification 20x).The cells in four grids of $0.1mm^3(1mm^2 \cdot 0,1mm)$ were counted and divided by four to obtain a mean. The counted number was multiplied by 10^4 to account for cells per ml. That number was multiplied by two to account for the added volume of tryphan blue. If the original 5 µl was taken from a total of 5 ml volume, the number calculated now representing cells per ml, must be multiplied by 5 to yield total cell count. Thus the final factors in this scenario would be 10^5 multiplied by the cells counted in the $0.1mm^3$ chamber.

2.2 qRT PCR

2.2.1 Primer design

In order to amplify small fragments of the genes found in table 1 primers were designed using eurofinsgenomic's web portal which is based on the Prime+ GCG Wisconsin Package software originally made by Irv Edelman [99]. Specifications used were a 3' G/C clamp, 3ºC melting temperature difference, primer length of 20-24 base pairs, primer GC content within 40-60% a melting temperature within 50-65 degrees, and a size range from 80 to 300 base pairs. The mRNA sequence were found through NCBI's web page, and exon-exon boundaries were noted to make sure the primers spanned at least one junction. If isoforms of the gene existed, the needle algorithm from EMBOSS or the clustal omega algorithm were used to align them and common exons were chosen.

Clustal alignment of Ganglioside-induced differentiation-associated protein 1-like 1 (GDAP1L1) showed less than optimal match for the isoform variant 5 in the last exon chosen. None of the other junctions had optimal alignment for all 5 isoforms. No data on this variant was available, thus the option was to choose primers that implemented most transcript variants (excluding variant 5). Expression of some variants are limited to specific tissues, and leaving out one variant can yield a skewed picture of the expression rate of the gene. This is seen with the two variants

of the Timothy Syndrome (1 and 2), where exon 8 has a high occurrence in heart whereas exon 8A has not [100].

2.2.2 Total purification

The iPSC pellets and RNA samples were obtained from Agata Antonina Rita Impellizzeri and Srjan Djurovic (Oslo University Hospital). The cells were collected at different time points of differentiation treatments. The samples were prepared by Agata as described in table 3. The cell pellet samples were purified by total RNA isolation using the Qiagen RNeasy plus micro protocol. See appendix 7.5

	Control (1)	Control (2)	SCZ (3)	SCZ (4)
IPS cells	Cell pellet	RNA	RNA	RNA
7 days in to differentiation	Cell pellet	Cell pellet	Cell pellet	Cell pellet
Neuronal stem cells	Cell pellet	Cell pellet	Cell pellet	Cell pellet

Table 3: Sample conditions, of the samples gotten from the Djurovic lab.

2.2.3 Making cDNA with Qiagen QuantiTect Reverse Transcription kit

This Qiagen kit was used to make complimentary DNA (cDNA) of total mRNA in the cell samples, using a mix of random primers and oligo(dT). RNA amount was measured on a nanodrop 2000, the protocol was followed accordingly, see 7.6.

2.2.4 qPCR using FastStart Essential DNA green master on lightcycler (Roche)

The samples were prepared on a 96 well qPCR plate and mixed with the components provided by the FastStart Essential DNA green master kit (Roche) according to the protocol in 7.7. The qPCR was initially ran on a LightCycler 96 machine (Roche) to obtain curves (fluorescent signal to cycle number), but the limited access to the raw data points form each cycle of this machine became a problem. These data points are necessary for the subsequent analysis with the LinReg software, and is available with the LightCycler 480 (Roche) which I also had access to.

2.2.5 Primer testing

In order to test primers, a qPCR run on one of the control cDNA samples was conducted. The first order of 27 primer pairs was prepared for the first test to study the melting curve, tendency to generate primer dimers, and C_t value. One pair for each gene were chosen where the mentioned parameters were considered: (1) The best melting curve has one clear peak with little noise, indicating a single amplicon (2) the water sample was checked for fluorescence, and any sign of that would be indicative of dimerization of the primers (3) the $C_t$value was an initial test for which primer concentrations were ideal in the PCR reaction (see 7.7) and to get an indication of the amount of complementary(c)DNA in the samples. The ideal $C_t$values should be between cycle 20 and 35.

Two dilutions of the sample were also made in this initial test of primers, to figure out how to optimally dilute the isolated RNA samples for good results from these experiments. The best primers for each of the ten genes (nine experimental and one reference) were implemented in another test where three dilutions of one RNA sample were made to produce a standard

curve. The standard curve of the housekeeping gene, Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), working as a reference. The efficiency of the primers of the experimental genes of interest (GOI) and the GADPH primers must be near equal for a credible result from the $2^{-\Delta\Delta Ct}$ method [54].

The primer verification method from Livak and co-workers (2001) was carried out with a dilution series of five dilutions (1, 0.5, 0.1, 0.05, 0.01) in a final test, since the results from the series of three dilutions did not make any sense and suspicion of too few data points rose. The results from the five series did however not make any more sense. This was likely due to improper C_t range, and due to limited sample material a standard curve-free method was adopted.

Efficiency measures of the amplification curve are obtainable if studying the exponential phase of the fluorescence to cycle curve. The maximal doubling of fragments for each cycle yields a certain incline of the slope in a window of linearity in this curve. A software analyzing each data point of the curve was then used to obtain a percentage of this maximum(100%) which is defined as the amplification efficiency (LinReg).

A fold change between 1.8 and 2 (efficiency 80-100%) is considered sufficient for a good primer pair, see 7.8 for primer overview. In this test 30 more primers were added and some excluded from the crude results of the first tests. The POLR2A reference gene was also added for accuracy of the normalization and for redundancy, see discussion section 4.1.1.

2.2.6 qPCR control samples

The samples were studied, based on melting curves, the negative control (noRT sample), and C_t value. Three technical replicates - equally mixed samples made to exclude variation occurring in the qPCR machinery - were made and reproducibility between runs were considered. Obvious outliers based on biological and technical reasoning were excluded.

2.2.7 The Pfaffl modified -ΔΔCt method for relative quantification of expression

Expression levels were calculated from the qPCR results based on the modified $2^{-\Delta\Delta Ct}$ method suggested by Pfaffl [101]. In order to account for the variance in efficiency of amplification between the primer pairs the efficiency is implemented in this modified version. Of the three technical replicates, an average was calculated for the reference genes. For calibration in the Pfaffl formula, a pooled sample of all the cDNA samples were made. Expressional fold change compared to the calibrator was calculated with the average value for the reference gene, but for all the three replicates of the GOI. The arithmetic mean of the resulting fold change was calculated before analyzing the results. A repeated measure non-parametric one way analysis of variance (ANOVA) test, Friedman's test, followed by a posthoc test for significance, the non-parametric Dunn's test, was conducted.

2.3 CRISPR/Cas9

2.3.1 Designing guide RNAs

Three guideRNAs were designed to target the DNA close to the SNP target. The guides were named Seq1 to Seq3 (figure 4 and table 4). The Benchling[Biology Software](2016), was used to generate sequence 1 while the Sanger web page (http://www.sanger.ac.uk/htgt/wge/search_by_seq) was used to create sequence 2 and 3. Additionally, the primers for the CACNA1C gene exon 8A were made by the Benchling software (table 4).

Figure 4: a) Seq 1 is a CRISPR guide designed by using the Benchling tool, the Seq 2 and Seq 3, are CRISPR guides designed by using HTGT WGE Sanger Institute tool for CRISPR (they have the Sanger id's:1091677949 and 1091677950). a) and b) illustrates the primers used for the PCR amplification of the Exon 8 of the CACNA1C gene, see table 4for the sequences.

type	Sequence
seq1	5' - GGT CAA TGA TGC CGT AGG AA - 3'
seq2	5' - TT GTT ACA CTA ATC ATC ATA GGG - 3'
seq3	5' - GT CAT TTT TTG TAC TTA ACT TGG - 3'
Fwd primer	5' - TGA GGT GAC CCA TTT TCT CAG ACA – 3'
Rev primer	5'- CTC AGA GAT AGA CTG CTC AGT CTA TG - 3'

* Red color indicate the PAM sequence, this is not included in the sequence from Benchling.

Table 4: gRNA target sequence for the three CRISPR/Cas9 guides targeting exon 8A of the CACNA1C gene. Forward (Fwd) and reverse (Rev) primers for the CACNA1C gene flanking all the three guide targets.

2.3.2 Transfection with Lipofectamine

In order to transfect the cells, CRISPR/Cas9 complex was assembled by first mixing crRNA and tracrRNA to form the gRNA in Nuclease-free duplex buffer (IDT, USA) to a final concentration of 1 µM. The mixture was heated and kept at 95 ° C for 5 min during continuous stirring, before cooling to room temperature. This procedure should ensure hybridizing of the overlapping, specially modified 16 nucleotides of the Alt-R tracrRNA and the Alt-R-crRNA from IDT. Next, the Cas9 protein was dissolved in Complete Growth Medium (CGM, see table 2) without antibiotics ($^-$), together with the gRNA at room temperature for 5 min. The Ribonucleo protein (RNP) complex in the CGM^- was mixed to a final concentration of 60 nM.

The transfecting agent was made by adding 1,2 µl Lipofectamine CRISPRMAX (Thermo Scientific, Oslo) (2,4% (v/v) of the volume), to 23,8 µl of CGM^-(47,6% (v/v) of volume). From the RNP solution, 25 µl was pipetted to the lipofectamine reagents (50% (v/v) of the volume, final 30nM RNP concentration). and incubated for 5-10 minutes.

The cells were grown to 80-90% confluency before harvesting. The harvested cells were suspended in CGM$^-$ to a concentration of about 400 000 cells/ml, decided by a cell count and appropriate dilution. The 96-well plate were filled with transfecting agent (50µl per well) and incubated for 20 min before the cells suspended in CGM^- (100 µl per well) were added. The solution with cells and transfecting agent was incubated at 37 degrees for 48 hours to permit CRISPR/Cas9 transfection. Thereafter the cells were collected and the DNA purified.

2.3.3 Viability and transfection efficiency

In order to test the viability of SH-SY5Y cells after Lipofectamine 3000 (Thermo Scientific, Oslo) treatment, the following transfection protocol for green fluorescent protein (GFP) was used. Cells were first seeded in 20 wells of a 96-well plate, to reach about $0.7 - 2 \cdot 10^4$ cells/well on the following day. The next day, the cells were treated with the following concentrations of DNA and lipofectamine:

1. The Lipofectamine volumes from table 5 were diluted in Opti-MEM to final volumes of 5 µl.

2. A master mix was prepared for all wells, Opti-MEM (5 µl per well) and P3000 reagent (0.4 µl per well) from the lipofectamine kit, and was then mixed with the listed weights of GFP plasmid DNA (table 5), taken from a 2 µg/ml stock solution.

		Lipofectamine volume (µl)			
		0.15	0.3	0.6	1.2
DNA Weight (µg)	0.2	1	6	11	16
	0.3	2	7	12	17
	0.4	3	8	13	18
	0.5	4	9	14	19
	1	5	10	15	20

Table 5: Lipofection optimization. The different conditions in all of the 20 different wells (indicated by a number 1-20), in the transfection test protocol. Each well is treated with a range of different concentrations of GFP plasmid DNA and lipofectamine.

3. The reagents were mixed and incubated for 10 - 15 min at room temperature.

4. The DNA:lipofectamine complex were added to the cells and incubated for three days, before cells were photographed in a fluorescent microscope (Olympus IX71, magnification 20x).

In order to examine transfection level and survival of cells two images from each well, the same frame with fluorescence and regular light, were acquired using a Olympus IX71 microscope (20x magnification). The frame was arbitrarily chosen, and was covering about one fifth of a well of a 96 well plate. Pictures were analyzed using ImageJ and live and dead cells were counted within a fixed size frame where the marker tool in ImageJ were used, with the *Measure* feature under the *Analyze* tab, to acquire an Area of ~ 50 in the *Measure* window.

Cells were excited by a blue laser (470 nm, Sutter instruments Lambda 10-2) to obtain fluorescent signal from the cells expressing GFP. Images were taken with a camera, attached to the microscope (Olympus IX71, magnification 20x). The fluorescent signal was analyzed in the ImageJ software by using the threshold tool, under the *Image* $\rightarrow$ *Adjust* $\rightarrow$ *Threshold* tab. The upper slider was set to about 250 and the lower one to 1017, but this was adjusted by sliding the upper slider until a small change yielded a relatively large increase in unspecific noise particles in the picture, and then drag the slider back right before the change. Pictures are then transformed into fields of black representing cells, and the rest of the image is white enabling the *Analyze Particle* feature of ImageJ to be used, under the *Analyze* $\rightarrow$ *Analyze Particle* tab. The smallest particle known to be a cell was marked and Area measured, again using the measuring tool. The area value was put into the particle analyzing software "Size(cm^2)" parameter. Checking off the Results, and Summary boxes and choosing *Outlines* in the *Show* roll-down window, sums up the measures of the particles, give out a particle count, and marks the outlines of all the particles in a new image. The count was noted.

2.3.4 Verification of genome edit

Extraction of DNA was carried out using a DNeasy Miniprep kit (Qiagen). The DNA yield measured (ng/µl) was obtained using a spectrophotometer (NanoDrop 2000, Thermo Scientific), measuring absorbance at 260/280 nm. The DNA was, if possible, diluted to a concentration of 20 ng/µl, in Tris EDTA buffer provided in the kit. For protocol see 7.1.

Cycle stage	Temperature (°C)	Time (min)
Initial denaturation	98	0:30
Denature	98	0:10
Anneal	60-70(gradient)	0:30
Extend	72	0:30
Final extension	72	2:00

Table 6: PCR Gradient temperatures and time for each cycle.

PCReaction for amplification of the altered gene segment

Primers for PCR were ordered from ThermoFisher. Literature was searched for previous amplification of the exon 8A region, but only a reverse primer from Yazawa and co-workers (2011) was suitable: 5'- CTC AGA GAT AGA CTG CTC AGT CTA TG – 3' as the forward primer was too close to the target **[102]**. The forward primer was calculated using the primer3 web tool (http://bioinfo.ut.ee/primer3-0.4.0/): 5' - TGA GGT GAC CCA TTT TCT CAG ACA – 3'. This PCR product of 1222 bp was then verified using genomic DNA from SH-SY5Y cells to test annealing temperature of primers.

A New England Biolab (NEB) PCR Q5 Hi-Fi polymerase kit was used in accordance with the manufacturer's protocol, see 7.2. The mixture was pipetted to PCR tubes (VWR) that were put into a programmable PCR machine (Eppendorf Mastercycler). Cycling routines are found in 6. Optimal annealing temperatures were decided by running a program with a gradient of a 10 °C difference, between minimum and maximum temperature, yielding one degree difference between adjacent wells. Eight samples were tested ranging from $60 - 67$ °C in annealing temperature. The cycle are found in table 6.

The test was run for both primers; the Hypoxanthine-guanine phosphoribosyltransferase (HPRT) gene included in the control part of the CRISPR-Kit provided by IDT and the experimental primers for the exon 8A of the CACNA1C gene. Subsequent primer runs were conducted with the a 65°C temperature for the CACNA1C gene, and 64°C for the HPRT gene.

Purification of PCR product with ethanol precipitation

The PCR product was purified using a generic ethanol precipitation protocol:

1. Add 1/10th of the solution volume of sodium acetate (NaOAC)
2. Add one volume isopropanol
3. Vortex the solution and spin at maximum speed for 15 minutes at room temperature
4. Decant alcohol and add 500 µl 70% ethanol and vortex
5. Spin at maximum speed for 5 minutes at room temperature
6. Decant alcohol and remove trace ethanol with a quick spin with the lid open

7. Dry the pellet briefly

8. Re-suspend the pellet in 30 µl of water for the following T7 Endonuclease 1 (T7E1) treatment

The DNA concentrations were measured using a NanoDrop 2000 spectrophotometer from Thermo Scientific as described above (2.3.4).

Heteroduplex formation

The purified PCR product from a non-treated sample was mixed with the purified PCR product of a treated one in equimolar amounts. A solution described in 7.3 (first point), was made and the resulting mixture of potentially different helices of DNA was heat-treated in the Eppendorf Mastercycler with the following programmed ramp cycle to produce hetero duplexes:

Temperature	Time	Temperature ramp
95°C	10 min	
95°C to 85°C		(–2.0°C/sec)
85°C	1 min	
85°C to 75°C		(–0.3°C/sec)
75°C	1 min	
75°C to 65°C		(–0.3°C/sec)
65°C	1 min	
65°C to 55°C		(–0.3°C/sec)
55°C	1 min	
55°C to 45°C		(–0.3°C/sec)
45°C	1 min	
45°C to 35°C		(–0.3°C/sec)
35°C	1 min	
35°C to 25°C		(–0.3°C/sec)
25°C	1 min	
4°C	Hold ∞	

Table 7: Ramp Cycle for heteroduplex formation.

T7 endonuclease 1 and the EnGen Kit

In order to asses the genome editing efficiency of the CRISPR/Cas9 treatment, the T7 Endonuclease protocol accompanying the nuclease from NEB was then followed from point 3, see 7.3.

In order to assess if the CRISPR/Cas9 was successful, the EnGen Kit (New England Biolabs) containing complete set of controls was used. The kit supplied with its own polymerase enabled

direct T7E1 treatment after PCR, omitting the ethanol precipitation. The components were compatible with the T7E1 without this purification step. For protocol see 7.4.

Gel electrophoresis

In order to separate fragments of DNA of different length a gel electrophoresis was done. The gel was mixed using the following components:

- 1X TAE buffer[40mM Tris-Acetate (pH 8,3) 1mM EDTA]
- Agarose
- SYBR green gel-loading dye or Ethidium Bromide (EtBr)
- Gel casting tray and tape
- Electrophoresis chamber
- 100bp ladder from NEB
- Loading Dye [10 mM Tris-Hcl (pH8,0), 10mM EDTA (pH8,0), 50% (w/v) sucrose 0,15 (w/v)]

For casting the gel 50 ml TAE buffer was mixed with 0.5-1 grams of agarose, depending on the percentage of the gel (1-2%) made. The agarose crystals were dissolved in the TAE buffer in a microwave oven for about 60 sec. Afterwards, the beaker was cooled in the sink while swirled, and 5 µl SYBR green/1-2 µl EtBr was added. After additional cooling, to about 40 degrees, the liquid was poured into a prepared cast and a comb for casting of loading wells was put into the liquid. The preparation was set aside for solidification for about half an hour.

For loading the gel the finished cast was placed in the electrophoresis chamber and the chamber was filled until the cast was completely immersed in TAE buffer. The DNA fragments were prepared on ice with 1/6th the volume loading dye before the samples were loaded into the wells. An additional well was loaded with a 100 bp ladder from New England Biolabs (NEB) or a 50 bp Gene ruler from (Thermofisher).

3 Results

An initial search for gene expression that might be telling of the functional state of a neuron was done, and as I've gotten to learn about the methods in more detail, questions about how they were conducted here have been raised. Some questions, I have been trying to study further with the available data, and some questions still remain elusive. The next part of the project, which has been the genome editing, was carried out without much luck and as this work comes to an end the remaining question raised by this work still stands. Can the genome editing procedure presented here work if a lower lipofectamine concentration had been used?

3.1 qRT PCR

3.1.1 Fold change analysis

Expressional analysis of the nine genes in table 1 was done on iPSC-derived neurons at different time points of the differentiation.The fold change in expression was calculated using the Pfaffl method [101] based on calibration sample and a reference gene to account for intersample differences. Cells from four individuals were obtained at the Oslo university hospital, from the group of Srjan Djurovic. Fibroblasts were processed to obtain iPSCs, and those were again differentiated into neurons. Cells from two controls and two schizophrenic patients were harvested as iPSCs, 7 days after differentiation started, and when differentiated into neurons. One iPSC sample (control 2, at the iPSC stage) showed large inconsistencies in melting curves and large variations between technical replicates and samples where no reverse transcriptase were added (the variation can be found in the appendix 7.9). This sample also had obvious precipitate when received from the university hospital lab. The sample was for these reasons omitted in the analysis which was done on samples from the three remaining individuals.

In order to test the differences observed between the groups for statistical significance a non-parametric one-way ANOVA(analysis of variance) was conducted. This test, the Friedman test, was run to check the variance for one gene's expression in all the individuals, for all the time points. All nine target genes went through this test. The cyclin dependent kinase inhibitor 2D (CDKN2D) gene showed significantly different variability between the three groups figure 5 (χ^2 = 6.000, p-value 0.0278 Friedman's test). The difference was found to be between the iPSC group and the neuronal cell group when running a posthoc test (adjusted p-value of 0.0429 in Dunn's multiple comparisons test). No difference was found between groups for the other eight target genes as indicated in figure 6. The lack of difference in eight out of the nine genes, supposedly up-regulated in mature neurons, suggests that these genes might not be up-regulated after all, or other factors have impacted the credibility of the results (see discussion).

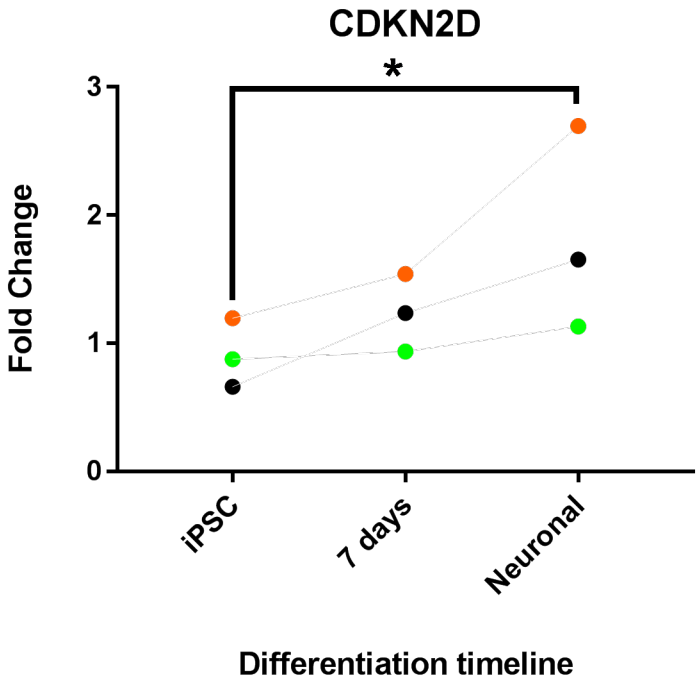


Figure 5: The relative expression of the CDKN2D gene normalized against the POLR2A gene. Each dot represent cells from one individual, where the cells are harvested from a well of a 6-well plate. The orange dots show expression levels from the first control individual, the black dots are generated from the first schizophrenic individual, and the green dots are from the second schizophrenic individual. Cells from each individual are collected at three different stages of differentiation. The CDKN2D gene express statistical significance in expression-level-difference between the iPSC, and the Neuronal cell group (Adjusted p-value 0.0429* from Dunn's multiple comparisons test). The difference between the 7 days of differentiation group, and the other groups are not significant.

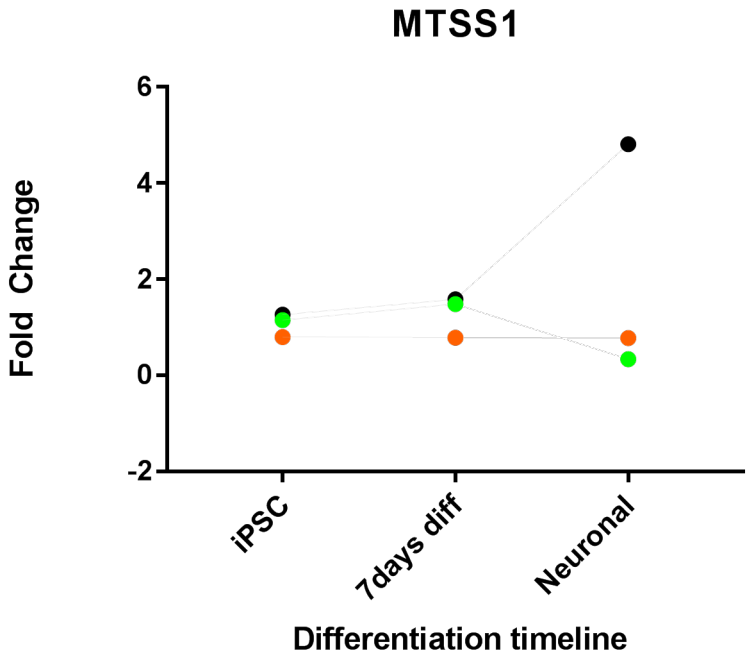


Figure 6: Example of one of the none significant genes' (MTSS1) expression levels at the different time points and for the differnt individuals, equal to what is shown in figure 5. The dots in this figure are represented by the same individuals and normalized using the same gene (POLR2A), only in this graph we are looking at the normalized relative expression level of the MTSS1 gene.

Gene	Test value	P-value	degrees of freedom(df)
CDKN2D	6	0.0278	2
CKMT1B	4.667	0.1944	2
MTSS1	0.6667	0.9444	2
RGS9	4.667	0.1944	2
TRAPPC6B	4.667	0.1944	2
TUBB4A	0.6667	>0.9999	2
SCN9A	2.667	0.3611	2
PCLO	0.6667	0.9444	2
GDAP1L1	0.6667	0.9444	2

Table 8: The test results from the Friedman's test, for all the nine tested genes. The dependent variable is the differentiation stage, and the independent variable is the expression level of each gene. Only the CDKN2D gene are statistically significant when analysing the variance in the groups of the different stages of differentiation.

3.1.2 Assessing the reference gene

An additional normalization test of the differentiation stages available was done for the second reference gene, GAPDH. The figure 7 shows the resulting progression through the stages. There should be no fold change in a reference gene, and even though the difference between the groups are non-significant (p-value 0,4, and the test result is 2, Mann-Whitney U), the observed difference seen in the data might occlude the results of the other genes, with which the POLR2A gene was used to normalize. A non-parametric T-test was done here as the data points obtained for the GAPDH gene were not sufficient to conduct the Friedman test.

Although non-significant, the expected value of two perfectly stable reference genes normalized against one another are expected to be, stable. I.e the foundation of using them as normalizers are based on the assumption that they stay stable, like the black dots, and not like the orange ones (figure 7).

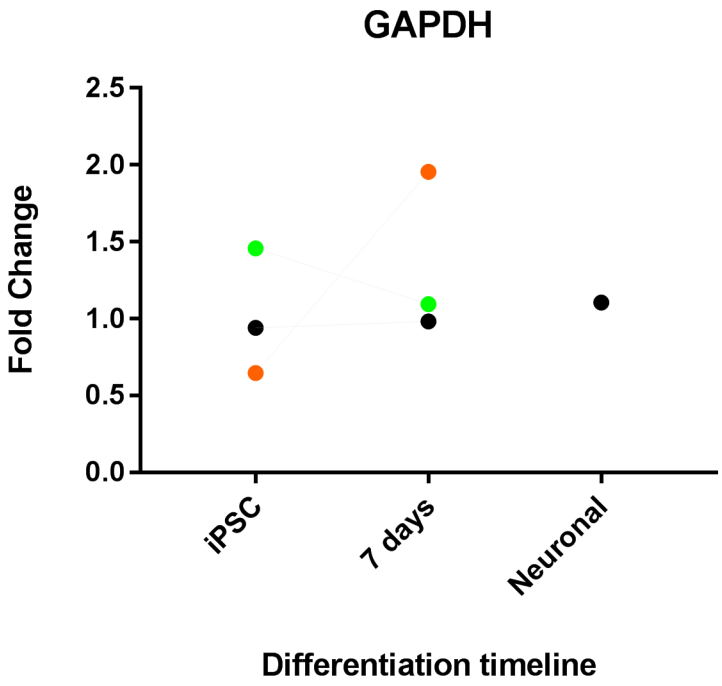


Figure 7: Normalization of the GAPDH reference gene with the use of the POLR2A reference gene. The differences seen in the groups are non-significant when a Mann-Whitney U, non-parametric T-test is ran (p-value 0,4, and the test result is 2).

3.2 Genome editing with CRISPR/Cas9

The procedures in generating a genome edit and verifying it are many, and there are many steps along the way that needs proper controlling if one are to figure out where a potential error occurred. This work is not finished yet, but initial data and information acquired should increase the speed and workflow of how this method can be sorted out and optimized satisfactorily.

3.2.1 Verification: SH-SY5Y and HEK293

In order to test genome editing with CRISPR/Cas9 a neuronal cell line was transfected with the genome editing system. The failed expression/targeting of CRISPR/Cas9 in the SH-SY5Y (this was verified without a positive control), led to the decision to repeat the experiment with HEK293 cells. Before the realization of the problem with the lack of a positive control, the SH-SY5Y results was dismissed as negative results. This faulty conclusion led to a new test on the same samples, and thus the two experiments on two different cell lines are found in the same figure (figure 8). These results are showing the final T7E1 treatment on a gel. The verification by the positive control ensures the working of the T7E1, and thus indicate that the genome editing was not successful for either cell line. Mixing of DNA, PCR (polymerase chain reaction) product, from non-treated WT (wild type) cells and CRISPR/Cas9 treated cells potentially creates heteroduplexes if genome editing has made the cell genomes distinctly different from the WT, detectable with the T7E1. The HEK293 cells are known to be easy to transfect and thus increasing the likelihood of success with such transfection and the following genome editing. The kit manufacturer (Integrated DNA Technologies) also tested the kit using HEK293 cell line [103].

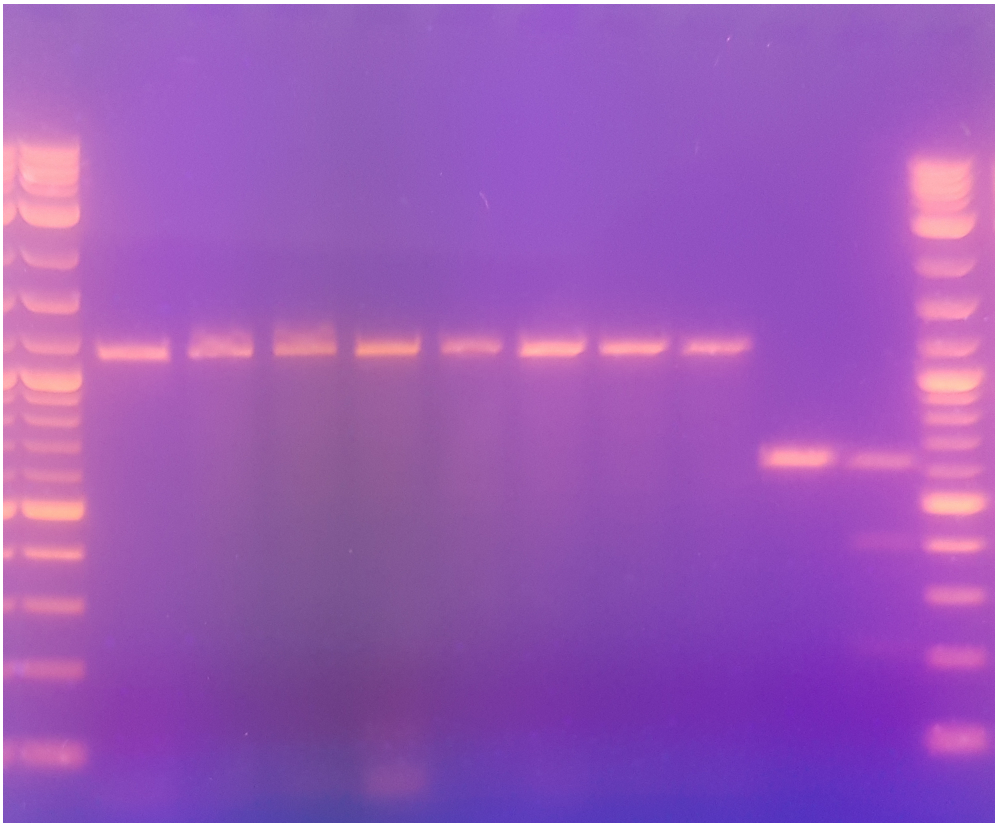


Figure 8: The figure shows a T7E1 test for CRISPR experiment ran on an SH-SY5Y-, and a HEK293 cell line. The two flanking lanes are ladders, then from right to left, the first two (lane 2-3) are the T7E1 controls: T7E1 treated control DNA, control without T7E1 added. The next lanes are DNA from the HEK293 cells from right(lane 4-6): negative CRISPR control, positive control for CRISPR, a negative control with just WT (wild type) DNA. Lane (7-11) are from the SH-SY5Y cell line from right: positive control, negative control with just WT DNA, a negative CRISPR control, and a final lane equal to 7, but without any added T7E1.

3.2.2 Primer tests

In order to test primers used for the PCR amplification of a fragment containing the sequence with the potential genome edit, PCR conditions were optimized using a gradient program. The gradient was conducted to find the optimal amplification annealing temperature of the primer pairs in the PCR kit (Q5 Hi-Fi polymerase NEB). The resulting optimal temperature of 64 °C can be seen in the sixth lane (from right) of figure 9, where a clear band together with minimal unspecific product were obtained. The recommended annealing temperature of 67 degrees, by the manufacturer, delineates the importance of running primer tests [103], although it should be mentioned that the manufacturer's suggested PCR kit is not the same as the one I have used. A gradient run was also conducted for the CACNA1C primers, the results which can be seen in figure 10 reveals an optimal temperature of 65 °C in the third lane (from right).

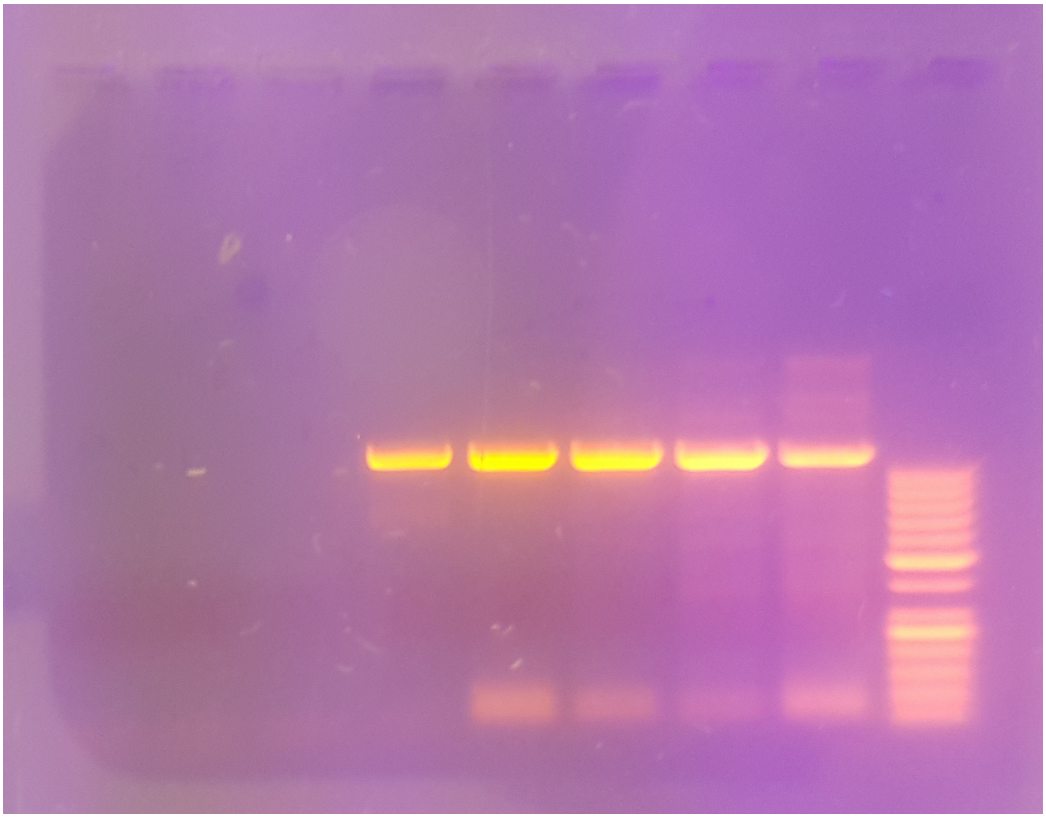


Figure 9: PCR gradient test for finding the functional annealing temperature for optimized PCR runs of the HPRT primers. Figure 2 (From right): lane 1; Ladder, lane 2-9 has a similar, one degree celsius rise per well as in figure 10, which is ranging from 60-67 ºC. Lane 6, 64 ºC has the least unspecific product, and a sharp band, and is used in the rest of the experiments.

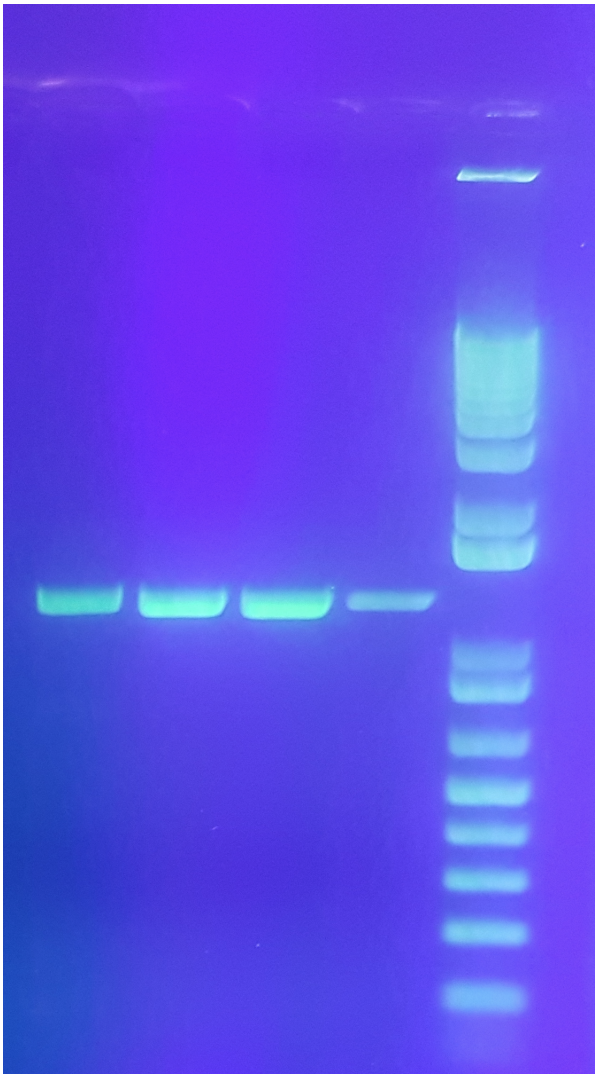


Figure 10: PCR gradient test for finding the functional annealing temperature for optimized PCR runs of the CACNA1C primers. From right: lane 1 is a Ladder, lane 2-5 has a one degree celsius decline per well which is ranging from 66-63 ^{o}C. Lane 3, 65 ^{o}C, has the sharpest band, and is used in the rest of the experiments.

3.2.3 Cell-free cleavage

In order to search the non-working assay for errors, the CRISPR/Cas9 system was assembled and performed on the PCR product as a cell-free test. figure 11 shows a gel with the cleavage product of the HPRT PCR amplicon treated with the CRISPR/Cas9 system with the guide for the HPRT gene. A similar test was done on the CACNA1C PCR amplicon with two of the guides (figure 12).

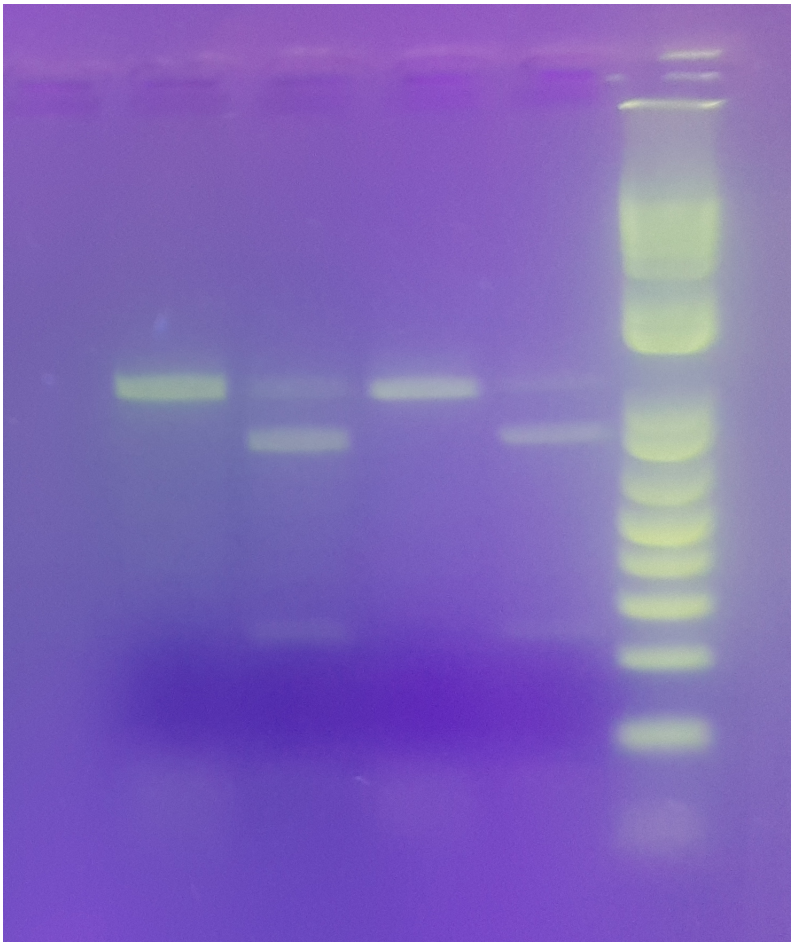


Figure 11: «In vitro» cleavage of the WT PCR product of the HPRT gene with CRISPR/Cas9 to verify the working in a cell free system. The first lane (from right) is a 100 bp ladder, the second lane is in vitro cleavage of the PCR amplified HPRT gene, the third lane PCR amplified HPRT gene, but without the added CRISPR/Cas9 complex. The fourth and the fifth lane is replicates of the 2. and 3. lane.

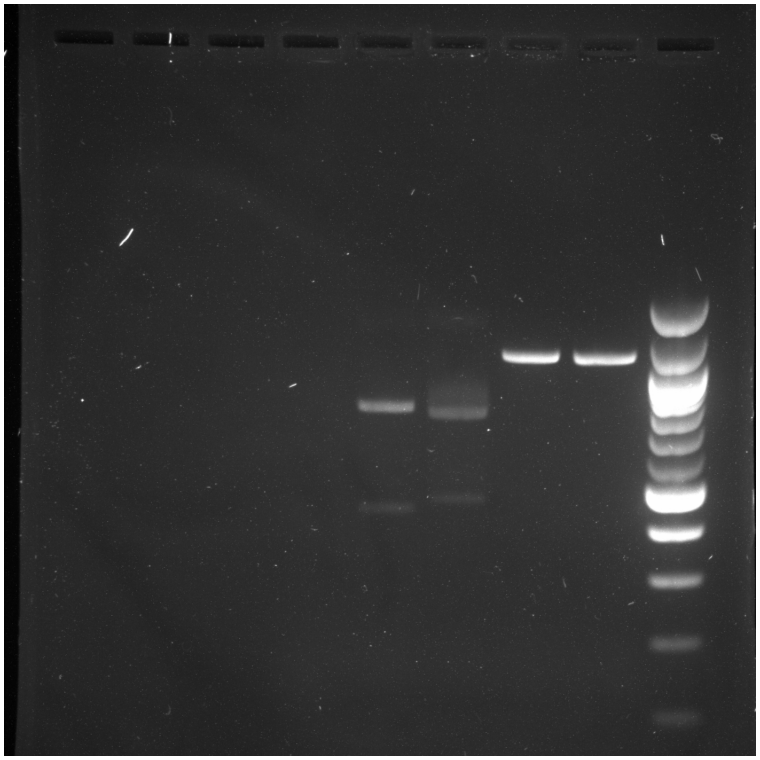


Figure 12: «In vitro» cleavage of the WT PCR product of the CACNA1C gene with two of the CRISPR/Cas9 guides to verify the working in a cell free system (Seq 1 and Seq 3). From right, lane 1: a 100 bp ladder from NEB. Lane 2 and 3: PCR amplicon without the added guide. Lane 4: PCR amplicon cleavage by Seq 1, yielding near total cleavage of the 1222 bp band in to one 491 bp and one 731 bp fragment. Lane 5: in vitro cleavage by Seq 3, also yielding near total cleavage producing two fragments of 428 bp and 794 bp.

3.3 Transfection test

3.3.1 Viability after Lipofectamine treatment

In order to reveal the reason for the unsuccessful CRISPR/Cas9 delivery and expression, a test of cell survival after lipofectamine transfection was conducted. In 20 wells with equally treated SH-SY5Y cells, different concentrations of DNA and lipofectamine were added. The viability of the treatment is shown in figure 13 as the ratio of dead to live cells in one well examined by counting the cells using ImageJ software. Each well represent a green dot in the plot and a linear regression analyses was performed showing an R^2 of 0,7415, a p-value of $< 0{,}0001$, and an over all F score of 51,64. This shows a clear correlation between Lipofectamin concentration and cell viability. For Lipofection treatment with an added 1.2 µl, the viability of the cells fell quite sharply.

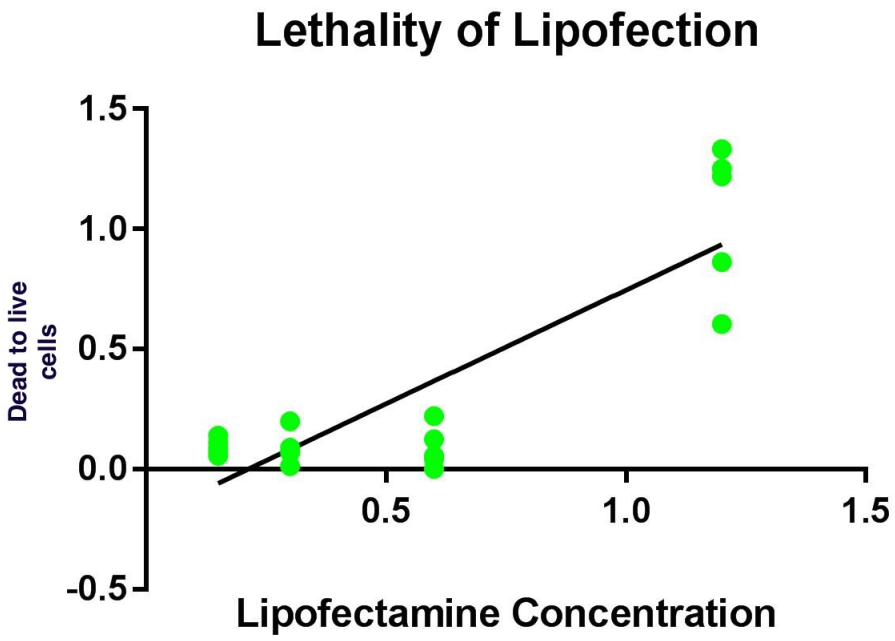


Figure 13: The graph shows the plotting of dead/live cells against the lipofectamine concentration. Every green point is the count of dead cells divided by a count of live cells in one picture taken of a part of a well, of a 96-well plate. A clear correlation with cell death and lipofectamine concetration is seen, with an R^2 of 0,7415, p-value of $<$ 0,0001, and an F score of 51,64.

3.3.2 Fluorescence after lipofectamin treatment

The increased cell death after Lipofectamin treatment (figure 13) raises the question of whether the cells that die are the ones transfected or if there is an equal proportion of cells being transfected independent of lethality. Correlating the flourescent to live cell ratio against lipofectamine concentration, does not show significant correlation (p = 0,2438, F score is 1,452, and the R^2 is 0,07466, figure 14), indicating that the the transfection procedure has the same lethality for cells that ends up being transfected and cells that do not.

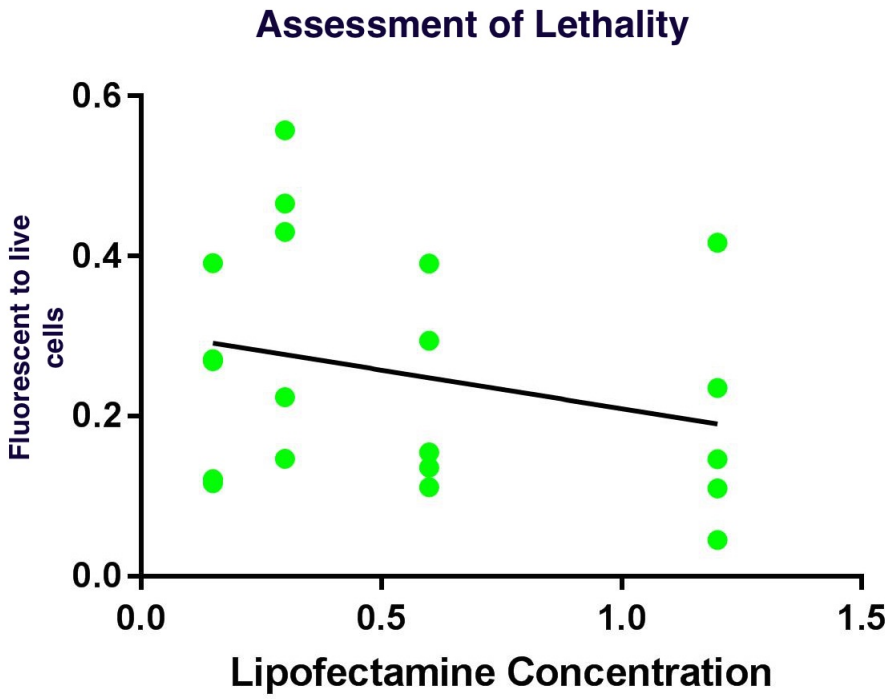


Figure 14: The graph shows the plotting of the ratio of fluorescent to live cells against the lipofectamine concetration, the linear regression has a slope which is not significantly different from zero. Every green point is the count of fluorescence divided by a count of live cells in one picture taken of a well of a 96-well plate. The p value is 0,2438, F score of 1,452, and an R^2 of 0, 07466.

Discussion

This work was an effort to find biological markers that can be used to distinguish poorly differentiated cells from ones that have reached a higher level of maturation. Genes were chosen as such markers based on their up-regulation in highly functional neurons [32]. The absence (except from one gene) of the expected pattern of increased expression levels as cells matured, opens for questions regarding the method. Another aim in this study was to establish the technique of CRISPR/Cas9 and implement an HDR (Homology Directed Repair) template to acquire a precise single nucleotide mutation in the SH-SY5Y cell line as a proof of principle that we, in our lab, could utilize this technique. The proof would not only be shown in a mutation detection assay, but also in the different patch clamp voltage traces obtained from the manipulated cells compared to the non-manipulated ones [104]. Instead of the success story, here comes the troubleshooting of potential things that did go wrong and suggestions on how to, first of all, get the CRISPR/Cas9 system into the cells. The resulting observation that too high lipofectamine concentrations affect the viability of cells, points toward a potential source of error in these experiments.

4 Discussion of methods

4.1 qRT PCR

4.1.1 Normalization

Normalization by means of adjusting variability between samples, is done to account for the differences in amount of transcripts that occur when e.g. two samples have different number of cells. The procedure also normalize for other differences that might affect the results, such as the RT (reverse transcriptase) procedure and/or RNA extraction. Several different normalization techniques are possible to use, such as normalization against total RNA, an internal reference(housekeeping) gene, or an externally added reference [52, 105].

Total RNA which includes rRNA, small RNA constructs, tRNA as well as mRNA, have been shown to vary [106]. The actual mRNA levels only constitutes a small fraction (1-5% according to Thermo Scientific [107]). This makes total RNA content an uncertain representation of the mRNA levels in a cell. The so called housekeeping genes are also found to vary in expression levels [55, 56]. In cancer cells where gene expression is often very dysregulated, this could be a significant problem. Thus, an assay founded on the assumption that some genes have fixed expression may be unreliable. The geNorm tool was built to get around this problem by analyzing several reference genes and measuring variance in expression levels was suggested to predict the most stably expressed ones [55]. However, this procedure has been criticized because the tendencies of genes to be co-regulated [108], meaning that the seemingly equal expression of two genes might not be caused by both being stable. A suggested way to get around this possible hidden variable is to test genes that have widely different functions [108]. The final implementation of several reference genes in the fold change calculation can also mitigate this problem [57]. The third option of adding an external reference has also gained footing [105]. When working with cancer cells and also extreme treatment conditions, housekeeping genes can't be assumed to be stably expressed, and an external reference might serve as a better option. This can also work as a quality test for the reference genes, for the stability of the reference genes themselves [105].

The method for normalization that I have used is not verifying the stability of the selected reference gene (POLR2A). In future experiments I need to include multiple reference genes where

the assumptions of their stable expression across the treatment conditions can be ascertained [109]. The genes chosen preferably should have different functions [108]. Then can a proper assessment of the stability be made with NormFinder, geNorm or the like [110, 55].

4.1.2 RNA quality

Quality of the RNA is difficult to measure, and is one important reason for why normalization of the data is necessary. For every step of the protocol, chances are that one introduces more variation between samples. The sample should be free of proteins, organic molecules, and residual components from the solutions used that are not RNA, this can be checked by spectrophotometry (260nm/280 nm) [111]. A high protein concentration would however have to be present in order for it to be showing on the ratio, as just a small percentage of the amino acid residues are absorbing at 280 nm compared to every nucleic acids absorbance at 260 nm. The 230 nm absorbance will be more informative on the organic impurities, like absorbance from phenol compounds and/or carbohydrate contaminants. This method is solely used to detect purity of samples, i.e the nucleotides conformation or integrity is not measured. Nucleotides might as well be dNTPs, and thus a way of testing the integrity of transcripts is needed. Measures of the RNA samples in this project yielded A260nm/280 nm close to 2 which is good values for single stranded RNA, one sample (control 2) was however off (1,52) supporting the decision of removing this sample from the analysis. The ubiquitous RNases impair the quality of the sample if present, especially for long-term storage. Then freeze thaw cycles should be kept at a minimum. Other factors from the extraction can influence the RT (reverse transcription) or the PCR amplification [112, 113]. I froze cells immediately and only a necessary transportation of the samples between the lab of Srdjan and ours were conducted. When the RNA was isolated it was thawed on ice to limit degradation. The isolated RNA which was not used immediately was quickly frozen to -80 degrees after the procedure.

Sufficient amount of substrate is of course essential, but when a signal can be detected after RNA isolation, consistency of the protocol is the most important aspect to consider. Using the same kit for all the samples, reliable results can be obtained with samples that even have undergone partial degradation [114]. Thus the quality *per se* is not considered a major issue as fluctuations usually can be neutralized by normalization. Despite this, detection of small differences in expression levels, and especially low levels of transcripts, might be affected more by degradation. That is, an uneven degradation of different transcripts in a sample with only 10 copies, constitute a bigger ratio of total transcripts than do one with 500 copies. Thus the recommended quality test of the RNA integrity is standard procedure for a good qPCR experiment according to the miqe-guide by Hugget end co-workers (2013) [115].

The limited time I had at my disposal for learning and design my experiments in this phase of the thesis, and the additional lack of knowledge on how to practice the technique in our group, has impacted the experiments and some realizations have first announced their presence after the experiments were finished. It has especially affected the RNA integrity testing. For future experiments two procedures will be included. First, the RNA amounts would be diluted to try to obtain more similar concentrations for the testing of the RNA quantity at the nanodrop spectrophotometer, to see if this remains linear over variable concentrations. Although this has been shown to be relatively stable [116], variability is potentially introduced in the pipetting of very different volumes. Second, the integrity tests should be conducted although that might require some acquiring of equipment that we do not possess in the lab. One method that can be used is based on the ratio measure between different subunits of rRNA [117]. Another technique is based on using a housekeeping gene to test ratio of RNA on the 3' versus the 5', this might serve as a better test for mRNA integrity testing [118, 119] .

4.1.3 cDNA synthesis

The generation of cDNA from RNA can be done in several different ways and the selected primers used are important for the specificity of the assay. Random primers and short fragments that have an unspecific binding capability to many sequences, are frequently used [51]. This feature is thought to convert a representative ratio of all RNA in the sample. A problem with this technique is the potential loss of low expressed transcripts, as conversion of RNA that are of no interest when studying expression, competes with the conversion of the mRNA that we are interested in. Another common alternative is using oligo(dT) primers that bind specifically to the poly-A-tail of mRNAs [51]. This, however, produces a bias toward the final 3' end of the mRNA because degradation of these molecules results in copying of the part still attached to the poly-A-tail. A final option is to use specific primers, as in a regular PCR but this requires large amounts of RNA substrate as every test needs to be done on the RNA sample, rather than the cDNA sample. In the experiments presented here a mix of random primers and oligo(dT) primers were used. Signals were obtained with all the primer pairs, and thus sufficient RNA to DNA conversion was assumed to have occurred.

More and more common is the use of a mixture of oligo(dT) and random hexameres [120, 121]. This mitigates the problem of loosing weakly expressed transcripts in the ocean of rRNA [122, 123], whilst copying of fragments that have lost the poly-A-tail is ensured. This technique however, still requires careful design of primers at the 3' part of the mRNA. Additionally and maybe most important of all is the consistent use of the same kit or even the same lot for samples that is compared as it is easily imaginable that variability may occur in this step if the ratio of the primer types differ. All this points in the direction, lack of knowledge ascertains the need for caution. The mixture ratio of random- to oligo(dT) primers might influence the results as random primers are shown to vary and overestimate mRNA levels [124]. The problem of random primers binding at several places, leading to short cDNA constructs, will happen using mixed primers as well as with just random primers alone. In this thesis relatively short fragments were generated in the qPCR run, making this a minor problem.

4.1.4 Primer design

Although a lot is known about the genome, and we can take clear measures to optimize primers by design, the only real assurance for the quality of a primer is gained from running tests. The design is however prospecting some important aspects which can rule out obvious problems. First, a short fragment ensures an efficient replication, and enables quick cycle steps of temperature exposure in the PCR cycles. Polymerases have certain rates at which the fragment is generated, and thus time and fragment length dictate the length of the cycle. Short cycles decrease unspecific products, and possible copying of gDNA containing an extra intron in between the sites of primer annealing. An effort to design equal amplicons on all genes tested lowers the variability in the assay, as the same cycle time applies to all reactions on one plate. Second, the annealing temperature of the two primers used should be similar. A specific amplicon is generated by use of specific primers, and precise annealing temperature helps generate the specificity [125]. Since qPCR runs often implement multiple primer pairs a limited difference between melting temperatures, not only for the two primers, but also for primer pairs are ideal. This limits the problem with the unavoidable difference in optimal annealing temperature for all simultaneuous reactions on one plate. Third, tools that design primers also take in to account the possibility of self-annealing primers, as well as the formation of primer dimers [99, 126], features of DNA that might hamper with the efficiency of the primer, and the fluorescent signal. A plethora of design tools exist, the eurofinsgenomic's web entry based on Irv Edelman's algorithm [99] was used for the design of the qPCR primers described in this thesis. The National Center for Biotechnology

Information (NCBI) database was used to acquire sequence information about the 3' most parts of the transcript and this part of the sequence was used in the designing of primers [127].

4.1.5 Avoidance of gDNA

The exclusion of gDNA is important and is ensured in several steps. First by the DNase addition in an RNA purification procedure and then by a gDNA wipeout solution in the making of cDNA (see materials and methods). Additionally, primers are designed to span an exon-exon junction with an intron preferably so big that copying of the total genomic region including the intron is less likely and avoided or limited by the aforementioned short cycle step. A final assurance is made by running a no-reverse transcriptase (noRT) sample along with the RT-treated cDNA sample as a negative control. The no-RT sample can also be indicative of primer dimers as the dimerization would produce fluorescence by the binding of the SYBR green component. This is however better tested in a template free sample which is done in an initial primer test where also primer efficiency is measured.

4.2 Cell culturing

4.2.1 Subculturing

In the recommendations from the ATCC (American type culture collection) regarding the SH-SY5Y cells, there is a centrifugation step in the cell splitting procedure. This cell line is a semi-adherent line and free-floating cells should ideally be collected by centrifugation, if a robust representation of the cells are to be passaged. As the goal of this thesis was to patch clamp the cells, this was deliberately omitted as the non-adherent cells are non-patchable. This could potentially result in a selection pressure that bias the results.

4.2.2 4.1.2 Cryopreservation

In the first round of cryopreservation I added the DMSO to cells in pre-heated medium by mistake. The permeating properties of DMSO might facilitate the influx of any substances in the surrounding media, with potential harmful consequences [128, 129]. Although removal of antibiotics in the cryoprotective medium is done as a precaution for this, the cold treatment of DMSO is what is usually done to minimize this flux [130]. The DMSO in combination with other solutes protect the integrity of the membrane by vitrification rather than forming crystals [131]. For this vitrification, DMSO's permeating properties are important when freezing, but it is a double edged sword because of the increased permeability of the membrane when temperature rises. A centrifugation step is part of the resuscitation protocol to remove DMSO before culturing.

An additional round of cryopreservation was conducted to ensure healthy vials with viable cells, and the vials of the first round was marked. Despite this mistake, the viability of the cells did not seem to have decreased much, and resuscitation of these vials have provided viable cultures.

4.3 Genome Editing

4.3.1 CRISPR/Cas9 editing

The promise of a footprint free treatment of cells, and a lowered off-target rate makes the protein transfection a tempting option for the CRISPR assay [132, 133, 134]. The nuclear localization signals (NLS) brings the complex to nucleus, and no transcription-translation procedure is delaying the genome editing [135]. After some time, all cas9 proteins will be degraded and the

nuclease will stop cleaving. The cheap generation of protein, and even cheaper construction of the guideRNA, enables quick generation of different assemblies that can target any site in the genome [60]. This also opens up for quick multiplexing of many targets which is a great advantage when dealing with MFD (multifactorial disorders) which is the goal for our assay.

The down side of protein transfection is that you lack an indicator of success rate of transfection whereas a plasmid can implement a GFP (green fluorescent protein) and even a selectable marker to identify and select for transfection efficiency and genome editing [136]. Visualization of transfection in combination with fluorescent-activated cell sorting (FACS) may greatly improve the efficiency for considering transfection and genome editing successes, enabling a faster screen and identification of cells to be tested for function by other methods.

In my study I have found that the protein based method raises far more problems than advertised by the kit manufacturers and the promises of a new technique. This have made me skeptical of the advantages of implemented markers over the advantages provided by a simple system. The low success rate of a HDR (Homology directed repair) event in a cell requires high success rates in the other methods leading up to the editing [69, 73]. The low rate is constantly being improved in an effort to increase efficiency by methods ranging from inhibition of the NHEJ (Non Homologous End-joining) system [137], to improving the HDR (Homology Directed Repair) system by manipulating the HDR template [138]. The specific editing of cells in a SNP fashion is a difficult endeavor, but despite that i did not come to undertake this in my thesis, this must be the goal for proper investigation of the SNPs relevant for the psychiatric disorders with which we want to study.

4.3.2 Delivery method

Electroporation was tested using a BTX ECM 630 electroporation machine. In my hands, electroporation harmed the cells and left almost no surviving cells thus low yield of transfection was obtained. This made me try lipofectamine instead. I now realize that electroporation optimization protocols can improve the results, and that more modern equipment (Neon transfection system) comes with preprogrammed, built-in, optimization protocols [139], protocols that should be tested before dismissing a method. In contrast, other work has actually shown electroporation to be more efficient and consistent with different cell types than lipid based transfection [93].

Transfections with viral vectors are known to have high transfection efficiency. However, because the long term goal of the present project is to manipulate differentiated iPSCs we decided to aim for the least potential stress for the cells. Viral vectors may induced immune responses [140] or in itself be toxic for the cells [141]. Moreover, the acquiring of viruses is expensive, because they must be specially packed and titrated for every new target sequence [90]. It takes less effort to generate a plasmid with the proper sgRNA [142], but least of all is the effort of buying a ready made 36 nt (nucleotide) crRNA. When Cas9 and tracrRNA is acquired, the only thing needed for targeting a new site is the 20 nt target sequence with the added tracrRNA-complimentary region of 16 nt [103]. Currently viral vectors have limited carrying capacity and greater sizes can be obtained by the plasmid-only transfection [143, 125]. The opportunity of greater carrying capacity is important when studying MFDs (multifactorial diseases) as combinations of several SNPs may be of interest. This can be acquired by insertion of several sgRNAs and/or one or more donor DNA templates of more substantial length.

The reason for choosing lipofectamine over electroporation was not founded in thorough testing of the electroporation technique and that might be necessary before assurance of which technique is best suited for the assay. The lipofectamine transfection test reveals shortcomings with the way the lipofection protocol was carried out, and additional testing of a more optimal concentration of lipofectamine should be conducted before a certain conclusion can be drawn.

4.3.3 CRISPR/Cas9 verification assay

In order to verify if the genome editing has worked it is necessary to run some sort of verification. Verification assays of genome edits are somewhat different with regards to what change is incorporated. Large deletions can be identified with gel analysis that shows difference in DNA fragment lengths based on whether or not there is a deletion [144]. The SNP mutations in an HDR experiment can be detected by inserting a silent mutation to generate a restriction site which later can be used to verify an edit [89]. For short InDel mutations, the heteroduplex formation procedure followed by enzymatic cleavage by specific mismatch detecting endonucleases, can split a fragment and distinguish a true heteroduplex and an actual edit. In this case, the T7E1 (T7 endonuclease 1) has been shown to best cut mismatches of more than 2 bp, whereas a surveyor nuclease (another commonly used endonuclease) is more efficient with single mismatched duplexes [145]. Since the NHEJ results in changes that adds 0-25 bps, and the cutting by the CRISPR/Cas9 complex tend to cut until it no longer recognizes the target sequence, the resulting heteroduplex of mixed mutant and WT (wild type) DNA will mostly carry more than 2 bp mismatches. Thus the best endonuclease enzyme for genome edit verification will be the T7E1 [145, 75]. The results are often verified by sequencing as well [29] but the endonuclease method will also give information on the efficiency of the genome edit [146, 147]. My goal was to identify the best of three sgRNAs to choose one for the proceeding HDR experiments, and thus enzyme verification was chosen over sequencing.

Although to a lesser extent, the T7E1 can cleave a difference generated by an introduction of a mismatch in an early cycle of the PCR by an inaccurate polymerase, and does not distinguish this from a genomic difference acquired in the CRISPR assay. High fidelity polymerases are thus commonly used to limit risk of poor amplification [148, 149]. The known length of the fragments generated from a CRISPR experiment might be used to exclude such a mishap, but this is by no means guaranteed. Reports of *in vitro* recombinant events is also a possible source of error but can be mitigated by lowering cycle number or input DNA concentration [150]. The priming of a truncated product generated in too short extension cycles or a polymerase starting before ideal temperature has been reached might serve as inducers of variability and recombinant generation of such chimeric PCR products [150, 151]. One way to avoid this is by using hot start polymerases where antibodies inhibit replication until a certain temperature is reached [151].

The EnGen kit from NEB used late in the present study provides a hot start polymerase. In addition it contains a positive control for the T7E1. This is essential for proving negative results. Earlier in this work, I was not aware of this kit with control for the T7E1. Evidently, much of my work was trials with the aim to acquire a positive result with no means of verifying the negative ones. The kit was ordered once I realized this and it turned out that this kit also spared the need for the laborious experimental procedures with ethanol precipitation to purify the PCR product as discussed below section 4.3.5.

4.3.4 PCR Primers for genome edit analysis

Primers used were designed to yield an 800 bp fragment (recommendation was 600-1000 bp [103]) of the CACNA1C gene. When designing more than one pair of primers, it turned out that two primers across the designed pairs worked better together (data not shown). The product ended up being 1222 bp, quite a bit longer than the recommended maximum. Since I did not get a good result from the control kit which had a more optimal PCR fragment length, this possible error source was not pursued any further. When designing the primers for PCR, the following considerations were taken: (1) primers were designed so that the CRISPR induced mutation was not present too close to the edges of the fragment (recommendation was >100 bp from the edge) because if the fragment cleaved by the T7E1 is too short, it would travel far in the gel and possibly

be difficult to detect and (2) the center nucleotides of the fragment were avoided as a centered mutation would produce two equally long bands that would be impossible to distinguish from each other. A good separation on a gel is essential for a successful experiment and should yield three separate bands; the parental band and two additional ones generated from the cleavage of the parental one. The visualization is especially important if a quantitative evaluation of the bands were to be done. For the HPRT control gene the expected fragment sizes would be about 1083 bp for the parental band and about 827 bp and 256 bp for the cleavage product. The length might differ slightly as different mutational outcomes will produce variable lengths.

4.3.5 Ethanol Precipitation

A precipitation protocol is conducted in order to clean up the PCR product for molecules incompatible with the T7E1 enzyme. While this was essential for experiments before commercial kits became standard to use, consistency of measurements and thereby the efficiency and quality of data have improved using commercial products. To familiarize with generic protocols might however illuminate the "black box" experience commercial kits often gives. The problem with this non-commercial protocol became the decanting step, which potentiated a clear risk of loosing the whole PCR amplified product. A careful pipetting technique was adopted instead, sucking up the supernatant from the bottom and avoiding the side of the eppendorf tube where the tiny pellet would be. As mentioned above, I acquired the EnGen kit late in the experiments and this was a clear time and effort efficient investment, as the compatibility between the PCR reagents and the T7E1 made the Ethanol Precipitation protocol superfluous.

5 Discussion of Results

The establishment of a technique was a goal of this thesis work, and to plan how to use it to introduce single nucleotide mutations into the genome of cells. As I got stuck, with not managing to transfect the cells, this goal was not fulfilled. The other important goal of this work was to detect measurable differences between stages of differentiated neurons. A means of having a quality control of neuronal state, before studying differentiated cells assumed to be neuron-like. Also this work has turned out incomplete as the scientific literature was reviewed.

5.1 qRT PCR

First we asked whether it was possible to use quantitative PCR to identify the functional state of a neuronal cell population differentiated from iPSCs, based on the expression rates of a set of genes known to correlate with the functional state of neurons [32]. In order for such a procedure to be useful, we depended on the existence of a clear difference in expression levels of genes from one functional stage to the next. If this is the case, the expression fingerprint can be used to define the maturation stage of neuronal populations. This could either be a threshold for the combination of genes or single genes tested with the importance being its predictive power. Since only a single gene showed significant difference between differentiation stages based on expression (CDKN2D figure 5), this is the only gene with potential predictive value. The gene inhibits the cell cycle progression, and from neuronal cells we know that division ceases when a mature stage is reached, making the up-regulation of this gene in mature cells biologically reasonable [34].

In addition to the necessity of gene expression to be specific to maturation stages, they also need to be distinguishable. One data point obtained from a sample should be predictive of the state, and no ambiguity of which group it belonged to can be accepted. We see that for one time point for SCZ sample 2 figure 5 (green dot representing the neuronal stage), this is not the case. If a threshold value were set at the highest value obtained from the iPSC group (orange - iPSC sample), this green time point would fall in the iPSC group and not the neuronal group.

This raises the question if the test populations were representative of a high functional state of neurons. The only quality control we had in this experiment was the time of differentiation, a factor that previously has been shown to be highly variable [32]. We had to trust the research group of Djurovic who delivered the samples and who did additional morphological verifications of the cells. Their verification may have been limited to visual inspection of the cells and more objective measures should be applied. Ideally time of differentiation should be the only initial information needed, and the qPCR results could enable verification of the cellular state. If we are to make an objective measure of maturation stage we are dependent on a trustable base line condition obtained from before differentiation and from mature neurons. Without trusting that the samples we obtained were from mature neurons it is difficult to conclude if the assay can work. To address this question, electrophysiological measurements of the putative mature cells must be acquired. This can be a means to verify their matureness in addition to data from transcriptomics, immunohistochemistry and so on. This is impossible to do with the samples already treated, but could be implemented in a future test. Moreover, could the way that the qRT PCR experiments were conducted affect the results? The question is discussed in detail in the following section, and in section section 4.1

5.1.1 Reliability of the qPCR experiment

The reliability of qRT PCR results obtained in this thesis is likely questionable. As discussed in section 4.1, stability of a reference gene is essential for how trustworthy the results are, and thus verification of stability is essential. The reference gene, POLR2A, used was recommended

by an experienced researcher in a neighboring research group at the Department of Biosciences. However, it turned out that the researcher had only used this references gene for analysis of different tissues and cell lines compared to the present study. Therefore, the POLR2A may not be suited as a reference gene for the cells used and therefore may be unsuited as verification of stability.

The subsequent effort to investigate the quality of this reference gene was done by trying to normalize against another reference (GAPDH). I had initially chosen, GAPDH as the reference gene, before I was told by the same researcher that the gene is known to vary and is therefore not suited to test for stability. Variability of GAPDH has been reported in the literature [55, 56]. Other reports also clearly states that one should select genes that are stable in the specific tissue or cells of study because the variation between samples may be largely dependant on the treatment [55, 152, 109, 153]. Even the common use of a standard set of reference genes, to test which is best, is criticized as other genes might serve as better normalizers [109]. In retrospect, I should have reviewed the literature more thoroughly but due to time constraints and my supervisors' lack of knowledge and therefore guidance, I mistakingly trusted the expertise next door.

The result from the normalization of the reference gene with respect to another reference gene, when nothing is known about their stability, can at best only be an indication of the stability between the two. Both genes can be unstable, one may be while the other is not or both can be stable or co-related. Only similarity between the two may potentially be predicted from the test used for the results in figure 7. However, since it cannot be claimed that the expression of the reference genes (POLR2A and GAPDH) are stable, I have no means of knowing how the chosen reference gene is deviating from true stability.

The difference of the CDKN2D gene between iPSC and neuronal groups can, considering the uncertainty of stability of the reference, either be "pushed" by the normalization and become significantly different although it is not. Another alternative is that the other eight genes (or some of them) may in reality show difference between groups because they could have been equalized by the improper normalization, and perhaps all or some of the genes actually would show significant difference if verified reference genes were used. This requires proper verification of the stability of adequate reference genes and increased sample sizes from the three differentiating stages before any conclusions can be drawn.

The work by Bardy and co-workers (2016), found that the nine genes tested in the present study distinguished functional states of neuronal cells, not iPSCs or seven days differentiated cells. It may very well be that seven days old differentiated neurons are far from the maturation levels of neurons investigated by Brady and co-workers. This questions if at all the nine genes selected would be likely candidate genes for the identification of maturation stages. The poor choice of method may be explained by constraints in obtaining samples and also the efforts to do qPCR measurements of a series of 8 stages of increasing neuronal developmental state. I received my iPSC samples and differentiated neuron samples from our collaborator at OUS but during the process, it became clear that there was a lack of knowledge concerning the experimentation and pretreating of the samples before delivery. Samples were lost due to bad preparation, mislabeling and wrong deliveries compared to what was conveyed to me unfortunately contributing to less meaningful results.

A final issue with the study design is the use of both control cells and cells from patients with SCZ for the initial testing described here. Emerging evidence points in the direction that SCZ may be caused by developmental disturbances [154, 155, 156]. The control assay described in this thesis should be thoroughly tested and established on control cells before testing of cells from SCZ patients. It is likely that the genes to be used for identifying differentiation stages may turn out to have different expression patterns in the two groups. A GWA study have found correlation with one (PCLO) of the genes used in in this expression analysis [48]. If such

differences are found, expression analysis might be included in the final assay. Then, however, it will be used to distinguish groups rather than working as a quality control. When considering the potential differences between these groups the procedure of generating stem cells, which might be considered removing the epigenetic fingerprint, may substantially change the difference that we seek to find between the groups [157]. The advances in techniques of differentiating cells where direct differentiation without the iPSC transition state, might be a way to maintain the epigenetic fingerprint [158]. This must however not undermine the possibility of producing differentiated neurons that have an epigenetic profile reminiscent of a fibroblast profile (or other type of cells that are initially taken from controls or patients to be differentiated), rather than a complete neuron. These problems warrants the investigation of the epigenetic fingerprints of iPSC derived neurons, SCZ patients cells, and the neurons differentiated directly from fibroblast (or other cell type).

5.2 CRISPR/Cas9 Editing

By understanding and utilizing genome editing techniques, we hope to be able to use that in the study of variants' influence on cell function. In the finalized assay we envision optical electrophysiology done on these iPS cells and geneomic engineering of the cells by CRISPR/Cas9. Our limited access to stem cells and their relatively high fragility, made us think that the technique best could be established on a more robust cell line with excitable properties.

Several attempts to transfect cells with the CRISPR/Cas9 complex and conduct genome editing were carried out. The final representation in figure 8 shows a faint band in the positive control lane indicating that the T7E1 did cut the mismatch and thus mismatches between heteroduplexes in the other lanes were not present, or not obvious enough to show on this gel. If the clarity of the assay is this poor for the control, it is clear that it needs optimization. A 100% efficient CRISPR/Cas9 editing would result in a mixing of equal amounts of WT duplex and duplexes deviating from the WT, in the heteroduplex formation. Thus 100% efficiency would be the equivalent of the control lane and not much wiggle room for poorer efficiency exist for the assay's detection ability. The use of a capillary electrophoresis machine (Agilent 2100 Bioanalyzer) have been shown to detect as poor as 5% editing efficiency when used with the T7E1 [145], and might serve as a solution to this low detection ability.

Even though the control kit from the Integrated DNA Technologies (IDT) providing the CRISPR/Cas9 protein complex did not result in genome editing, it should not be judged as a poor method. Other factors might, and probably did, occlude the results. The HEK293 cells were tested as a cell type positive control for the kit, as this cell line was recommended by the kit manufacturer who had got it to work. Since I was unable to obtain any signs of editing with the HEK293 cells other factors that I have failed to identify may have caused the lack of editing. From the lack of results from experiments with the HEK293 cells, it is impossible to say that my lack of success with the SH-SY5Y cell line was due to the kit. The most probable source of error lays in the hands of the experimenter. In an effort to rinse those hands of any blame, an optimization protocol for transfection of GFP plasmids was conducted, Transfection with lipofectamine.

5.2.1 Transfection with lipofectamine

The resulting evidence of toxicity of the Lipofectamine points toward loss of efficiency of the CRISPR/Cas9 system when using too high concentrations of the membrane fusing agent. The idea of doing such an experiment was born out from the discovery that the IDT (provider of the CRISPR/Cas9 system) protocol, and Lipofectamine protocol have large discrepancies with

regards to the recommended concentration of Lipofectamine (they suggests 1,2 µl and 0,15 - 0,3 µl, respectively).

Although the ratio of transfected to live cells seems to be the same for all groups of different lipofectamine concentration, the high lethality surely inhibits transfection just by killing cells and lower the total amount of cells substancially. This implicates that the transfection efficiency could be highly reduced, but it should not be absent. It should be mentioned that different lipofectamine products were used, although the lipofectamine are the same, they differ with regards to one additional reagent (P3000 for Lipofectamine 3000, and CRISPRMAX Plus reagent for Lipofectamine CRISPRMAX). This could imply differences between the two. It is difficult to acquire information on the contents of these proprietary compounds, and so it is difficult to know the validity of generalization of these results across differnt products . It should in this context also be mentioned that the IDT recommandations is the use of the RNAiMAX (Lipofectamine variant) for the transfection, and not the CRISPRMAX specially made for transfecting CRISPR/Cas9 complexes, which I used in these experiments. Even so, all these protocols from Thermo Scientific are equal regarding the volume added for the lipofectamine reagent, namely $\sim 0,3$µl for one well of a 96-well plate.

The combined effect of the poor quality of the aforementioned CRISPR verification assay, and this possible inhibiting effect of the Lipofectamine concentration on the transfection efficiency, might have added up to obscure the working of the assay. It is in the nature of negative results, to be difficult to establish compared to positive ones.

6 Future Perspectives

The neurons-in-a-dish studies needs a functional state quality control (QC). The QC might consist of expression analysis of the genes studied in this thesis, but verification of the procedure must be ensured to have a high quality and 100 % predictive power. When studying delicate differences between groups, it is important to rid the assay off all the confounders that potentially might affect the result, like the different functional states of immature neurons. Although it must be considered that such states may be the mere outcome of the differences as well, making it a delicate piece of work.

We imagine an assay that is based on patient cells grown in a dish, transfected with indicators allowing both stimulation and detection of activity by use of light. In such an all-optical-electrophysiology system, we believe that features can be studied based on communication between cells, but also with regards to development. Developing a scoring table for how a dish is "behaving", one can start to compare schizophrenic versus control cell based dishes. Then by introducing genome editing, correction of some of the mutations found to correlate with disease in the GWA studies can be done. The further comparing of rectified cells versus patient cells might be hard to detect, but by introducing computational power, we hope that algorithms containing variables representative of the SNP's effect on the cell, can be manipulated to have a predictive power of what is to be seen in the dish.

A ranking system must be developed for the SNP mutations, including synergistic effects of specific combinations of SNPs. When this is worked out, drug's efficiency on patients must be compared to variants, to find correlation between effective drug and SNP variants. In combining this knowledge, diagnostic tools can be developed to subdivide disease into groups representative of the most severe SNP or combination of SNPs carried by a patient. As this knowledge about the SNPs is elucidated therapeutic targets can be developed based on variants, and the end goal would be to utilize the assay as a diagnostic tool for the prescription of the best suited therapeutic for each patient.

References

[1] Coryell, W. & Zimmerman, M. The heritability of schizophrenia and schizoaffective disorder: a family study. *Archives of general psychiatry* **45**, 323–327 (1988).

[2] Purcell, S. M. *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–752 (2009).

[3] of the Psychiatric Genomics Consortium, S. W. G. *et al.* Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).

[4] WHO. WHO | Mental Disorders (2017). URL `http://www.who.int/mediacentre/factsheets/fs396/en/`.

[5] WHO, G. Who methods and data sources for global burden of disease estimates 2000-2011. *Geneva: Department of Health Statistics and Information Systems* (2013).

[6] Insel, T. R. Assessing the economic costs of serious mental illness (2008).

[7] Association, A. P. *et al. Diagnostic and statistical manual of mental disorders (DSM-5®)* (American Psychiatric Pub, 2013).

[8] Lieberman, J. A. *et al.* Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *N Engl j Med* **2005**, 1209–1223 (2005).

[9] Carlsson, A. & Lindqvist, M. Effect of chlorpromazine or haloperidol on formation of 3-methoxytyramine and normetanephrine in mouse brain. *Basic & Clinical Pharmacology & Toxicology* **20**, 140–144 (1963).

[10] Liebzeit, K. A., Markowitz, J. S. & Caley, C. F. New onset diabetes and atypical antipsychotics. *European Neuropsychopharmacology* **11**, 25–32 (2001).

[11] Miller, D. D. *et al.* Extrapyramidal side-effects of antipsychotics in a randomised trial. *The British Journal of Psychiatry* **193**, 279–288 (2008).

[12] Kessler, R. C. *et al.* Age of onset of mental disorders: a review of recent literature. *Current opinion in psychiatry* **20**, 359 (2007).

[13] Consortium, . G. P. *et al.* A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).

[14] Hinds, D. A. *et al.* Whole-genome patterns of common dna variation in three human populations. *Science* **307**, 1072–1079 (2005).

[15] Johnson, R. C. *et al.* Accounting for multiple comparisons in a genome-wide association study (gwas). *BMC genomics* **11**, 724 (2010).

[16] Metzker, M. L. Sequencing technologiesâthe next generation. *Nature reviews genetics* **11**, 31–46 (2010).

[17] Masters, J. R. Hela cells 50 years on: the good, the bad and the ugly. *Nature Reviews Cancer* **2**, 315–319 (2002).

[18] Mayshar, Y. *et al.* Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. *Cell stem cell* **7**, 521–531 (2010).

[19] Martin, G. R. 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* **78**, 7634–7638 (1981).

[20] Hikabe, O. *et al.* Reconstitution in vitro of the entire cycle of the mouse female germ line. *Nature* **539**, 299–303 (2016).

[21] Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *cell* **126**, 663–676 (2006).

[22] Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *science* **318**, 1917–1920 (2007).

[23] Park, I.-H. *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141–146 (2008).

[24] Malik, N. & Rao, M. S. A review of the methods for human ipsc derivation. *Pluripotent Stem Cells: Methods and Protocols* 23–33 (2013).

[25] Chun, Y. S., Byun, K. & Lee, B. Induced pluripotent stem cells and personalized medicine: current progress and future perspectives. *Anatomy & cell biology* **44**, 245–255 (2011).

[26] Ebert, A. D. & Svendsen, C. N. Human stem cells and drug screening: opportunities and challenges. *Nature Reviews Drug Discovery* **9**, 367–372 (2010).

[27] Yu, X. *et al.* Improved delivery of cas9 protein/grna complexes using lipofectamine crisprmax. *Biotechnology letters* **38**, 919–929 (2016).

[28] Tang, Z.-H. *et al.* Genetic correction of induced pluripotent stem cells from a deaf patient with myo7a mutation results in morphologic and functional recovery of the derived hair cell-like cells. *Stem cells translational medicine* **5**, 561–571 (2016).

[29] Wu, Y. *et al.* Correction of a genetic disease in mouse via use of crispr-cas9. *Cell stem cell* **13**, 659–662 (2013).

[30] Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *cell* **131**, 861–872 (2007).

[31] Morton, A., Hammond, C., Mason, W. & Henderson, G. Characterisation of the l-and n-type calcium channels in differentiated sh-sy5y neuroblastoma cells: calcium imaging and single channel recording. *Molecular brain research* **13**, 53–61 (1992).

[32] Bardy, C. *et al.* Predicting the functional states of human ipsc-derived neurons with single-cell rna-seq and electrophysiology. *Molecular psychiatry* **21**, 1573–1588 (2016).

[33] Belinsky, G. S., Moore, A. R., Short, S. M., Rich, M. T. & Antic, S. D. Physiological properties of neurons derived from human embryonic stem cells using a dibutyryl cyclic amp-based protocol. *Stem cells and development* **20**, 1733–1746 (2011).

[34] Okuda, T. *et al.* Molecular cloning, expression pattern, and chromosomal localization of human cdkn2d/ink4d, an inhibitor of cyclin d-dependent kinases. *Genomics* **29**, 623–630 (1995).

[35] Sauter, A. & Rudin, M. Determination of creatine kinase kinetic parameters in rat brain by nmr magnetization transfer. correlation with brain function. *Journal of Biological Chemistry* **268**, 13166–13171 (1993).

[36] Glassmann, A. *et al.* Developmental expression and differentiation-related neuron-specific splicing of metastasis suppressor 1 (mtss1) in normal and transformed cerebellar cells. *BMC developmental biology* **7**, 111 (2007).

[37] He, W., Cowan, C. W. & Wensel, T. G. Rgs9, a gtpase accelerator for phototransduction. *Neuron* **20**, 95–102 (1998).

[38] Rahman, Z. *et al.* Cloning and characterization of rgs9-2: a striatal-enriched alternatively spliced product of the rgs9 gene. *Journal of Neuroscience* **19**, 2016–2026 (1999).

[39] Jones, S., Newman, C., Liu, F. & Segev, N. The trapp complex is a nucleotide exchanger for ypt1 and ypt31/32. *Molecular biology of the cell* **11**, 4403–4411 (2000).

[40] Yu, S. & Liang, Y. A trapper keeper for trapp, its structures and functions. *Cellular and Molecular Life Sciences* **69**, 3933–3944 (2012).

[41] Ethell, I. M., Hagihara, K., Miura, Y., Irie, F. & Yamaguchi, Y. Synbindin, a novel syndecan-2–binding protein in neuronal dendritic spines. *J Cell Biol* **151**, 53–68 (2000).

[42] Simons, C. *et al.* A de novo mutation in the β-tubulin gene tubb4a results in the leukoencephalopathy hypomyelination with atrophy of the basal ganglia and cerebellum. *The American Journal of Human Genetics* **92**, 767–773 (2013).

[43] Cox, J. J. *et al.* An scn9a channelopathy causes congenital inability to experience pain. *Nature* **444**, 894–898 (2006).

[44] Fertleman, C. R. *et al.* Scn9a mutations in paroxysmal extreme pain disorder: allelic variants underlie distinct channel defects and phenotypes. *Neuron* **52**, 767–774 (2006).

[45] Yang, Y. *et al.* Mutations in scn9a, encoding a sodium channel alpha subunit, in patients with primary erythermalgia. *Journal of medical genetics* **41**, 171–174 (2004).

[46] Mukherjee, K. *et al.* Piccolo and bassoon maintain synaptic vesicle clustering without directly participating in vesicle exocytosis. *Proceedings of the National Academy of Sciences* **107**, 6504–6509 (2010).

[47] Hek, K. *et al.* The pclo gene and depressive disorders: replication in a population-based study. *Human Molecular Genetics* **19**, 731–734 (2009).

[48] Choi, K. H. *et al.* Gene expression and genetic variation data implicate pclo in bipolar disorder. *Biological psychiatry* **69**, 353–359 (2011).

[49] Shield, A. J., Murray, T. P. & Board, P. G. Functional characterisation of ganglioside-induced differentiation-associated protein 1 as a glutathione transferase. *Biochemical and biophysical research communications* **347**, 859–866 (2006).

[50] VanGuilder, H. D., Vrana, K. E. & Freeman, W. M. Twenty-five years of quantitative pcr for gene expression analysis. *Biotechniques* **44**, 619 (2008).

[51] Bustin, S., Benes, V., Nolan, T. & Pfaffl, M. Quantitative real-time rt-pcr–a perspective. *Journal of molecular endocrinology* **34**, 597–601 (2005).

[52] Bustin, S. A. Absolute quantification of mrna using real-time reverse transcription polymerase chain reaction assays. *Journal of molecular endocrinology* **25**, 169–193 (2000).

[53] Wong, M. L. & Medrano, J. F. Real-time pcr for mrna quantitation. *Biotechniques* **39**, 75–88 (2005).

[54] Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative pcr and the 2- $\delta\delta$ct method. *methods* **25**, 402–408 (2001).

[55] Vandesompele, J. *et al.* Accurate normalization of real-time quantitative rt-pcr data by geometric averaging of multiple internal control genes. *Genome biology* **3**, research0034–1 (2002).

[56] Eissa, N. *et al.* Stability of reference genes for messenger rna quantification by real-time pcr in mouse dextran sodium sulfate experimental colitis. *PloS one* **11**, e0156289 (2016).

[57] Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. & Vandesompele, J. qbase relative quantification framework and software for management and automated analysis of real-time quantitative pcr data. *Genome biology* **8**, R19 (2007).

[58] Boch, J. *et al.* Breaking the code of dna binding specificity of tal-type iii effectors. *Science* **326**, 1509–1512 (2009).

[59] Kim, Y.-G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to fok i cleavage domain. *Proceedings of the National Academy of Sciences* **93**, 1156–1160 (1996).

[60] Jinek, M. *et al.* A programmable dual-rna–guided dna endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).

[61] Makarova, K. S. *et al.* Evolution and classification of the crispr–cas systems. *Nature Reviews Microbiology* **9**, 467–477 (2011).

[62] Sorek, R., Kunin, V. & Hugenholtz, P. Crisprâa widespread system that provides acquired resistance against phages in bacteria and archaea. *Nature Reviews Microbiology* **6**, 181–186 (2008).

[63] Nishimasu, H. *et al.* Crystal structure of cas9 in complex with guide rna and target dna. *Cell* **156**, 935–949 (2014).

[64] Graham, D. B. & Root, D. E. Resources for the design of crispr gene editing experiments. *Genome biology* **16**, 260 (2015).

[65] Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A. 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**, 5429–5433 (1987).

[66] Jansen, R., van Embden, J. D., Gaastra, W. & Schouls, L. M. Identification of a novel family of sequence repeats among prokaryotes. *Omics: a journal of integrative biology* **6**, 23–33 (2002).

[67] Mojica, F. J., Díez-Villaseñor, C., Soria, E. & Juez, G. Biological significance of a family of regularly spaced repeats in the genomes of archaea, bacteria and mitochondria. *Molecular microbiology* **36**, 244–246 (2000).

[68] Pourcel, C., Salvignol, G. & Vergnaud, G. Crispr elements in yersinia pestis acquire new repeats by preferential uptake of bacteriophage dna, and provide additional tools for evolutionary studies. *Microbiology* **151**, 653–663 (2005).

[69] Jiang, F., Zhou, K., Ma, L., Gressel, S. & Doudna, J. A. A cas9–guide rna complex preorganized for target dna recognition. *Science* **348**, 1477–1481 (2015).

[70] Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. Dna interrogation by the crispr rna-guided endonuclease cas9. *Nature* **507**, 62–67 (2014).

[71] Kanaar, R., Hoeijmakers, J. H. & van Gent, D. C. Molecular mechanisms of dna double-strand break repair. *Trends in cell biology* **8**, 483–489 (1998).

[72] Takata, M. *et al.* Homologous recombination and non-homologous end-joining pathways of dna double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *The EMBO journal* **17**, 5497–5508 (1998).

[73] Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with crispr-cas9. *Science* **346**, 1258096 (2014).

[74] Irion, U., Krauss, J. & Nüsslein-Volhard, C. Precise and efficient genome editing in zebrafish using the crispr/cas9 system. *Development* **141**, 4827–4830 (2014).

[75] Lieber, M. R. The mechanism of double-strand dna break repair by the nonhomologous dna end-joining pathway. *Annual review of biochemistry* **79**, 181–211 (2010).

[76] Yamtich, J. & Sweasy, J. B. Dna polymerase family x: function, structure, and cellular roles. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **1804**, 1136–1150 (2010).

[77] Gu, J. *et al.* Xrcc4: Dna ligase iv can ligate incompatible dna ends and can ligate across gaps. *The EMBO journal* **26**, 1010–1023 (2007).

[78] McElhinny, S. A. N. *et al.* A gradient of template dependence defines distinct biological roles for family x polymerases in nonhomologous end joining. *Molecular cell* **19**, 357–366 (2005).

[79] Bebenek, K., Garcia-Diaz, M., Blanco, L. & Kunkel, T. A. The frameshift infidelity of human dna polymerase λ implications for function. *Journal of Biological Chemistry* **278**, 34685–34690 (2003).

[80] Domínguez, O. *et al.* Dna polymerase mu (pol μ), homologous to tdt, could act as a dna mutator in eukaryotic cells. *The EMBO journal* **19**, 1731–1742 (2000).

[81] Tippin, B., Kobayashi, S., Bertram, J. G. & Goodman, M. F. To slip or skip, visualizing frameshift mutation dynamics for error-prone dna polymerases. *Journal of Biological Chemistry* **279**, 45360–45368 (2004).

[82] Ma, Y., Pannicke, U., Schwarz, K. & Lieber, M. R. Hairpin opening and overhang processing by an artemis/dna-dependent protein kinase complex in nonhomologous end joining and v (d) j recombination. *Cell* **108**, 781–794 (2002).

[83] Gauss, G. H. & Lieber, M. R. Mechanistic constraints on diversity in human v (d) j recombination. *Molecular and cellular biology* **16**, 258–269 (1996).

[84] Sonoda, E., Hochegger, H., Saberi, A., Taniguchi, Y. & Takeda, S. Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. *DNA repair* **5**, 1021–1029 (2006).

[85] Zapotoczny, G. & Sekelsky, J. Human cell assays for synthesis-dependent strand annealing and crossing over during double-strand break repair. *G3: Genes, Genomes, Genetics* **7**, 1191–1199 (2017).

[86] Krejci, L., Altmannova, V., Spirek, M. & Zhao, X. Homologous recombination and its regulation. *Nucleic acids research* **40**, 5795–5818 (2012).

[87] Liang, P. *et al.* Crispr/cas9-mediated gene editing in human tripronuclear zygotes. *Protein & cell* **6**, 363–372 (2015).

[88] Straimer, J. *et al.* Site-specific genome editing in plasmodium falciparum using engineered zinc-finger nucleases. *Nature methods* **9**, 993–998 (2012).

[89] Cong, L. *et al.* Multiplex genome engineering using crispr/cas systems. *Science* **339**, 819–823 (2013).

[90] Kim, T. K. & Eberwine, J. H. Mammalian cell transfection: the present and the future. *Analytical and bioanalytical chemistry* **397**, 3173–3178 (2010).

[91] Dalby, B. *et al.* Advanced transfection with lipofectamine 2000 reagent: primary neurons, sirna, and high-throughput applications. *Methods* **33**, 95–103 (2004).

[92] Chu, Y., Masoud, M. & Gebeyehu, G. Transfection reagents (2016). US Patent 9,358,300.

[93] Liang, X. *et al.* Rapid and highly efficient mammalian cell engineering via cas9 protein transfection. *Journal of biotechnology* **208**, 44–53 (2015).

[94] Paganin-Gioanni, A. *et al.* Direct visualization at the single-cell level of sirna electrotransfer into cancer cells. *Proceedings of the National Academy of Sciences* **108**, 10443–10447 (2011).

[95] Blessing, D. & Déglon, N. Adeno-associated virus and lentivirus vectors: a refined toolkit for the central nervous system. *Current opinion in virology* **21**, 61–66 (2016).

[96] Senís, E. *et al.* Crispr/cas9-mediated genome engineering: an adeno-associated viral (aav) vector toolbox. *Biotechnology journal* **9**, 1402–1412 (2014).

[97] Cebrian-Serrano, A. & Davies, B. Crispr-cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. *Mammalian Genome* **28**, 247–261 (2017).

[98] Schambach, A., Zychlinski, D., Ehrnstroem, B. & Baum, C. Biosafety features of lentiviral vectors. *Human gene therapy* **24**, 132–142 (2013).

[99] Irv Edelman. Prime Algorithm (2002). URL `http://www.biology.wustl.edu/gcg/prime.html{\#}algorithm`.

[100] Splawski, I. *et al.* Severe arrhythmia disorder caused by cardiac l-type calcium channel mutations. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 8089–8096 (2005).

[101] Pfaffl, M. W. A new mathematical model for relative quantification in real-time rt–pcr. *Nucleic acids research* **29**, e45–e45 (2001).

[102] Yazawa, M. *et al.* Using induced pluripotent stem cells to investigate cardiac phenotypes in timothy syndrome. *Nature* **471**, 230–234 (2011).

[103] IDT. user guide Cationic lipid delivery of CRISPR ribonucleoprotein complex into mammalian cells genome editing (2016). URL `www.idtdna.com`.

[104] Dick, I. E., Joshi-Mukherjee, R., Yang, W. & Yue, D. T. Arrhythmogenesis in timothy syndrome is associated with defects in ca2+-dependent inactivation. *Nature communications* **7** (2016).

[105] Ellefsen, S., Stensløkken, K.-O., Sandvik, G. K., Kristensen, T. A. & Nilsson, G. E. Improved normalization of real-time reverse transcriptase polymerase chain reaction data using an external rna control. *Analytical biochemistry* **376**, 83–93 (2008).

[106] Tanaka, S., Furukawa, T. & Plotkin, S. A. Human cytomegalovirus stimulates host cell rna synthesis. *Journal of virology* **15**, 297–304 (1975).

[107] Thermofisher. mRNA Extraction. URL `https://www.thermofisher.com/no/en/home/life-science/dna-rna-purification-analysis/rna-extraction/rna-types/mrna-extraction.html`.

[108] Ståhlberg, A., Zoric, N., Åman, P. & Kubista, M. Quantitative real-time pcr for cancer detection: the lymphoma case. *Expert review of molecular diagnostics* **5**, 221–230 (2005).

[109] Hruz, T. *et al.* Refgenes: identification of reliable and condition specific reference genes for rt-qpcr data normalization. *BMC genomics* **12**, 156 (2011).

[110] Jensen, J. & Ørntoft, T. Normalization of real-time quantitative rt-pcr data: a model based variance estimation approach to identify genes suited for normalization-applied to bladder-and colon-cancer data-sets. *Cancer Research* **64**, 5245–5250 (2004).

[111] Fleige, S. & Pfaffl, M. W. Rna integrity and the effect on the real-time qrt-pcr performance. *Molecular aspects of medicine* **27**, 126–139 (2006).

[112] Rossen, L., Nørskov, P., Holmstrøm, K. & Rasmussen, O. F. Inhibition of pcr by components of food samples, microbial diagnostic assays and dna-extraction solutions. *International journal of food microbiology* **17**, 37–45 (1992).

[113] Wilson, I. G. Inhibition and facilitation of nucleic acid amplification. *Applied and environmental microbiology* **63**, 3741 (1997).

[114] Schoor, O. *et al.* Moderate degradation does not preclude microarray analysis of small amounts of rna. *Biotechniques* **35**, 1192–6 (2003).

[115] Huggett, J. F. *et al.* The digital miqe guidelines: minimum information for publication of quantitative digital pcr experiments. *Clinical chemistry* **59**, 892–902 (2013).

[116] Simbolo, M. *et al.* Dna qualification workflow for next generation sequencing of histopathological samples. *PLoS One* **8**, e62692 (2013).

[117] Van de Goor, T. The principle and promise of labchip technology. *Pharma Genomics* **3**, 16–18 (2003).

[118] Auer, H. *et al.* Chipping away at the chip bias: Rna degradation in microarray analysis. *Nature genetics* **35**, 292–293 (2003).

[119] Nolan, T., Hands, R. E. & Bustin, S. A. Quantification of mrna using real-time rt-pcr. *Nature protocols* **1**, 1559–1582 (2006).

[120] Huang, G. *et al.* cir-itch plays an inhibitory role in colorectal cancer by regulating the wnt/β-catenin pathway. *PloS one* **10**, e0131225 (2015).

[121] Park, N. J. *et al.* Characterization of salivary rna by cdna library analysis. *Archives of oral biology* **52**, 30–35 (2007).

[122] Curry, J., McHale, C. & Smith, M. T. Low efficiency of the moloney murine leukemia virus reverse transcriptase during reverse transcription of rare t (8; 21) fusion gene transcripts. *Biotechniques* **32**, 768–770 (2002).

[123] Ståhlberg, A., Kubista, M. & Pfaffl, M. Comparison of reverse transcriptases in gene expression analysis. *Clinical chemistry* **50**, 1678–1680 (2004).

[124] ZHANG, J. & BYRNE, C. D. Differential priming of rna templates during cdna synthesis markedly affects both accuracy and reproducibility of quantitative competitive reverse-transcriptase pcr. *Biochemical Journal* **337**, 231–241 (1999).

[125] Wu, Z., Yang, H. & Colosi, P. Effect of genome size on aav vector packaging. *Molecular Therapy* **18**, 80–86 (2010).

[126] Breslauer, K. J., Frank, R., Blöcker, H. & Marky, L. A. Predicting dna duplex stability from the base sequence. *Proceedings of the National Academy of Sciences* **83**, 3746–3750 (1986).

[127] NCBI, R. C. Database resources of the national center for biotechnology information. *Nucleic acids research* **45**, D12 (2017).

[128] Kolb, K., Jaenicke, G., Kramer, M. & Schulze, P. Absorption, distribution and elimination of labeled dimethyl sulfoxide in man and animals. *Annals of the New York Academy of Sciences* **141**, 85–95 (1967).

[129] Narula, P. The comparative penetrant-carrier action of dimethyl sulfoxide and ethyl alcohol in vivo. *Annals of the New York Academy of Sciences* **141**, 277–278 (1967).

[130] Imaizumi, K. *et al.* A simple and highly effective method for slow-freezing human pluripotent stem cells using dimethyl sulfoxide, hydroxyethyl starch and ethylene glycol. *PLoS One* **9**, e88696 (2014).

[131] Liu, D. & Pan, F. Advances in cryopreservation of organs. *Journal of Huazhong University of Science and Technology [Medical Sciences]* **36**, 153–161 (2016).

[132] Ramakrishna, S. *et al.* Gene disruption by cell-penetrating peptide-mediated delivery of cas9 protein and guide rna. *Genome research* **24**, 1020–1027 (2014).

[133] Fu, Y. *et al.* High-frequency off-target mutagenesis induced by crispr-cas nucleases in human cells. *Nature biotechnology* **31**, 822–826 (2013).

[134] Pattanayak, V. *et al.* High-throughput profiling of off-target dna cleavage reveals rna-programmed cas9 nuclease specificity. *Nature biotechnology* **31**, 839–843 (2013).

[135] Weninger, A., Glieder, A. & Vogl, T. A toolbox of endogenous and heterologous nuclear localization sequences for the methylotrophic yeast pichia pastoris. *FEMS yeast research* **15** (2015).

[136] Lakshmipathy, U. *et al.* Efficient transfection of embryonic and adult stem cells. *Stem cells* **22**, 531–543 (2004).

[137] Chu, V. T. *et al.* Increasing the efficiency of homology-directed repair for crispr-cas9-induced precise gene editing in mammalian cells. *Nature biotechnology* **33**, 543–548 (2015).

[138] Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L. & Corn, J. E. Enhancing homology-directed genome editing by catalytically active and inactive crispr-cas9 using asymmetric donor dna. *Nature biotechnology* **34**, 339–344 (2016).

[139] Moore, J. C. *et al.* Efficient, high-throughput transfection of human embryonic stem cells. *Stem cell research & therapy* **1**, 23 (2010).

[140] Nayak, S. & Herzog, R. W. Progress and prospects: immune responses to viral vectors. *Gene therapy* **17**, 295–304 (2010).

[141] Howard, D. B., Powers, K., Wang, Y. & Harvey, B. K. Tropism and toxicity of adeno-associated viral vector serotypes 1, 2, 5, 6, 7, 8, and 9 in rat neurons and glia in vitro. *Virology* **372**, 24–34 (2008).

[142] Jinek, M. *et al.* Rna-programmed genome editing in human cells. *elife* **2**, e00471 (2013).

[143] Kumar, M., Keller, B., Makalou, N. & Sutton, R. E. Systematic determination of the packaging limit of lentiviral vectors. *Human gene therapy* **12**, 1893–1905 (2001).

[144] Cho, S. W. *et al.* Analysis of off-target effects of crispr/cas-derived rna-guided endonucleases and nickases. *Genome research* **24**, 132–141 (2014).

[145] Vouillot, L., Thélie, A. & Pollet, N. Comparison of t7e1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. *G3: Genes, Genomes, Genetics* **5**, 407–415 (2015).

[146] Guschin, D. Y. *et al.* A rapid and general assay for monitoring endogenous gene modification. *Engineered Zinc Finger Proteins: Methods and Protocols* 247–256 (2010).

[147] Miller, J. C. *et al.* An improved zinc-finger nuclease architecture for highly specific genome editing. *Nature biotechnology* **25**, 778–785 (2007).

[148] Kleinstiver, B. P. *et al.* High-fidelity crispr–cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).

[149] Auer, T. O., Duroure, K., De Cian, A., Concordet, J.-P. & Del Bene, F. Highly efficient crispr/cas9-mediated knock-in in zebrafish by homology-independent dna repair. *Genome research* **24**, 142–153 (2014).

[150] Lahr, D. & Katz, L. A. Reducing the impact of pcr-mediated recombination in molecular evolution and environmental studies using a new-generation high-fidelity dna polymerase. *Biotechniques* **47**, 857–866 (2009).

[151] Birch, D. E. Simplified hot start pcr. *Nature* **381**, 445 (1996).

[152] Bonefeld, B. E., Elfving, B. & Wegener, G. Reference genes for normalization: a study of rat brain tissue. *Synapse* **62**, 302–309 (2008).

[153] Robinson, T., Sutherland, I. & Sutherland, J. Validation of candidate bovine reference genes for use with real-time pcr. *Veterinary immunology and immunopathology* **115**, 160–165 (2007).

[154] Remschmidt, H. Early-onset schizophrenia as a progressive-deteriorating developmental disorder: evidence from child psychiatry. *Journal of neural transmission* **109**, 101–117 (2002).

[155] Tebbenkamp, A. T., Willsey, A. J., State, M. W. & Šestan, N. The developmental transcriptome of the human brain: implications for neurodevelopmental disorders. *Current opinion in neurology* **27**, 149 (2014).

[156] Wu, J. Q. *et al.* Transcriptome sequencing revealed significant alteration of cortical promoter usage and splicing in schizophrenia. *PloS one* **7**, e36351 (2012).

[157] van den Hurk, M. *et al.* Transcriptional and epigenetic mechanisms of cellular reprogramming to induced pluripotency. *Epigenomics* **8**, 1131–1149 (2016).

[158] Vierbuchen, T. *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035–1041 (2010).

7 Appendix

7.1 DNeasy Miniprep protocol (Qiagen)

Protocol: Purification of Total DNA from Animal Blood or Cells (Spin-Column Protocol)

This protocol is designed for purification of total DNA from animal blood (with nucleated or nonnucleated erythrocytes) or from cultured animal or human cells.

Important points before starting

- If using the DNeasy Blood & Tissue Kit for the first time, read "Important Notes" (page 15).
- All centrifugation steps are carried out at room temperature (15–25°C) in a microcentrifuge.
- Vortexing should be performed by pulse-vortexing for 5–10 s.
- PBS is required for use in step 1 (see page 14 for composition). Buffer ATL is not required in this protocol.
- Optional: RNase A may be used to digest RNA during the procedure. RNase A is not provided in the DNeasy Blood & Tissue Kit (see "Copurification of RNA", page 19).

Things to do before starting

- Buffer AL may form a precipitate upon storage. If necessary, warm to 56°C until the precipitate has fully dissolved.
- Buffer AW1 and Buffer AW2 are supplied as concentrates. Before using for the first time, add the appropriate amount of ethanol (96–100%) as indicated on the bottle to obtain a working solution.
- Preheat a thermomixer, shaking water bath, or rocking platform to 56°C for use in step 2.

Procedure

1. **For blood with nonnucleated erythrocytes, follow step 1a; for blood with nucleated erythrocytes, follow step 1b; for cultured cells, follow step 1c.**

   Blood from mammals contains nonnucleated erythrocytes. Blood from animals such as birds, fish, or frogs contains nucleated erythrocytes.

1a. **Nonnucleated: Pipet 20 µl proteinase K into a 1.5 ml or 2 ml microcentrifuge tube (not provided). Add 50–100 µl anticoagulated blood. Adjust the volume to 220 µl with PBS. Continue with step 2.**

   Optional: If RNA-free genomic DNA is required, add 4 µl RNase A (100 mg/ml) and incubate for 2 min at room temperature before continuing with step 2.

1b. Nucleated: Pipet 20 µl proteinase K into a 1.5 ml or 2 ml microcentrifuge tube (not provided). Add 5–10 µl anticoagulated blood. Adjust the volume to 220 µl with PBS. Continue with step 2.

Optional: If RNA-free genomic DNA is required, add 4 µl RNase A (100 mg/ml) and incubate for 2 min at room temperature before continuing with step 2.

1c. Cultured cells: Centrifuge the appropriate number of cells (maximum 5 x 10^6) for 5 min at 300 x g. Resuspend the pellet in 200 µl PBS. Add 20 µl proteinase K. Continue with step 2.

When using a frozen cell pellet, allow cells to thaw before adding PBS until the pellet can be dislodged by gently flicking the tube.

Ensure that an appropriate number of cells is used in the procedure. For cell lines with a high degree of ploidy (e.g., HeLa cells), it is recommended to use less than the maximum number of cells listed in Table 1, page 16.

Optional: If RNA-free genomic DNA is required, add 4 µl RNase A (100 mg/ml), mix by vortexing, and incubate for 2 min at room temperature before continuing with step 2.

2. Add 200 µl Buffer AL (without added ethanol). Mix thoroughly by vortexing, and incubate at 56°C for 10 min.

Ensure that ethanol has not been added to Buffer AL (see "Buffer AL", page 18). Buffer AL can be purchased separately (see page 56 for ordering information).

It is essential that the sample and Buffer AL are mixed immediately and thoroughly by vortexing or pipetting to yield a homogeneous solution.

3. Add 200 µl ethanol (96–100%) to the sample, and mix thoroughly by vortexing.

It is important that the sample and the ethanol are mixed thoroughly to yield a homogeneous solution.

4. Pipet the mixture from step 3 into the DNeasy Mini spin column placed in a 2 ml collection tube (provided). Centrifuge at ≥6000 x *g* (8000 rpm) for 1 min. Discard flow-through and collection tube.*

5. Place the DNeasy Mini spin column in a new 2 ml collection tube (provided), add 500 µl Buffer AW1, and centrifuge for 1 min at ≥6000 x *g* (8000 rpm). Discard flow-through and collection tube.*

6. Place the DNeasy Mini spin column in a new 2 ml collection tube (provided), add 500 µl Buffer AW2, and centrifuge for 3 min at 20,000 x *g* (14,000 rpm) to dry the DNeasy membrane. Discard flow-through and collection tube.

It is important to dry the membrane of the DNeasy Mini spin column, since residual ethanol may interfere with subsequent reactions. This centrifugation step ensures that no residual ethanol will be carried over during the following elution.

* Flow-through contains Buffer AL or Buffer AW1 and is therefore not compatible with bleach. See page 8 for safety information.

Following the centrifugation step, remove the DNeasy Mini spin column carefully so that the column does not come into contact with the flow-through, since this will result in carryover of ethanol. If carryover of ethanol occurs, empty the collection tube, then reuse it in another centrifugation for 1 min at 20,000 x *g* (14,000 rpm).

7. **Place the DNeasy Mini spin column in a clean 1.5 ml or 2 ml microcentrifuge tube (not provided), and pipet 200 µl Buffer AE directly onto the DNeasy membrane. Incubate at room temperature for 1 min, and then centrifuge for 1 min at ≥6000 x *g* (8000 rpm) to elute.**

   Elution with 100 µl (instead of 200 µl) increases the final DNA concentration in the eluate, but also decreases the overall DNA yield (see Figure 2, page 21).

8. **Recommended: For maximum DNA yield, repeat elution once as described in step 7.**

   This step leads to increased overall DNA yield.

   A new microcentrifuge tube can be used for the second elution step to prevent dilution of the first eluate. Alternatively, to combine the eluates, the microcentrifuge tube from step 7 can be reused for the second elution step.

   Note: Do not elute more than 200 µl into a 1.5 ml microcentrifuge tube because the DNeasy Mini spin column will come into contact with the eluate.

7.2 Q5 Hi-Fi polymerase protocol (New England Biolabs)

PCR Using Q5® High-Fidelity DNA Polymerase (M0491)

Protocols.io also provides an interactive version of this protocol where you can discover and share optimizations with the research community.

1. Please note that protocols with Q5 High-Fidelity DNA Polymerase may differ from protocols with other polymerases. Conditions recommended below should be used for optimal performance.

Reaction Setup:
We recommend assembling all reaction components on ice and quickly transferring the reactions to a thermocycler preheated to the denaturation temperature (98°C). All components should be mixed prior to use. Q5 High-Fidelity DNA Polymerase may be diluted in 1X Q5 Reaction Buffer just prior to use in order to reduce pipetting errors.

COMPONENT	25 µl REACTION	50 µl REACTION	FINAL CONCENTRATION
5X Q5 Reaction Buffer	5 µl	10 µl	1X
10 mM dNTPs	0.5 µl	1 µl	200 µM
10 µM Forward Primer	1.25 µl	2.5 µl	0.5 µM
10 µM Reverse Primer	1.25 µl	2.5 µl	0.5 µM
Template DNA	variable	variable	< 1,000 ng
Q5 High-Fidelity DNA Polymerase	0.25 µl	0.5 µl	0.02 U/µl
5X Q5 High GC Enhancer (optional)	(5 µl)	(10 µl)	(1X)
Nuclease-Free Water	to 25 µl	to 50 µl	

Notes: Gently mix the reaction. Collect all liquid to the bottom of the tube by a quick spin if necessary. Overlay the sample with mineral oil if using a PCR machine without a heated lid.

Transfer PCR tubes to a PCR machine and begin thermocycling.

Thermocycling Conditions for a Routine PCR:

STEP	TEMP	TIME
Initial Denaturation	98°C	30 seconds
25–35 Cycles	98°C *50–72°C 72°C	5–10 seconds 10–30 seconds 20–30 seconds/kb
Final Extension	72°C	2 minutes
Hold	4–10°C	

*Use of the NEB T_m Calculator is highly recommended.

2. General Guidelines:

Template:
Use of high quality, purified DNA templates greatly enhances the success of PCR. Recommended amounts of DNA template for a 50 µl reaction are as follows:

DNA	AMOUNT
DNA Genomic	1 ng–1 µg
Plasmid or Viral	1 pg–1 ng

3. Primers:
Oligonucleotide primers are generally 20–40 nucleotides in length and ideally have a GC content of 40–60%. Computer programs such as Primer3 can be used to design or analyze primers. The best results are typically seen when using each primer at a final concentration of 0.5 µM in the reaction.

4. Mg^{++} and additives:
Mg^{++} concentration of 2.0 mM is optimal for most PCR products generated with Q5 High-Fidelity DNA Polymerase. When used at a final concentration of 1X, the Q5 Reaction Buffer provides the optimal Mg^{++}concentration.

Amplification of some difficult targets, like GC-rich sequences, may be improved by the addition of 1X Q5 High GC Enhancer. The Q5 High GC Enhancer is not a buffer and should not be used alone. It should be added only to reactions with the Q5 Reaction Buffer when other conditions have

failed.

5. Deoxynucleotides:
The final concentration of dNTPs is typically 200 µM of each deoxynucleotide. Q5 High-Fidelity DNA Polymerase cannot incorporate dUTP and is not recommended for use with uracil-containing primers or templates.

6. Q5 High-Fidelity DNA Polymerase concentration:
We generally recommend using Q5 High-Fidelity DNA Polymerase at a final concentration of 20 units/ml (1.0 unit/50 µl reaction). However, the optimal concentration of Q5 High-Fidelity DNA Polymerase may vary from 10–40 units/ml (0.5–2 units/50 µl reaction) depending on amplicon length and difficulty. Do not exceed 2 units/50 µl reaction, especially for amplicons longer than 5 kb.

7. Buffers:
The 5X Q5 Reaction Buffer provided with the enzyme is recommended as the first-choice buffer for robust, high-fidelity amplification. For difficult amplicons, such as GC-rich templates or those with secondary structure, the addition of the Q5 High GC Enhancer can improve reaction performance. The 5X Q5 Reaction Buffer is detergent-free and contains 2.0 mM Mg^{++} at the final (1X) concentration.

8. Denaturation:
An initial denaturation of 30 seconds at 98°C is sufficient for most amplicons from pure DNA templates. Longer denaturation times can be used (up to 3 minutes) for templates that require it.

During thermocycling, the denaturation step should be kept to a minimum. Typically, a 5–10 second denaturation at 98°C is recommended for most templates.

9. Annealing:
Optimal annealing temperatures for Q5 High-Fidelity DNA Polymerase tend to be higher than for other PCR polymerases. The **NEB T_m Calculator** should be used to determine the annealing temperature when using this enzyme. Typically, use a 10–30 second annealing step at 3°C above the T_m of the lower T_m primer. A temperature gradient can also be used to optimize the annealing temperature for each primer pair.

For high T_m primer pairs, two-step cycling without a separate annealing step can be used (see note 12).

10. Extension:
The recommended extension temperature is 72°C. Extension times are generally 20–30 seconds per kb for complex, genomic samples, but can be

reduced to 10 seconds per kb for simple templates (plasmid, *E. coli*, etc.) or complex templates < 1 kb. Extension time can be increased to 40 seconds per kb for cDNA or long, complex templates, if necessary.

A final extension of 2 minutes at 72°C is recommended.

11. Cycle number:
Generally, 25–35 cycles yield sufficient product. For genomic amplicons, 30-35 cycles are recommended.

12. 2-step PCR:
When primers with annealing temperatures ≥ 72°C are used, a 2-step thermocycling protocol (combining annealing and extension into one step) is possible.

13. Amplification of long products:
When amplifying products > 6 kb, it is often helpful to increase the extension time to 40–50 seconds/kb.

14. PCR product:
The PCR products generated using Q5 High-Fidelity DNA Polymerase have blunt ends. If cloning is the next step, then blunt-end cloning is recommended. If T/A-cloning is preferred, the DNA should be purified prior to A-addition, as Q5 High-Fidelity DNA Polymerase will degrade any overhangs generated.

Addition of an untemplated -dA can be done with *Taq* DNA Polymerase (NEB #M0267) or Klenow exo$^-$(NEB #M0212).

7.3 T7 Endonuclease 1 Digestion protocol (New England Biolabs)

T7 Endonuclease I digestion:

1. Assemble reactions as follows

Component	19 µl annealing reaction
DNA	200 ng
10X NEBuffer 2	2 µl
Nuclease-free Water	To 19 µl

2. Anneal the PCR products in a thermocycler using the following conditions:

Hybridization Conditions

Step	Temperature	Ramp Rate	Time
Initial Denaturation	95°C		5 minutes
Annealing	95-85°C	-2°C/second	
	85-25°C	-0.1°C/second	
Hold	4°C		Hold

3. Add T7 Endonuclease I to the annealed PCR products

Component	20 µl reaction
Annealed PCR product	19 µl
T7 Endonuclease I (M0302)	1 µl
Incubation Time	15 minutes
Incubation Temperature	37°C

4. Stop the reaction by adding 1.5 µl of 0.25 M EDTA.
5. Purify the reaction using 36 µl of Ampure XP beads according to the manufacturer's suggestion. This step is optional since 1 µl of the reaction will not interfere with analysis on an Agilent Bioanalyzer using DNA1000 reagents.
6. Elute the DNA fragments in 20 µl of water, recovering 15 µl.

7.4 EnGen kit protocol (New England Biolabs)

Protocols:

PCR – New users are encouraged to perform PCR and T7 Endonuclease I digestion using the included control template and primer mix. For each amplicon we recommend setting up three PCR reactions using the following templates:

a. *gDNA from targeted cells (e.g., Cas9, TALEN or ZFN transfected cells)*
b. *gDNA from negative control cells (e.g., non-specific DNA transfected cells)*
c. *Water (i.e., no template control)*

Amplification reactions for experimental samples

1. Thaw the kit components, mix and pulse-spin in microfuge each component prior to use.

2 Set up a 25 µl PCR reaction using up to 500 ng of genomic DNA as a template.

Assemble the following reaction at room temperature:

REAGENT	25 µl RXN	FINAL CONCENTRATION
Q5 Hot Start High-Fidelity 2X Master Mix	12.5 µl	1X
10 µM Forward Primer	1.25 µl	0.5 µM
10 µM Reverse Primer	1.25 µl	0.5 µM
Template DNA	variable	0.5–500 ng genomic DNA*
Nuclease-free water	to 25 µl	

*** To use cell lysate directly in PCR, lyse cells in QuickExtract or DNAzol Direct using 50 µl cells in each well of a 96-well plate (~40,000 cells) according to the manufacturers' recommendation. Dilute the lysate 1:5 in TE and use 2.5 µl of the diluted lysate.**

3. Gently mix the reaction. Collect all the liquid to the bottom of the tube with a brief spin. Transfer the tubes to a PCR machine and begin thermocycling using the following conditions:

CYCLE STEP	TEMP	TIME	CYCLES
Initial Denaturation	98°C	30 seconds	1
Denaturation Annealing Extension	98°C 50–72°C* 72°C	5 seconds 10 seconds 20 seconds	35
Final Extension	72°C	2 minutes	1
Hold	4–10°C		

* Please visit **Tmcalculator.neb.com** to determine correct annealing temperature.

Control reaction using included Control Template and Primer Mix.

1. Set up a 25 µl PCR reaction as follows:

REAGENT	25 µl RXN	FINAL CONCENTRATION
Q5 Hot Start High-Fidelity 2X Master Mix	12.5 µl	1X
Control Template and Primer Mix	2.5 µl	0.5 ng plasmid and 0.5 µM each primer
Nuclease-free water	10 µl	

2. Gently mix the reaction. Collect all the liquid to the bottom of the tube with a brief spin. Transfer the tubes to a PCR machine and begin thermo-cycling using the following conditions:

CYCLE STEP	TEMP	TIME	CYCLES
Initial Denaturation	98°C	30 seconds	1
Denaturation Annealing Extension	98°C 65°C 72°C	5 seconds 10 seconds 20 seconds	35
Final Extension	72°C	2 minutes	1
Hold	4–10°C		

3. Analyze a small amount of the PCR product on an agarose gel to verify amplification of a single product of the correct size. A DNA marker should also be run to help estimate amplicon concentration. The product of the control PCR reaction is ~600 bp.

Heteroduplex formation:

The products of the PCR reaction must be denatured and annealed in order to allow formation of heteroduplex between PCR products with and without mutations. T7 Endonuclease I digestion has been optimized for use with 5 µl of the PCR reaction, containing up to 250 ng of amplified DNA.

The following protocol applies to both experimental and control reactions:

1. Assemble the reaction as follows:

REAGENT	19 µl ANNEALING REACTION
PCR Reaction	5 µl
10X NEBuffer 2	2 µl
Nuclease-free water	12 µl

2. Denature and then anneal the products in a thermocycler using the following program*:

CYCLE STEP	TEMP	RAMP RATE	TIME
Initial Denaturation	95°C		5 minutes
Annealing	95–85°C 85–25°C	-2°C/second -0.1°C/second	
Hold	4°C		

* Alternatively, if a thermocycler is not available with these ramp speeds, the sample can be heated to 95ºC for 10 minutes and then allowed to cool slowly to room temperature.

3. Proceed to heteroduplex digestion.

Heteroduplex digestion:

The digestion reaction conditions have been optimized for 5 µl of the unpurified PCR reaction containing up to 250 ng of amplified DNA. Increased amounts of PCR reaction and/or DNA may lead to inaccurate estimates of editing efficiencies.

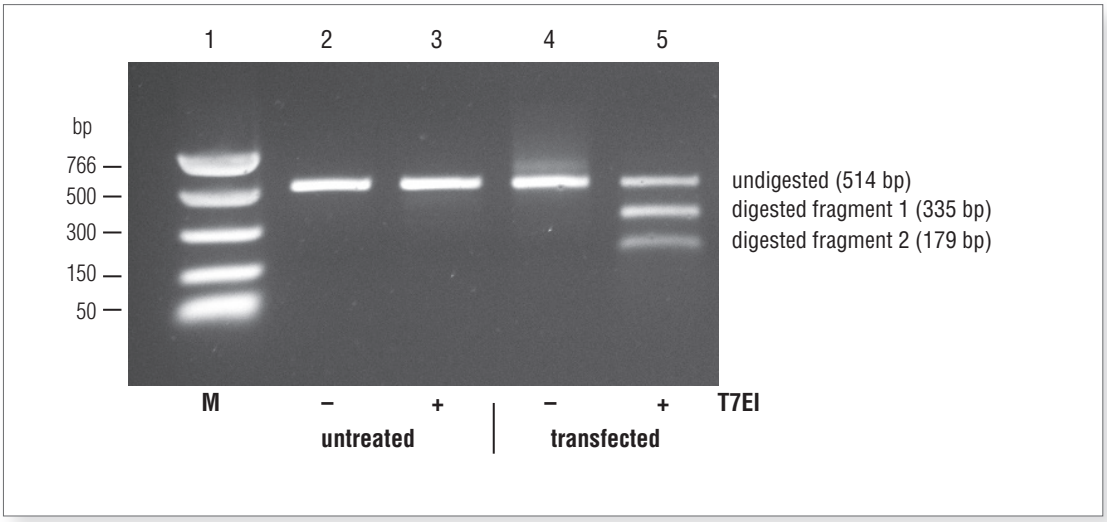
1. Set up each reaction as follows:

REAGENT	20 µl T7E1 REACTION
Annealed PCR Product	19 µl
EnGen T7 Endonuclease I	1 µl

2. Mix well and briefly spin. Incubate each reaction at 37°C for 15 minutes.
3. Following digestion, add 1 µl of Proteinase K and mix well.
4. Incubate for 5 minutes at 37°C to inactivate the T7 Endonuclease I.
5. Proceed with fragment analysis or store at –20°C until ready.

Optional: If needed, reactions can be purified prior to fragment analysis. For this we recommend the Monarch™ PCR & DNA Cleanup Kit (5 µg) (NEB #T1030).

Figure 2: Example of Mutation Detection on targeted 293 cells.



Genomic DNA was isolated from HEK 293 cells using Epicentre QuickExtract DNA extraction solution. Cells were either untreated (neg control) or transfected with Cas9 and guide RNA. Genomic DNA was amplified using Q5 Hot Start High-Fidelity Master Mix and denatured/annealed and digested with T7 Endonuclease I (T7E1) according to the recommended protocol. Lane 1: NEB PCR Marker (NEB #N3234), Lane 2: untreated genomic DNA, Lane 3: untreated genomic DNA digested with T7E1, Lane 4: DNA transfected with Cas9 and guide RNA, Lane 5: DNA transfected with Cas9 and guide RNA and digested with T7E1. Note that heteroduplexes can sometimes be seen running above the parental band, as seen in undigested test sample (lane 4).

Analysis of DNA Fragments:

1. Gel Analysis

   Add 4 µl of Gel Loading Dye, Purple (6X), no SDS (NEB #B7025) to the reaction and run on a 2% agarose gel stained with ethidium bromide. Run the included DNA ladder or an appropriate DNA size marker along side the sample for reference.

   Alternatively, samples can be analyzed using a fragment analyzer (e.g., Agilent Bioanalyzer or Advanced Analytical Technologies, Inc (AATI) Fragment Analyzer). For the Agilent Bioanalyzer, 1 µl of the reaction will not interfere when using the standard sensitivity Agilent DNA analysis kits. For the AATI Fragment Analyzer, 2 µl of the reaction can be used with the Standard Sensitivity NGS Fragment Analysis Kit (AATI Cat# DNF-473).

   Digestion of the control amplicon yields fragments of ~200 bp and ~400 bp in addition to the parental band.

2. Calculate the estimated % modification using the following formula:

   % Modification = 100 x [1-(1-fraction cleaved)$^{1/2}$]

   When calculating % modification for reactions with the control template where the starting material is known, the equation (100 x fraction cleaved) can be used.

   Where fraction cleaved = concentration of digested products/ (concentration of digested products + concentration of undigested band)

7.5 RNeasy micro protocol (Qiagen)

Things to do before starting

- If purifying RNA from cell lines rich in RNases, we recommend adding β-mercaptoethanol (β-ME) to Buffer RLT Plus before use. Add 10 µl β-ME per 1 ml Buffer RLT Plus. Dispense in a fume hood and wear appropriate protective clothing. Buffer RLT Plus containing β-ME can be stored at room temperature (15–25°C) for up to 1 month. Alternatively, add 20 µl of 2 M dithiothreitol (DTT) per 1 ml Buffer RLT Plus. The stock solution of 2 M DTT in water should be prepared fresh, or frozen in single-use aliquots. Buffer RLT Plus containing DTT can be stored at room temperature for up to 1 month.
- When processing <500 cells, carrier RNA may be added to the lysate before homogenization (see "Carrier RNA", page 15). Before using for the first time, dissolve the carrier RNA (310 µg) in 1 ml RNase-free water. Store this stock solution at –15 to –30°C, and use it to make fresh dilutions for each set of RNA preps. The concentration of this stock solution is 310 µg/ml (i.e., 310 ng/µl). To make a working solution (4 ng/µl) for 10 preps, add 5 µl stock solution to 34 µl Buffer RLT Plus and mix by pipetting. Add 6 µl of this diluted solution to 54 µl Buffer RLT Plus to give a working solution of 4 ng/µl. Add 5 µl of this solution to the lysate in step 3. **Do not add the carrier RNA to the lysate if purifying RNA for use in oligo-dT–based amplification.**
- Buffer RPE is supplied as a concentrate. Before using for the first time, add 4 volumes of ethanol (96–100%) as indicated on the bottle to obtain a working solution.
- Before using the kit for the first time, prepare 80% ethanol by mixing 24 ml ethanol (96–100%) and 6 ml RNase-free water (supplied). The procedure also requires 70% ethanol, which can be prepared by diluting ethanol (96–100%) with distilled water (not supplied).
- Buffer RLT Plus may form a precipitate during storage. If necessary, redissolve by warming, and then place at room temperature.

Procedure

1. Harvest cells according to step 1a or 1b.

1a. Cells grown in suspension (do not use more than 5 x 10^5 cells): Determine the number of cells. Pellet the appropriate number of cells by centrifuging for 5 min at 300 x *g* in a centrifuge tube (not supplied). Carefully remove all supernatant by aspiration, and proceed to step 2.

Note: Incomplete removal of cell-culture medium will inhibit lysis and dilute the lysate, affecting the conditions for DNA removal and RNA purification. Both effects may reduce RNA yield and purity.

1b. Cells grown in a monolayer (do not use more than 5 x 10^5 cells): Cells grown in a monolayer in cell-culture vessels can be either lysed directly in the vessel (up to 10 cm diameter) or trypsinized and collected as a cell pellet prior to lysis. Cells grown in a monolayer in cell-culture flasks should always be trypsinized.

To lyse cells directly:
Determine the number of cells. Completely aspirate the cell-culture medium, and proceed immediately to step 2.

Note: Incomplete removal of cell-culture medium will inhibit lysis and dilute the lysate, affecting the conditions for DNA removal and RNA purification. Both effects may reduce RNA yield and purity.

To trypsinize and collect cells:
Determine the number of cells. Aspirate the medium, and wash the cells with PBS. Aspirate the PBS, and add 0.10–0.25% trypsin in PBS. After the cells detach from the dish or flask, add medium (containing serum to inactivate the trypsin), transfer the cells to an RNase-free glass or polypropylene centrifuge tube (not supplied), and centrifuge at 300 x *g* for 5 min. Completely aspirate the supernatant, and proceed to step 2.

Note: Incomplete removal of cell-culture medium will inhibit lysis and dilute the lysate, affecting the conditions for DNA removal and RNA purification. Both effects may reduce RNA yield and purity.

2. Disrupt the cells by adding Buffer RLT Plus.

For pelleted cells, loosen the cell pellet thoroughly by flicking the tube. Add 350 µl Buffer RLT Plus. Vortex or pipet to mix, and proceed to step 3.

If processing ≤1 x 10^5 cells, 75 µl Buffer RLT Plus can be added instead. This allows cell pelleting in smaller tubes. Pipet up and down to lyse the cells.

Note: Incomplete loosening of the cell pellet may lead to inefficient lysis and reduced RNA yields.

For direct lysis of cells grown in a monolayer, add 350 µl Buffer RLT Plus to the cell-culture dish. Collect the lysate with a rubber policeman. Pipet the lysate into a microcentrifuge tube (not supplied). Vortex or pipet to mix, and ensure that no cell clumps are visible before proceeding to step 3.

If processing ≤1 x 10^5 cells, 75 µl Buffer RLT Plus can be added instead. This may be necessary for multiwell plates and cell-culture dishes. Pipet up and down to lyse the cells.

3. Homogenize the lysate according to step 3a, 3b, or 3c.

See "Disrupting and homogenizing starting material", page 13, for more details on homogenization. If processing ≤1 x 10^5 cells, they can be homogenized by vortexing for 1 min. After homogenization, proceed to step 4.

Note: If only 75 µl Buffer RLT Plus was used in step 2, transfer the lysate to a new 1.5 ml microcentrifuge tube, and adjust the volume to 350 µl with Buffer RLT Plus. Vortex for 1 min to homogenize and proceed to step 4.

Note: If processing <500 cells, 20 ng carrier RNA (5 µl of a 4 ng/µl solution) may be added to the lysate before homogenization. Prepare the carrier RNA as described in "Things to do before starting".

Note: Incomplete homogenization leads to significantly reduced RNA yields and can cause clogging of the gDNA Eliminator and RNeasy MinElute spin columns. Homogenization with the TissueRuptor or QIAshredder homogenizer generally results in higher RNA yields than with a syringe and needle.

3a. Pipet the lysate directly into a QIAshredder spin column (not supplied) placed in a 2 ml collection tube, and centrifuge for 2 min at full speed. Proceed to step 4.

3b. Place the tip of the TissueRuptor disposable probe into the lysate and operate the TissueRuptor at full speed until the lysate is homogenous (usually 30 s). Proceed to step 4.

Note: To avoid damage to the TissueRuptor and disposable probe during operation, make sure the tip of the probe remains submerged in the buffer.

3c. Pass the lysate at least 5 times through a blunt 20-gauge needle (0.9 mm diameter) fitted to an RNase-free syringe. Proceed to step 4.

4. Transfer the homogenized lysate to a gDNA Eliminator spin column placed in a 2 ml collection tube (supplied). Centrifuge for 30 s at ≥8000 x *g* (≥10,000 rpm). Discard the column, and save the flowthrough.

Note: Make sure that no liquid remains on the column membrane after centrifugation. If necessary, repeat the centrifugation until all liquid has passed through the membrane.

The remaining steps of this protocol allow the purification of RNA molecules longer than 200 nucleotides. If purification of total RNA containing small RNAs such as miRNA is desired, follow steps D1–D6 in Appendix D on page 45 instead of steps 5–11 in this protocol.

5. **Add 1 volume (usually 350 µl) of 70% ethanol to the flow-through from step 4, and mix well by pipetting. Do not centrifuge. Proceed immediately to step 6.**

   Note: The volume of 70% ethanol to add may be less than 350 µl if some lysate was lost during homogenization and DNA removal.

   Note: When purifying RNA from certain cell lines, precipitates may be visible after addition of ethanol. This does not affect the procedure.

6. **Transfer the sample, including any precipitate that may have formed, to an RNeasy MinElute spin column placed in a 2 ml collection tube (supplied). Close the lid gently, and centrifuge for 15 s at ≥8000 x *g* (≥10,000 rpm). Discard the flow-through.***

   Optional: If recovery of protein is desired, keep the flow-through on ice and follow steps E1–E5 in Appendix E on page 47.

   Reuse the collection tube in step 7.

7. **Add 700 µl Buffer RW1 to the RNeasy MinElute spin column. Close the lid gently, and centrifuge for 15 s at ≥8000 x *g* (≥10,000 rpm) to wash the spin column membrane. Discard the flow-through.***

   Reuse the collection tube in step 8.

   Note: After centrifugation, carefully remove the RNeasy MinElute spin column from the collection tube so that the column does not contact the flow-through. Be sure to empty the collection tube completely.

8. **Add 500 µl Buffer RPE to the RNeasy MinElute spin column. Close the lid gently, and centrifuge for 15 s at ≥8000 x *g* (≥10,000 rpm) to wash the spin column membrane. Discard the flow-through.**

   Reuse the collection tube in step 9.

   Note: Buffer RPE is supplied as a concentrate. Ensure that ethanol is added to Buffer RPE before use (see "Things to do before starting").

9. **Add 500 µl of 80% ethanol to the RNeasy MinElute spin column. Close the lid gently, and centrifuge for 2 min at ≥8000 x *g* (≥10,000 rpm) to wash the spin column membrane. Discard the collection tube with the flow-through.**

   Prepare the 80% ethanol with ethanol (96–100%) and the RNase-free water supplied with the kit.

   Note: After centrifugation, carefully remove the RNeasy MinElute spin column from the collection tube so that the column does not contact the flow-through. Otherwise, carryover of ethanol will occur.

* Flow-through contains Buffer RLT Plus or Buffer RW1 and is therefore not compatible with bleach. See page 6 for safety information.

10. **Place the RNeasy MinElute spin column in a new 2 ml collection tube (supplied). Open the lid of the spin column, and centrifuge at full speed for 5 min. Discard the collection tube with the flow-through.**

    To avoid damage to their lids, place the spin columns into the centrifuge with at least one empty position between columns. Orient the lids so that they point in a direction opposite to the rotation of the rotor (e.g., if the rotor rotates clockwise, orient the lids counterclockwise).

    It is important to dry the spin column membrane, since residual ethanol may interfere with downstream reactions. Centrifugation with the lids open ensures that no ethanol is carried over during RNA elution.

11. **Place the RNeasy MinElute spin column in a new 1.5 ml collection tube (supplied). Add 14 µl RNase-free water directly to the center of the spin column membrane. Close the lid gently, and centrifuge for 1 min at full speed to elute the RNA.**

    As little as 10 µl RNase-free water can be used for elution if a higher RNA concentration is required, but the yield will be reduced by approximately 20%. Do not elute with less than 10 µl RNase-free water, as the spin column membrane will not be sufficiently hydrated.

    The dead volume of the RNeasy MinElute spin column is 2 µl: elution with 14 µl RNase-free water results in a 12 µl eluate.

    For RT-PCR and real-time RT-PCR with the purified RNA, QIAGEN offers a range of optimized, ready-to-use kits that provide highly specific and sensitive results. For details, visit www.qiagen.com/PCR . For whole transcriptome amplification (WTA) of limited amounts of RNA, we recommend the QuantiTect Whole Transcriptome Kit. For details, visit www.qiagen.com/goto/WTA .

7.6 QuantiTect Reverse Transcription protocol (Qiagen)

QuantiTect® Reverse Transcription Kit

The QuantiTect Reverse Transcription Kit (cat. nos. 205310, 205311, 205313, and 205314) should be stored immediately upon receipt at –20°C in a constant-temperature freezer.

For more information, please refer to the *QuantiTect Reverse Transcription Handbook*, which can be found at www.qiagen.com/handbooks.

For technical assistance, please call toll-free 00800-22-44-6000, or find regional phone numbers at www.qiagen.com/contact.

Notes before starting

- Dissolve any precipitates in gDNA Wipeout Buffer by vortexing. If necessary, briefly incubate at 37°C until the precipitates dissolve.
- Set up all reactions on ice to minimize the risk of RNA degradation.
- RNase inhibitor and dNTPs are already included in the kit components. Do not add additional RNase inhibitor or dNTPs.
- RT Primer Mix (supplied) or gene-specific primers (not supplied) should be used. RT Primer Mix is optimized to provide high cDNA yields for all regions of RNA transcripts. For gene-specific primers, we recommend using a final concentration of 0.7 µM.
- Separate denaturation and annealing steps are not necessary before starting the reverse-transcription reaction.
- If using a reaction volume of 200 µl or greater for reverse transcription, make sure the reaction tube is efficiently heated (e.g., if using a heating block, carefully fill each well with a drop of water so that heat can be efficiently transferred from the block to the tube).
- After reverse transcription, the reaction must be inactivated by incubation at 95°C for 3 min.
- If working with RNA for the first time, refer to Appendix A of the *QuantiTect Reverse Transcription Handbook*.
- If you have purchased the QuantiTect Reverse Transcription Kit in order to perform additional reverse-transcription reactions with the FastLane® Cell cDNA Kit, follow the protocol in the *FastLane Cell cDNA Handbook*. Do not follow this protocol.

January 2011

1. Thaw template RNA on ice. Thaw gDNA Wipeout Buffer, Quantiscript® Reverse Transcriptase, Quantiscript RT Buffer, RT Primer Mix, and RNase-free water at room temperature (15–25°C). Mix each solution by flicking the tubes. Centrifuge briefly to collect residual liquid from the sides of the tubes, and then keep on ice.

2. Prepare the genomic DNA elimination reaction on ice according to Table 1. Mix and then keep on ice.

   Note: If setting up more than one reaction, prepare a master mix of gDNA Wipeout Buffer and RNase-free water with a volume 10% greater than that required for the total number of reactions to be performed. Distribute the appropriate volume of master mix into individual tubes, followed by each RNA sample.

   Note: The protocol is for use with 10 pg to 1 µg RNA. If using >1 µg RNA, scale up the reaction linearly. For example, if using 2 µg RNA, double the volumes of all reaction components for a final 28 µl reaction volume.

Table 1. Genomic DNA elimination reaction components

Component	Volume/reaction
gDNA Wipeout Buffer, 7x	2 µl
Template RNA, up to 1 µg*	Variable
RNase-free water	Variable
Total reaction volume	14 µl

* This amount corresponds to the entire amount of RNA present, including any rRNA, mRNA, viral RNA, and carrier RNA present, and regardless of the primers used or cDNA analyzed.

3. Incubate for 2 min at 42°C, then place immediately on ice.

   Note: Do not incubate at 42°C for longer than 10 min.

4. Prepare the reverse-transcription master mix on ice according to Table 2. Mix and then keep on ice. The reverse-transcription master mix contains all components required for first-strand cDNA synthesis except template RNA.

   Note: If setting up more than one reaction, prepare a volume of master mix 10% greater than that required for the total number of reactions to be performed. Distribute the appropriate volume into individual tubes.

   Note: If using >1 µg RNA, scale up the reaction linearly. For example, if using 2 µg RNA, double the volumes of all reaction components for a final 40 µl reaction volume.

Table 2. Reverse-transcription reaction components

Component	Volume/reaction
Reverse-transcription master mix Quantiscript Reverse Transcriptase*	1 µl
Quantiscript RT Buffer, 5x†‡	4 µl
RT Primer Mix‡	1 µl
Template RNA Entire genomic DNA elimination reaction (step 3)	14 µl (added at step 5)
Total reaction volume	20 µl

* Also contains RNase inhibitor.

† Includes Mg^{2+} and dNTPs.

‡ For convenience, premix RT Primer Mix and 5x Quantiscript RT Buffer in a 1:4 ratio if RT Primer Mix will be used routinely for reverse transcription. This premix is stable when stored at –20°C. Use 5 µl of the premix per 20 µl reaction.

5. Add template RNA from step 3 (14 µl) to each tube containing reverse-transcription master mix. Mix and then store on ice.

6. Incubate for 15 min at 42°C.

   Note: In some rare cases (e.g., if the RT-PCR product is longer than 200 bp or if analyzing RNAs with a very high degree of secondary structure), increasing the incubation time up to 30 min may increase cDNA yields.

7. Incubate for 3 min at 95°C to inactivate Quantiscript Reverse Transcriptase.

8. Place the reverse-transcription reactions on ice and proceed directly with real-time PCR. For long-term storage, store reverse-transcription reactions at –20°C.

   Note: For details on performing real-time PCR after reverse transcription, refer to Appendix C of the *QuantiTect Reverse Transcription Handbook*. For details on appropriate controls, see Appendix D. We recommend using a Rotor-Gene® Kit, QuantiFast® Kit, or QuantiTect Kit for real-time PCR.

7.7 FastStart Essential DNA green master protocol (Roche)

LightCycler® 96 Instrument Protocol

The following procedure is optimized for use with the LightCycler® 96 Instrument.

Run Editor				
Detection Format			Reaction Volume [µl]	
Dyes 1: SYBR Green I			20	
Programs				
	Temp. [°C]	Ramp [°C/s]	Duration [s]	Acquisition Mode
Pre-Incubation	95	4.4	600[2]	None
3-Step Amplification	No. of Cycles: 45			
	95	4.4	10[1][3]	None
	60 primer dependent[4]	2.2	10[1][3]	None
	72	4.4	10[1][3][5]	Single
Melting	95	4.4	10	None
	65	2.2	60	None
	97	0.1	1	5 Readings/°C

[1] For well-established assays you may shorten the amplification times to: 95°C for 10 seconds, 60°C for 10 seconds, and 72°C for 10 to 20 seconds. Forty-five cycles are suitable for most assays. If the assay is optimized and has steep amplification curves and early crossing points (even when target concentrations are low), 40 cycles should be sufficient. Reducing the number of cycles will reduce the time required for the assay (fast protocol).

[2] For some assays, a pre-incubation of 300 seconds is sufficient (fast protocol). However, if high polymerase activity is required in early cycles, a 600-second period is recommended, especially for higher reaction volumes and when working with unpurified cDNA samples as template. Do not use more than 2 µl unpurified cDNA sample.

[3] For greater precision in target quantification experiments, it can be advantageous (in some cases) to choose longer annealing and extension times for the amplification cycles. This is especially recommended for higher reaction volumes.

[4] For initial experiments, set the target temperature (the primer annealing temperature) 5°C below the calculated primer Tm.

[5] Calculate the hold time for the PCR elongation step by dividing the amplicon length by 10 (*e.g.,* a 150 bp amplicon requires 15 seconds elongation time). Do not exceed the hold time for elongation below 10 seconds.

Preparation of the PCR Mix

Follow the procedure below to prepare one 20 µl standard reaction.

⚠ ***Always wear gloves during handling.***

1. Thaw one vial of FastStart Essential DNA Green Master (Vial 1) and Water, PCR Grade (Vial 2).
   – Briefly spin vials in a microcentrifuge before opening to ensure recovery of all the contents.
   – Mix carefully by pipetting up and down and store on ice.

   ⚠ ***Keep the Master Mix protected from light.***

2. Prepare a 10x concentrated solution of the PCR primers.

3. In a 1.5 ml reaction tube on ice, prepare the PCR mix for one 20 µl reaction by adding the following components in the order listed below:

Reagent	Volume (µl)
Water, PCR Grade (Vial 2)	3
PCR Primer, 10x conc.	2
Master Mix, 2x conc. (Vial 1)	10
Total Volume	**15**

ⓘ *To prepare the PCR mix for more than one reaction, multiply the amount in the "Volume" column above by z, where z = the number of reactions to be run + sufficient additional reactions.*

4 Mix carefully by pipetting up and down. Do not vortex.
- Pipette 15 µl PCR mix into each reaction vessel of a LightCycler® 8-Tube Strip or LightCycler® 480 Multiwell Plate.
- Add 5 µl of the DNA template.
- Close the reaction vessels.

5 Place the LightCycler® 480 Multiwell Plate in a standard swinging-bucket centrifuge with suitable adapter.
- Balance it with a suitable counterweight, such as another LightCycler® 480 Multiwell Plate, or
- Place the 8-Tube Strips into a standard multiwell plate 96 (MWP) and balance them in the centrifuge.
- Centrifuge at 1,500 × *g* for 0.5 to 2 minutes.

6 Load the reaction vessels into the the LightCycler® Nano or LightCycler® 96 Instrument.

7 Start the PCR program described above.

If you use reaction volumes different from 20 µl, it may be advantageous to adapt the hold times of all amplification steps.

Two-Step RT-PCR

FastStart Essential DNA Green Master can also be used to perform two-step RT-PCR. In two-step RT-PCR, the reverse transcription of RNA into cDNA is separated from the other reaction steps and is performed outside the LightCycler® Nano or the LightCycler® 96 System. Subsequent amplification and online monitoring is performed according to the standard LightCycler® Nano or LightCycler® 96 System procedure, using the cDNA as the starting sample material. The Transcriptor First Strand cDNA Synthesis Kit* is recommended for reverse transcription of RNA into cDNA. Synthesis of cDNA is performed according to the instructions provided with the kit.

⚠ ***For initial experiments, we recommend running undiluted, 1:10 diluted, and 1:100 diluted cDNA template in parallel to determine the optimum template amount. If you use undiluted cDNA as template, use a 10-minute pre-incubation.***

2.3. Other Parameters

Prevention of Carryover Contamination

Uracil DNA N-Glycosylase (UNG) is suitable for preventing carryover contamination in PCR. This carryover prevention technique involves incorporating deoxyuridine triphosphate (dUTP, a component of the Master Mix in this kit) into amplification products, then pretreating later PCR mixtures with UNG. If a dUTP-containing contaminant is present in the later PCRs, it will be cleaved by a combination of the UNG and the high temperatures of the initial denaturation step; it will not serve as a PCR template.

⚠ ***To ensure optimal results in carryover prevention reactions with the FastStart Essential DNA Green Master, always use LightCycler® Uracil-DNA-Glycosylase*.***

Follow the Instructions for Use for the enzyme.Since your target DNA template contains thymidine rather than uridine, it is not affected by this procedure.

The use of UNG may influence the melting temperature (Tm) in melting curve analysis.

7.8 Primer Overview

Primers

	Oligo Name	Sequence 5’- 3’	Primer pair
2	Fwd ncbi 1 CDKN2D	TCCCAGCCCCTCGCC	
3	Rev ncbi 1 CDKN2D	TGCCAAACATCATGACCTGC	
4	fwd eur 1 2 3 CDKN2D	CTCAACCGCTTCGGCAAGAC	2
5	rev eur 1 CDKN2D	GCTGCCAGAAAGCTGACCAC	2
6	rev eur 2 CDKN2D	CTGACCACAGCAGTGTGACC	
7	rev eur 3 CDKN2D	TGACCACAGCAGTGTGACCC	
8	Fwd CKMT1B 1	AGGCAAATCAGAGGTGGAGC	5
9	Rev CKMT1B 1	GTGTGGATGACAGGTGTGGG	5
10	Fwd CKMT1B 2	TCCCAAAGATCCTGGAGAAC	
11	Rev CKMT1B 2	GCCATCATTAGAATGAGCAGAAC	
12	fwd CKMT1B 3	CGAGGAAGGCAGAGATTCCC	
13	Rev CKMT1B 3	TACTCAGCGCTCGGGGGATA	
14	Fwd GDAP1L1 1 2 3	TGGACCAGATTGAGGCGGAG	
15	Rev GDAP1L1 1	CGGAAAGCATTGGGGATGAC	
16	Rev GDAP1L1 2	GGACATCAGCGAGGGTGAAG	
17	Rev GDAP1L1 3	GTGAAGGCACAGCCACAGAG	
18	Fwd MTSS1 1 3	TACAGCACCCAGACAACCAC	13
19	Rev MTSS1 1 2	CTGGACGCTTGGCTTGGAAC	
20	Fwd MTSS1 2	GGCTACAGCACCCAGACAAC	
21	Rev MTSS1 3	CGAACTCCTGCTGATCTGCC	13
22	Fwd RGS9 1 3	CTTCTCCTCTTGTCTCCCAC	14
23	Rev RGS9 1 2	GACCCTCTGGTTCTGCATTC	14
24	Fwd RGS9 2	GATGACAATCCGACACCAAG	
25	Rev RGS9 3	GGGTTCTGCATGTCCTTCAC	
26	Fwd TRAPPC6B 1 3	AATCAGAGGTGGCTTATCAAAC	19
27	Rev TRAPPC6B 1	ACTCCTGTACAAACATCGAAC	
28	Fwd TRAPPC6B 2	TGTGGCTTAATCAGAGGTGG	
29	Rev TRAPPC6B 2 3	GCCTTGCATTTCAGTATGTTC	19
30	Fwd TUBBA4A 1	GAACATGGCATCGACCCCAC	

31	Rev TUBB4A 1 2	GGAACTCCTCGCGGATCTTAC	21
32	Fwd TUBBA4A 2	TCGACCCCACAGGCACATAC	21
33	Fwd PCLO 1	ATTCCAGTCAGAGCAGCCAG	
34	Rev PCLO 1	AGACAAAGGGACAGAGAACAG	
35	Fwd PCLO 2	CCCTGACCCATCAAAGGACAT	
36	Rev PCLO 2	AAGCTGCCATGCTGAGGAAT	
37	Fwd PCLO 3	GACCCATCAAAGGACATGCAG	
38	Rev PCLO 3	TAAGCTGCCATGCTGAGGAA	
39	Fwd SCN9A 1 2 3	CCCTCAGACACTACTACTTCAC	27
40	Rev SCN9A 1	ATCCCCTTTGCTCCTTTGAC	
41	Rev SCN9A 2	TAAACAACGCAGGAAGGGAC	
42	Rev SCN9A 3	CAAGACGGATCACTCGGAAC	27
43	Fwd PCLO 1 2	ATCCAGTTCCGTTCCCAGCC	29
44	Rev PCLO 1	TCTTGTACTCAGCACTTGCATTC	
45	Rev PCLO 2 3	ACAACCATGACTTGACCTCTCC	29
46	Fwd PCLO 3	AGCAAAAAGAAGCACGGCAG	
47	Fwd PCLO 4	ACAAATTTAGCAGAAGCTGGAC	
48	Rev PCLO 4	CCTCAAAACACTCCCCACAC	
49	Fwd GADP1L1 1	GCGACAACATCATCAGTGAC	32
50	Rev GADP1L1 1	CGTACTTGGGGATCATGGAG	32
51	Fwd GADP1L1 2	CTTTCCTGCCTCTGATTCCG	
52	Rev GADP1L1 2	GCATGAACCAGGGCTCCTT	
53	Fwd TUBB4A 1	ACTGCAAGGATTCTGGCGAA	
54	Rev TUBB4A 1	TCGATGCCATGTTCGTCACT	
55	Fwd TUBB4A 2 3	TCGCGGTTACTGCAAGGATT	
56	Rev TUBB4A 2	GTGGGGTCGATGCCATGTTC	
57	Rev TUBB4A 3	GTCGATGCCATGTTCGTCAC	
58	LC-hPOLR2A-F3932	TGCCACAGACAGACAACAAG	ref
59	LC-hPOLR2A-R4158	GACATAGGAGCCATCAAAGGAG	ref
60	Fwd GAPDH	AGGGCTGCTTTTAACTCTGGT	Extra ref
61	Rev GAPDH	CCCCACTTGATTTTGGAGGGA	Extra ref

Primer pair	Primers	Amplicon from gene:	Pair selected for the final assay	Exon-exon boundry	Accession number	PCR fragment length
1	2+3	CDKN2D				
2	4+5	CDKN2D	X	1-2(last)	NM_001800.3	272 bp
3	4+6	CDKN2D				
4	4+7	CDKN2D				
5	8+9	CKMT1B	X	9-10(last)	NM_020990.4	120 bp
6	10+11	CKMT1B				
7	12+13	CKMT1B				
8	14+15	GDAP1L1				
9	14+16	GDAP1L1				
10	14+17	GDAP1L1				
11	18+19	MTSS1				
12	19+20	MTSS1				
13	18+21	MTSS1	X	14-15(last exon)	NM_001282971.1	115 bp
14	22+23	RGS9	X	1-2(tot.exons:19)	NM_001081955.2	178 bp
15	24+23	RGS9				
16	22+25	RGS9				
17	26+27	TRAPPC6B				
18	28+29	TRAPPC6B				
19	26+29	TRAPPC6B	X	5-6(last exon)	NM_001079537.1	123 bp
20	30+31	TUBB4A				
21	32+31	TUBB4A	X	3-5 (last exon)	NM_001289123.1	390 bp
22	33+34	PCLO				
23	35+36	PCLO				
24	37+38	PCLO				
25	39+40	SCN9A				
26	39+41	SCN9A				
27	39+42	SCN9A	X	26-27(last exon)	NM_002977.3	141 bp
28	43+44	PCLO				
29	43+45	PCLO	X	13-16(Tot.exons:20)	NM_014510.2	264 bp
30	45+46	PCLO				
31	47+48	PCLO				
32	49+50	GDAP1L1	X	2-3(Tot.exons:6)	NM_001256737.1	276 bp
33	51+52	GDPA1L1				
34	53+54	TUBB4A				
35	55+56	TUBB4A				
36	55+57	TUBB4A				
polr2a	58+59	POLR2A	X	24-25(Tot.exons:29)	NM_000937.4	248 bp
Gapdh	60+61	GAPDH	X	2-4 (Tot.exons:8)	NM_001256799.2	206 bp

7.9 Raw Ct values

Sample:		**1** Ctr 1+ Ctrl 2	**1** Scz 1+ Scz2	**2** Ctr 1+ Ctrl 2	**2** scz 1+ scz 2	**3** Ctr 1+ Ctrl 2	**3** scz 1+ scz 2	
Ctr 1	A1	0	0	0	0	0	0	
	B1	25,32	29,93	23,76	25,17	21,99	29,21	
	C1	25,5	29,05	23,64	24,58	21,36	29,6	
	D1	25,16	28,17	23,51	24,6	19,87	29,57	POLR2A
	E1	37	0	0	0	0	0	
	F1	36,2	27,48	25,57	24,73	17,8	23,68	
	G1	35,16	27,6	25,62	24,66	17,52	23,93	
	H1	33,9	27,52	24,87	24,1	18,24	23,53	
	A2	0	37	0	0	37,5	37	
	B2	30,82	35,92	28,57	30,16	37,96	35,67	
	C2	30,9	37	28,71	30,02	24,95	34,17	
	D2	30,78	35,3	28,85	30,55	25,47	34,06	CDKN2D
	E2	36,14	0	0	0	31,89	37	
	F2	35,16	33,28	31,07	31,67	22,14	29,55	
	G2	34,14	33,76	31,95	31,51	22,69	30,14	
	H2	33,56	33,74	30,55	29,23	22,81	28,57	
	A3	0	37	0	0	0	0	
	B3	23,89	24,2	25,54	26,44	25,47	27,57	
	C3	23,76	24,21	25,48	26,34	25,51	27,47	
	D3	23,69	24,18	25,34	26,21	24,84	27,46	CKMT1B
	E3	31,2	37	0	0	38,41	0	
	F3	27,19	21,7	24,84	26,07	23,98	28,95	
	G3	27,15	21,7	24,91	25,88	23,81	29,2	
	H3	27,13	21,74	24,83	26,2	24	29,13	
	A4	0	37	0	0	33,94	37	
	B4	28,51	30,86	27,02	27,04	24,97	29,63	
	C4	0	31,6	26,99	27,09	24,74	29,6	
	D4	0	31,22	27,02	26,92	24,46	29,61	MTSS1
	E4	27,75	37	0	37	31,77	0	
	F4	27,75	30,14	27,65	26,96	22,46	28,42	
	G4	27,2	29,84	27,57	26,91	22,32	28,24	
	H4	26,33	29,92	27,88	26,71	23,79	28,32	
	A5	0	37	0	0	0	0	
	B5	29,06	31,83	30,06	30,46	25,84	32,63	
	C5	28,79	31,11	29,66	30,83	25,71	32,1	
	D5	28,09	31,02	29,96	29,65	25,4	32,01	RGS9
	E5	0	0	37	0	0	0	
	F5	34,02	30,75	30,7	30,31	24,61	29,11	
	G5	34,03	31,71	30,77	29,97	23,62	29,16	
	H5	33,14	31,04	31,94	29,06	23,7	28,86	
	A6	0	0	0	0	0	0	
	B6	26,51	26,73	25,11	25,32	20,95	27,48	
	C6	26,59	26,79	25,58	25,9	20,79	27,72	
	D6	26,45	26,79	24,97	25,75	20,87	27,34	TRAPPC6B
	E6	25,35	0	0	0	0	0	
	F6	0	24,26	26,3	25,54	19,35	24,7	
	G6	34,28	24,32	26,52	25,48	19,59	24,71	
	H6	35,3	24,25	26,04	25,34	19,69	24,61	