BUILDING AN INS-1 cDNA LIBRARY FOR A GENOME-WIDE CRISPR-Cas9 SCREEN

Idongesit Ekpo

Follow this and additional works at: https://scholarsarchive.byu.edu/studentpub_uht

BYU ScholarsArchive Citation
Ekpo, Idongesit, "BUILDING AN INS-1 cDNA LIBRARY FOR A GENOME-WIDE CRISPR-Cas9 SCREEN" (2020). Undergraduate Honors Theses. 163. https://scholarsarchive.byu.edu/studentpub_uht/163

This Honors Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
ABSTRACT

BUILDING AN INS-1 cDNA LIBRARY FOR A GENOME-WIDE CRISPR-Cas9 SCREEN.

Idongesit Ekpo
Department of Chemistry and Biochemistry
Bachelor of Science

By the year 2040, an estimated 642 million people are expected to have diabetes globally. Diabetes results from an elevation of metabolic stressors, such as glucotoxicity, lipotoxicity, oxidative stress and apoptosis. In type 2 diabetes, these stressful conditions contribute to the malfunction and loss of functional insulin-producing β-cells. Current treatment methods for diabetes include insulin therapy, islet transplant and anti-diabetes medication. These treatments are not curative and ignore other factors that contribute to the pathogenesis of diabetes beyond insulin resistance and islet β-cell failure. Previous research on β-cells has focused on ways to replace functional β-cell mass, trigger β-cell proliferation, protect β-cells from stressors and enhance glucose-stimulated insulin secretion. While these treatments are effective, they neglect the possibility of unknown contributing factors. A more effective method is to probe the genetic variation presented in diabetics so that scientists can understand the disease better and develop curative methods. One way of doing this is through gene editing using clustered regular interspaced short palindromic repeats (CRISPR)-CRISPR associated protein 9(Cas9).

CRISPR-Cas 9 is a gene-editing tool that modifies DNA using a guide RNA and an endonuclease called Cas9. The guide RNA directs the Cas9 protein to cut a target
DNA sequence that is complementary to its RNA sequence. This guided mechanism makes CRISPR-Cas9 the most precise tool for genetic manipulation. Because of its convenience, CRISPR-Cas9 has been used to create many forward genetic screens. Using this knowledge of the connection between β-cells and diabetes, and the CRISPR-Cas9 mechanism, we enzymatically generated a CRISPR guide-RNA library of the rat pancreatic islet β-cell insulinoma (INS-1) cell line. This library can be used to create a CRISPR-Cas9 knockout (KO) forward genetic screen of all the genes in the INS-1 cell line required for insulin secretion, β-cell viability, proliferation, and growth. In the future, this screen will help us identify genes involved in the various mechanistic pathways that contribute to diabetes. For the library synthesis, we will use the total RNA from INS-1 cells to generate the guide RNA libraries.
ACKNOWLEDGMENTS

I am grateful to my advisor Dr Tessem for his support, guidance, and feedback he has provided throughout this project. I would like to thank Dr Hill for his continuous guidance and coaching. I am grateful to everyone in the Tessem and Hill labs who contributed their time and efforts to this work especially Courtney Smith, Kyle Hendricks, Nathaniel Barton, and Joshua Yates. Finally, I would like to thank my family and friends for constantly supporting me especially on hard days.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>i</td>
</tr>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>vi</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>viii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>x</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xi</td>
</tr>
<tr>
<td>I. Background</td>
<td>1</td>
</tr>
<tr>
<td>II. Materials and Methods</td>
<td>5</td>
</tr>
<tr>
<td>III. Results</td>
<td>12</td>
</tr>
<tr>
<td>IV. Discussion</td>
<td>18</td>
</tr>
<tr>
<td>V. References</td>
<td>21</td>
</tr>
<tr>
<td>VI. Appendices</td>
<td>27</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

FIGURE 1: Brief Illustration of The Mechanism of the CRISPR-Cas9 system in INS-1 β-cells……………………………………………………………………………………......4

FIGURE 2: Simple Illustration of Main Experimental Steps ……………………………6

FIGURE 3: Total INS-1 RNA Post Isolation …………………………………………...12

FIGURE 4: Modifications to Amplification Conditions Yield High-Quality INS-1 cDNA…………………………………………………………………………………………14

FIGURE 5: INS-1 cDNA Normalization Shows Effective Normalization for Some cDNA Samples………………………………………………………………………………16

FIGURE 6: INS-1 CRISPR Single Guide DNA library……………………………...….…17

FIGURE 7: Future INS-1 CRISPR/Cas9 screen workflow ……………………………20
LIST OF TABLES

TABLE 1: Primer Sequences for Normalization qPCR Analysis..........................10

TABLE 2: Modified INS-1 cDNA Synthesis Amplification Conditions...............13

TABLE 3: ΔCt Values Showing Failed Normalization .................................15

TABLE 4: Modified INS-1 cDNA Normalization Amplification Conditions .........16

TABLE 5: ΔCt Values Showing Efficient Normalization ..............................17
BACKGROUND

An estimated 415 million people worldwide suffer from diabetes, with the vast majority suffering from type 2 diabetes (T2D).\textsuperscript{1,2} T2D is characterized by hyperglycaemia and hyperlipidaemia due to muscle, adipose, and liver tissue insulin resistance.\textsuperscript{3,4} Insulin resistance usually has a detrimental effect on β-cell function and viability.\textsuperscript{5} β-cells are endocrine cells found in the islet of Langerhans in the pancreas. They are responsible for producing, storing and secreting insulin.\textsuperscript{6} Under high blood glucose conditions, β-cells secrete insulin, but insulin resistance compromises this response.\textsuperscript{7,8}

Insulin resistance throughout the body leads to overworking and exhaustion of β-cells resulting in the loss of functional β-cell mass.\textsuperscript{9,10} Studies indicate that decreased functional β-cell mass is an early and essential step in T2D disease progression. The β-cell dysfunction observed in T2D results in decreased and poorly controlled insulin secretion and ultimately leads to β-cell death.\textsuperscript{9} The loss of β-cell function and mass is driven by various metabolic and stress-related factors including elevated glucotoxicity, and lipotoxicity.\textsuperscript{9} Consequently, lipotoxicity and glucotoxicity function as both a cause and a symptom of T2D, representing a destructive positive feedback loop. Strategies for replacing functional β-cell mass as a cure for diabetes require some means of protecting β-cells from continued stress and damage from hyperglycaemia, hyperlipidaemia, and oxidative stress. Understanding the effects of these stressors and their mechanisms of action on the β-cells is vital to developing a potential cure for T2D.

To understand the connection between the β-cell transcriptome and diabetes, Kutlu et al.\textsuperscript{11} generated a detailed transcriptome of pancreatic β-cells called the Beta Cell Gene Atlas (BCGA) using massively parallel signature sequencing (MPSS) and
microarrays. While this inventory has been used to improve β-cell GWAS (genome-wide associated studies) it has limitations. Some of these limitations include an inability to identify 7% of genes that lack the DpnII restriction site, generation of false positives and interference of high G+C content in MPSS.

One method that overcomes the limitations from previous studies is gene editing. An effective gene-editing tool is the *Streptococcus pyogenes* clustered regular interspaced short palindromic repeats (CRISPR)-CRISPR associated protein9 (Cas9) system. CRISPR-Cas9 is a bacterial adaptive immune response system used to attack invading pathogens. It consists of a non-coding RNA called the single guide RNA (sgRNA) and an endonuclease known as Cas9. The sgRNA directs Cas9 to the target DNA sequence by locating and binding its complementary DNA sequence. For Cas9 to function, the target DNA must have an adjacent PAM (protospacer adjacent motif) site. The PAM sequence for the *S.pyogenes* Cas9 is 5'-NGG-3'. Unlike other genome editing tools such as the zinc finger nucleases (ZFN) and the Transcription activator-like effector nucleases (TALENs), the CRISPR-Cas9 system is precise, easy to design, and able to modify several genes simultaneously. Figure 1 contains a brief illustration of how the CRISPR-Cas9 system works.

Current improvements in the CRISPR-Cas9 technology has made gene modification on a genome-wide scale feasible in mammalian cells. Using a kit that enzymatically generates CRISPR sgRNA libraries, we created a CRISPR sgRNA library from the INS-1 832/13 cell-line. We hypothesise that it will have around 200,000 distinct sgRNAs targeting about 7000 genes. This will allow us to target and knock out any gene in the INS-1 genome required for β-cell functionality. To generate this CRISPR
sgRNA library, we isolated total RNA from the INS-1 cells and synthesised cDNA from
the RNA. After this, we normalized the cDNA and quantified the normalization. Then
using sgRNA library assembly by ligation on magnetic beads (SLALOM™) technology
for enzymatically generating CRISPR sgRNA libraries, we created an INS-1 guide RNA
library. This INS-1 CRISPR sgRNA library will be used to generate a genome-wide
forward genetic screen of genes directly involved in the T2D pathway. Our study uses the
knowledge of the mechanics of the CRISPR-Cas9 tool to enzymatically generate an INS-
1 CRISPR sgRNA library. This library will assist with our subsequent plans to study the
influence of genetics on hyperlipidaemia, hyperglycaemia, insulin secretion, and growth
and viability of \( \beta \)-cells.

We aimed to generate an INS-1 CRISPR sgRNA library that can be used to
identify all \( \beta \)-cell genes that play a role in the pathways of different metabolic stressors
and conditions that are characteristic of T2D. With this library, we can look towards
possible curative approaches to treating diabetes.
Figure 1. Brief Illustration of The Mechanism of the CRISPR-Cas9 system in INS-1 β-cells. i) A crRNA that is complementary to the DNA target, and a tracrRNA sequence that associates with the Cas9 protein make up the single guide RNA (sgRNA). ii) This sgRNA binds the Cas9 protein. iii) The Cas9-sgRNA causes a target-specific double-stranded DNA cleavage. iv) The cleavage site is repaired by nonhomologous end joining (NHEJ). Adapted from Takarabio.com.
MATERIALS AND METHODS

Experimental Design Overview

This project aimed to generate an INS-1 CRISPR sgRNA library that will use the CRISPR-Cas9 technology to find β-cell genes involved in the pathways of metabolic pressures like glucotoxicity, lipotoxicity, oxidative stress, and apoptosis in β-cells. This project has laid a foundation for future studies involving the identification of these genes and their specific contributions to these stress pathways.

We cultured INS-1 832/13 cells in complete RPMI 1640 medium. The cell culture conditions are included in the cell passaging section of this paper. Then we harvested these cells at a passage number of 52. This low passage number ensures that the morphology, sensitivity, growth rate, and transfection efficiency of these cells are not altered.19

Following the cell harvest, we isolated total RNA from these cells and used it to synthesize cDNA using the two-step reverse transcriptase PCR from Evrogen. Then we normalized the cDNA to achieve equal distribution of gene transcripts. We quantified this normalization using qPCR and analysed the results. Finally, we generated a CRISPR sgRNA library with the normalized cDNA. Figure 2 shows a simple illustration of this project.
**Cell Passaging**

We cultured INS-1 832/13 cells from a stock solution at a passage number of 50 in RPMI 1640 (+L-glutamine) supplemented with 2 g Sodium Bicarbonate, 10 ml penicillin/streptomycin, 10 ml 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), 10% foetal bovine serum (FBS) and 22 ml Ins-1 supplement. These cells were grown in a T75 flask at 37°C and in a 5% CO₂ incubator. The cells underwent two more passages before we prepared them for total RNA isolation. At the time of total RNA isolation, these cells were healthy and at passage number of 52. INS-1 cells are efficient under 100 passages.

**Total RNA Isolation**

We harvested the cells by aspirating the culture media and washing them twice with 5 ml of phosphate buffered saline (PBS). After the PBS washes, we treated the cells with 2 ml of trypsin for seven minutes. Then, we added 8 ml of PBS to the culture flask. We aliquoted this cell solution into ten marked Eppendorf tubes each containing 1ml of the cell solution. We centrifuged these tubes, aspirated the PBS and added 0.5 ml of Sigma Aldrich TRI Reagent® solution. This cell harvesting protocol is an adaptation of
the one found on the Tessem Lab Website. Using Zymo Research Direct-zol™ RNA MiniPrep Kit, we isolated total RNA from the INS-1 cells following the manufacturer’s protocol. Then we confirmed the quality and concentration of isolated RNA using Thermo Fisher Scientific™ NanoDrop™ One/OneC Microvolume UV-Vis Spectrophotometer and gel electrophoresis. We determined the purity of the RNA samples through their 260/280 and 260/230 absorbance ratios. 260/280 and 260/230 ratios of 2.0±0.2 shows that the RNA samples are pure. Denaturing agarose gels of total mammalian RNA shows two bright 28S and 18S rRNA bands. We used a non-denaturing agarose gel to check the integrity of our RNA samples and confirmed these two bands.

**cDNA Synthesis**

From the isolated total RNA, we synthesized cDNA using the Evrogen Mint-2 cDNA synthesis kit, according to the manufacturer’s protocols. For the first strand synthesis, we used the 5’-end adapter PlugOligo-1 5’-AAGCAGTGTTATCAACGCAG AGTACGCGGG-3’ and the 3’-end adapter CDS-1 5’-AAGCAGTGTTATCAACGCAG AGTAC(T)30VN -3’. We chose these adapters because they lacked the 5’-CCGG-3’ sequence which would interfere with our CRISPR gRNA library generation as this technology relies on the 5’-CCGG-3’ HpaII site to generate a CRISPR gRNA library. These adapters also have a poly(T) sequence that is necessary for synthesising cDNA from mRNA.

For the second strand amplification, we made some modifications to the conditions suggested in the manufacturer’s protocol as shown in table 2. The average
number of cycles that we needed to amplify the second strand cDNA was 24. After the
cDNA synthesis, we purified the cDNA samples with the Zymo Research DNA Clean &
Concentrator™-5 Kit. Then we measured the concentration and quality of the cDNA
using a spectrophotometer. We also ran the cDNA on a 1.5% (w/v) agarose/EtBr gel in
1X TAE buffer. Colour adjustments were made on the gel images using ImageJ and the
gel images were labelled using Microsoft PowerPoint.

**cDNA Normalization**

Following the generation of the double-stranded(ds) cDNA from total RNA, we
normalized the cDNA using the Evrogen Trimmer-2 cDNA normalization kit. The
normalization step was important because the transcript concentrations of all the genes
must be equalized for this CRISPR sgRNA library to be used to generate a genome-wide
screen that is representative of the INS-1 β-cell transcriptome. High abundance
transcripts (thousands of mRNA copies per cell) that encode about 10 genes usually make
up 20% of the cellular mRNA. Transcripts in medium abundance (hundreds of mRNA
copies per cell) make up 40-60% of the cellular mRNA and the low abundance transcripts
make up the remaining 20-40% of the cellular mRNA. Without normalization, we
would get a library that is representative of only genes with high abundance transcripts.
Hence, the normalization process cannot be overlooked. This kit uses a duplex-specific
nuclease-based normalization process described by Zhulidov et al. to equalize the
transcript concentrations.

The normalization process involves three steps. The first involves denaturing the
ds cDNA at 98ºC and hybridizing it in at 68ºC for 6 hours. The next step involves
treating the hybridized cDNA with different concentrations (1x DSN, ¾x DSN, ½x DSN)
of the duplex-specific nuclease (DSN) at 68°C to find the optimal concentration of DSN needed for the normalization. For the last step, we amplified the normalized cDNA using PCR conditions that were modified from the manufacturer’s protocol. These modifications can be found in table 4. We amplified the normalized cDNA for 22 cycles in total. The number of cycles must be determined independently for each PCR tube due to experimental variations. The normalized cDNA was run on a 1.5% (w/v) agarose/EtBr gel in 1X TAE buffer. Colour adjustments were made on the gel images using ImageJ and the gel images were labelled using Microsoft PowerPoint.

**Measuring normalization using qPCR**

To ensure that the cDNA was properly normalized, we ran a quantitative analysis using quantitative real-time PCR (qPCR). Because the normalization process involves equalising transcript concentrations in cDNA samples, efficient normalization will show a relative decrease in the number of high abundance transcripts and a relative increase in the number of low abundance transcript. For this quantification, it is important to select genes that have high, intermediate, and low abundance transcripts to measure normalization effectively. One difficulty that arose here was the absence of trustworthy data on the transcriptome of INS-1 832/13 cells. We decided to select a few medium to high-expressing genes based on their functionality in INS-1 β-cells. These genes are Homeobox protein Nkx-6.1 (*Nkx6.1*), Insulin, Peptidylprolyl isomerase A (*Ppia*), Enolase 1 (*Eno1*), Beta-actin (*Actb*), Glyceraldehyde 3-phosphate dehydrogenase (*Gapdh*) and Succinate dehydrogenase complex, subunit A (*Sdha*). After selection, we ordered the primers of these genes from Integrated DNA Technologies. The accession numbers and sequences of the primers used in this analysis are included in table 1.
We performed the quantitative analysis using iTaq™ Universal SYBR® Green Supermix and a Bio-Rad CFX Connect Real-Time PCR Detection System thermal cycler. For each primer, we conducted triplicate qPCR reactions of the control cDNA (non-normalized), non-template control (NTC), \( \frac{3}{4} x \), \( \frac{1}{2} x \) and \( \frac{1}{4} x \) DSN treatments. The total number of reactions for all seven primers was 84. Each reaction consisted of 1\( \mu \)l control cDNA/ NTC/ \( \frac{3}{4} x \) DSN normalized cDNA/ \( \frac{1}{2} x \) DSN normalized cDNA, 10\( \mu \)l 2× SYBR Green Master Mix, 2\( \mu \)l forward and reverse primer mix (10\( \mu \)M), and 8\( \mu \)l nuclease-free water. We ran the reactions under the following PCR conditions, initial denaturation at 95°C for 3 min, followed by 39 cycles of denaturation at 95°C for 10 s and annealing at 59°C for 1 min. At the end of each run, we conducted a melt curve analysis of each PCR product between 65°C to 90°C with 0.5°C increments.

Analysis of qPCR Data

The cycle threshold (Ct) is the number of qPCR cycles needed for the fluorescence of a PCR product to pass a certain threshold or to be detected above background noise.\(^{27}\) Low Ct values mean that a gene is expressed highly because a higher concentration of cDNA will be amplified earlier in the PCR. The difference between the average Ct values of the target (normalized cDNA) and the control, \( \Delta \text{Ct} (\Delta \text{Ct} = \text{Ctc} – \text{Ctn}) \)
Ct(N) determines how well the cDNA samples have been normalized. A ΔCt value ≥ 5 shows effective normalization of high-level transcripts. For this analysis, we used an R-script from the Hill lab to rearrange and generate figures from the qPCR data. This script is included in the appendices.

**Generation of INS-1 CRISPR-Cas9 sgRNA Library**

We used the new, affordable and time-efficient technology to generate the INS-1 CRISPR-Cas9 sgRNA library known as SLALOM™ developed in the Hill lab at Brigham Young University. Unlike other enzymatic methods for generating CRISPR sgRNA libraries, it is cheap, easy to use, and does not require an inordinate amount of DNA. We followed the manufacturer’s protocol for this sgRNA library synthesis and ran products on a 1.5% EtBr agarose gel. The manufacturer’s protocol recommended using a concentration of DNA substrate with at least 10 pmol HpaII cut sites. We calculated this concentration using the NEBioCalculator® Linear DNA: Mass to Moles of Ends Converter. The formula from this calculation is included in the appendices. Colour adjustments were made on the gel images using ImageJ and the gel images were labelled using Microsoft PowerPoint. The concentration of DNA was measured using a spectrophotometer.
RESULTS

cDNA Synthesis Yields High-Quality INS-1 cDNA After Modifications to Amplification Conditions.

We synthesized INS-1 cDNA from total INS-1 RNA sample 1 (fig. 3) using the Evrogen Mint-2 cDNA synthesis kit. This RNA sample has a 260/230 ratio of 2.18 and a 260/280 ratio of 1.97. Its concentration was 515.1 ng/ul. Usually, cDNA generated from total mammalian RNA has a moderate smear and several bright bands that indicate the presence of abundant transcripts.29 This cDNA synthesis is a two-part protocol. We generated the single-stranded(ss) cDNA first and then amplified the ss cDNA to get INS-1 cDNA. Before the ds cDNA synthesis, we ran an evaluative PCR to determine the optimal amplification conditions for the INS-1 cDNA. Initially, we followed the manufacturer’s protocol and did not notice bright cDNA bands when we ran the cDNA products on a 1.2% EtBr agarose gel (fig. 4a and b).
After repeating this experiment without noticing significant differences in our product quality, we decided to modify certain amplification conditions as shown in Table 2. We also decided to increase the concentration of agarose in our gel from 1.2% to 2% to ensure proper resolution. We ran the PCR of the normal and modified amplification conditions simultaneously. To verify the quality of the cDNA products and the impact of our alterations, we ran the cDNA products from our normal PCR on a 2% EtBr agarose gel and the cDNA products from our modified PCR on a 1.2% EtBr agarose gel. The amplification took an average of 24 cycles to yield enough ds cDNA. Both amplifications yielded high-quality cDNA but the cDNA from the modified conditions looked sharper as shown in figure 4c and d. After determining which conditions were ideal for the INS-1 cDNA amplification, we amplified all the ss cDNA from the previous part of the experiment and ran them on a 1.5% gel (fig. 4e and f). The gel shows replicates of the various cDNA samples. We combined these replicates and proceeded to the normalization. The average yield of ds cDNA for this second strand amplification was about 350ng/µl.

<table>
<thead>
<tr>
<th></th>
<th>Normal Conditions</th>
<th>Modified Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Denaturation</td>
<td>95 °C 1 min</td>
<td>95 °C 45 s</td>
</tr>
<tr>
<td>Denaturation</td>
<td>95 °C 15 s</td>
<td>95 °C 15 s</td>
</tr>
<tr>
<td>Annealing</td>
<td>66 °C 20 s</td>
<td>66 °C 20 s</td>
</tr>
<tr>
<td>Elongation</td>
<td>72 °C 3 min</td>
<td>72 °C 2 min</td>
</tr>
</tbody>
</table>

Table 2. Modified INS-1 cDNA Synthesis Amplification Conditions.
cDNA Normalization and Quantitative Analysis (qPCR)

For the normalization, we used ds cDNA from samples 1, 3 and 4. The first time we conducted this experiment, we used replicates of the cDNA from sample 1 and labelled them A, B, and C. Because of the sensitive nature of the DSN enzyme, we had to conduct an ethanol precipitation of the cDNA after cleaning and concentrating it using the Zymo™ DNA clean and concentrate kit. The manufacturers of the normalization kit included an ethanol precipitation protocol. We used a combination of this protocol and one from The Hla Lab at Harvard University\(^ {30} \) to get maximum yield of cDNA post purification. The manufacturer’s protocol recommends treating each cDNA sample with different concentrations of DSN. We treated each replicate with 1x, ½x and ¼x amounts

---

**Figure 4. Modifications to Amplification Conditions Yield High-Quality INS-1 cDNA.**

a) Results of an evaluative PCR under normal conditions* to determine the optimal number of cycles for cDNA synthesis. This was run on a 1.2% EtBr agarose gel. Typical electrophoresis of optimally cycled cDNA should show a distinct smear and some bright bands that represent abundant transcripts. Only three lanes show smears, but no bright bands are visible. b) The products from 3a were cycled three more times but the results were still disappointing. c) Another evaluative PCR run under normal conditions on a 2% EtBr agarose gel showed more promising results. Most lanes show a moderately strong smear and several bright bands. d) Evaluative PCR under modified conditions run on a 1.2% EtBr agarose gel. Modified conditions show brighter bands. The smaller gel to the right shows the results of cycling samples 1 and 4 two more times. e and f) 1.5%EtBr agarose gel showing replicates of cDNA synthesized under modified PCR conditions for 24 cycles.
of DSN. We also had control cDNA samples that were not treated with DSN. After the
treatment, we amplified the treated and untreated fragments and ran them on a 1.5% gel.
Results of this amplification are shown in figure 5a. Normalized cDNA appears as a
bright smear on agarose gels showing equalized transcript concentrations. After
comparing the results from the normalization amplification, we decided to quantify the
½x and ¼x normalized cDNA from sample A using a qPCR. This quantification showed
that the normalization was inefficient(fig.5b). For genes with high abundance transcripts,
the ΔCt value post-normalization should be ≥5.24 Most of these genes had ΔCt values
below 5 (Table 3).

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>½x DSN</th>
<th>¼x DSN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eno1</td>
<td>3.880</td>
<td>3.143</td>
</tr>
<tr>
<td>Ppia</td>
<td>3.973</td>
<td>4.713</td>
</tr>
<tr>
<td>Nkx6.1</td>
<td>1.063</td>
<td>-0.303</td>
</tr>
<tr>
<td>Insulin</td>
<td>0.673</td>
<td>-0.050</td>
</tr>
<tr>
<td>Nr4a2</td>
<td>-1.730</td>
<td>-2.323</td>
</tr>
</tbody>
</table>

Table 3. ΔCt Values Showing Failed Normalization.

To ensure that these ΔCt values were not tainted by the quality of our primers, we
analysed the melt curves (not shown) for each primer. We eliminated Nr4a2 primers from
future quantifications because we could not confirm its expression levels in β-cells. We
repeated this normalization with slight changes to the amplification conditions as shown
in table 4. For the repeated normalization, we used cDNA samples 1, 3 and 4 and treated
them with 1x, ¾x and ½x amounts of DSN. Visual comparison of treated samples vs
controls on gel showed samples 4 ¾x and ½x DSN to be most normalized. These samples
were used for the quantification (fig 5c). We also added Gapdh, beta-actin and Sdha
primers to the quantification. This second normalization yielded better results compared
to the first. Most of the ΔCt values were above 5 (Table 5). Some of the primers’ melt curves (not shown) including those of Nkx6.1, Insulin and Eno1 were inconclusive so we did not use these results to determine the normalization efficiency even though they are included in the analysis (fig.5d). Overall, we noticed a convergence between pre- and post-normalized cDNA showing equalized transcript concentrations.

<table>
<thead>
<tr>
<th></th>
<th>Normal Conditions</th>
<th>Modified Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Denaturation</td>
<td>95 °C 1 min</td>
<td>95 °C 45 s</td>
</tr>
<tr>
<td>Denaturation</td>
<td>95 °C 15 s</td>
<td>95 °C 15 s</td>
</tr>
<tr>
<td>Annealing</td>
<td>66 °C 20 s</td>
<td>65 °C 15 s</td>
</tr>
<tr>
<td>Elongation</td>
<td>72 °C 3 min</td>
<td>72 °C 3 min 30 s</td>
</tr>
</tbody>
</table>

Table 4. Modified INS-1 cDNA Normalization Amplification Conditions.

Figure 5. INS-1 cDNA Normalization Shows Effective Normalization for Some cDNA Samples. a) 1.5% EtBr Agarose gel of three normalized INS-1 cDNA samples labelled A, B, and C. Each sample is a replicate of the original cDNA. The gel lanes show each sample treated with 1x, ½x and ¼x amounts of DSN. It also shows the control cDNA of each sample that was not treated with DSN. An even bright smear signifies efficient normalization. Some of the products showed efficient normalization including sample A and C’s ½x, and ¼x DSN treated products. b) Quantification of the normalization from a using sample A’s ½x, and ¼x DSN treated products showed that the normalization was inefficient. c) 1.5% EtBr Agarose gel of normalized INS-1 cDNA samples labelled A, B, and C. Each sample was treated with 1x, ¾x, and ½x amounts of DSN and the control was untreated as shown in the gel. Sample C’s ⅔x, and ½x DSN treated products showed the most effective normalization and these samples were used for the qPCR analysis. d) Quantification of the normalization showed effective normalization.
INS-1 CRISPR-Cas9 sgRNA library Shows Promising Results

Using the SLALOM™ method we generated an INS-1 CRISPR-Cas9 sgRNA library. The results from this library show that the SLALOM™ method works on INS-1 β-cells. We noticed two bands when we ran the sgRNA library on a gel which could be because one of the ligation steps worked partially (fig.6). Before this library is applied to other biological processes, we will optimize the library generation protocol for INS-1 cDNA.

![Figure 6. INS-1 CRISPR-Cas9 sgRNA library.](image)

The expected size for the INS-1 sgRNA library is 167 bp. Although the second lane has two strong bands, the top band is where we expect the sgRNA library to appear at 167 bp.
DISCUSSION AND FUTURE DIRECTIONS

Hyperglycaemia and hyperlipidaemia due to muscle, adipose, and liver tissue insulin resistance are characteristic of T2D.\textsuperscript{3} This project has brought us one step closer to understanding the effects of these stressors and their mechanisms of action on $\beta$-cells. The INS-1 CRISPR sgRNA library that we generated can be applied to future work including creation of a forward genetic CRISPR-Cas9 knockout screen of the INS-1 $\beta$-cell genome. This screen will help us understand more about the genes involved in the pre-diabetic stress conditions. Because our library was synthesised using the SLALOM\textsuperscript{TM} method, it can easily be ligated into a lentiviral plasmid which also codes for the Cas 9 protein\textsuperscript{31} necessary for CRISPR knockdowns in the genome. After the INS-1 $\beta$-cells are transduced with the lentivirus, we will subject them to glucolipotoxic and other stressful conditions to generate a population of cells that are resistant to these conditions. After this, we will sequence the genome of these cells and identify genes directly involved in protecting the cells from these stressors.

Apart from being used to generate a screen of genes that are directly involved in T2D stress conditions, this library can be applied to other $\beta$-cell viability assays. One such assay is the investigation of the effects of metabolites on $\beta$-cell viability. Previous research has shown that phytochemicals including Fagonia Oliveri, cocoa flavanols and other flavonoid microbial metabolites enhance glucose-stimulated insulin secretion (GSIS), boost insulin levels, increase respiration rates, increase the amount of electron-transport chain (ETC) proteins, and alleviate metabolic syndrome.\textsuperscript{32,33,34} We can treat the post-transduced INS-1 $\beta$-cells with these compounds and subject them to apoptotic and other stress conditions. These methods will help us discover how these compounds make
β-cells resistant to metabolic stressors. This library can also be applied to other cell viability assays including cell proliferation assays, ATP measurement assays and apoptotic assays.

Diabetes research has also focused on studying the impact of certain growth and transcription factors like IGFs, Nkx6.1, Neuro-D and Pdx1 on β-cell growth, identity, transcription regulation, proliferation, dysfunction, insulin production and insulin secretion. These studies have shed light on the molecular mechanisms that contribute to β-cell dysfunction. This library can be used to discover unknown factors that may also contribute to the mechanism of β-cell dysfunction. These factors may either act up- or downstream of these transcription and growth factors.

In the future, we plan to regenerate this library to achieve a cleaner product. We will also use this library to carry out a forward genetic CRISPR-Cas9 screen as illustrated in figure 7. This library will help us discover more about the pathways of pre-diabetic metabolic stressors. It will also contribute to other diabetes studies including β-cell viability, transcription factor, and stem cell studies. These studies can then be applied to developing possible curative approaches to diabetes.
Figure 7: Future INS-1 CRISPR/Cas9 screen workflow. A stepwise illustration of the conceptual workflow of the INS-1 CRISPR screen. Adapted from Thomsen and Mikkelsen. 39
REFERENCES


14. CRISPR-Cas9, TALENs and ZFNs - the battle in gene editing. Ptglab.com.


23. Methods to check RNA integrity - US.


27. QIAGEN. What is the threshold cycle or Ct value? - QIAGEN. Qiagen.com.


APPENDICES

R-Script For qPCR Data

```r
library(tidyverse)
library(stringi)

# TO DO: Save first page of Quantification Amplification Results file as "AmplificationData.csv" in working directory (be sure to change file type from .xlxs to .csv)
quantificationData = read_csv("AmplificationData.csv")
# TO DO: change 97 to number of reactions + 1
quantificationData <- gather(quantificationData, "sample", "Flouresence", 2:97)

# TO DO: Save first page of Quantification Cq Results as "CtData.csv" in working directory (be sure to change file type from .xlsx to .csv)
CtData = read_csv("CtData.csv")

geneCtData = data.frame(Gene = character(), Control = double(), Normalized = double(), Delta = double())
geneCtData2 = data.frame(Gene = character(), Control = double(), Normalized = double(), Delta = double())

###  TO DO: Repeat for each gene/primer pair ###
# Extracts gene amplification data
# TO DO: Change letter to letter of row for each sample
geneData = filter(quantificationData, grepl("A", sample))
types = c(rep("Untreated Control", 40*3), rep("Normalized Sample DSN 1/2", 40*3), rep("NTC", 40*3), rep("Normalized Sample DSN 1/4", 40*3))
geneData$`Sample Type` = types

# Extracts gene Ct data and calculated Delta Ct value
# TO DO: Change letter to letter of row for each sample
geneCt = filter(CtData, grepl("A", Well))
geneCt$SampleType = c(rep("Control", 3), rep("Normalized", 3), rep("NTC", 3), rep("Normalized 2", 3))
CtControl = mean(filter(geneCt, SampleType == "Control")$Cq)
CtSample = mean(filter(geneCt, SampleType == "Normalized")$Cq)
CtSample2 = mean(filter(geneCt, SampleType == "Normalized 2")$Cq)
DeltaCt = CtSample - CtControl
DeltaCt2 = CtSample2 - CtControl

# TO DO: Add gene name

geneCtData <- geneCtData %>% add_row(Gene = "[Gene Name]", Control = CtControl, Normalized = CtSample, Delta = DeltaCt)
geneCtData2 <- geneCtData2 %>% add_row(Gene = "[Gene Name]", Control = CtControl, Normalized = CtSample2, Delta = DeltaCt2)
```
# Plot first Normalized Sample
genePlot <- ggplot(filter(geneData, types != "Normalized Sample DSN 1/4"), aes(x = Cycle, y = Fluoresence, line = sample, colour = 'Sample Type')) +
  geom_line() +
  theme_bw() +
  # TO DO: Add Gene name to graph title
ggtitle("[Gene Name] DSN 1/2") +
  geom_vline(xintercept = CtControl, linetype = "dashed", colour = "blue") +
  geom_vline(xintercept = CtSample, linetype = "dashed", colour = "red") +
  theme(plot.title = element_text(hjust = 0.5))

genePlot + annotate("segment", x = CtControl, xend = CtSample, y = 1350, yend = 1350) +
  annotate("text", x = CtControl - 3, y = 1375, label = bquote(Delta ~"Ct = " ~ .(DeltaCt)))
# TO DO: Add File name for graph
ggsave("FileName.png")

# Plot second Normalized Sample
genePlot <- ggplot(filter(geneData, types != "Normalized Sample DSN 1/2"), aes(x = Cycle, y = Fluoresence, line = sample, colour = 'Sample Type')) +
  geom_line() +
  theme_bw() +
  # TO DO: Add gene name to title
ggtitle("[Gene Name] DSN 1/4") +
  geom_vline(xintercept = CtControl, linetype = "dashed", colour = "blue") +
  geom_vline(xintercept = CtSample2, linetype = "dashed", colour = "red") +
  theme(plot.title = element_text(hjust = 0.5))

genePlot + annotate("segment", x = CtControl, xend = CtSample2, y = 1350, yend = 1350) +
  annotate("text", x = CtControl - 3, y = 1375, label = bquote(Delta ~"Ct = " ~ .(DeltaCt2)))
# TO DO: Add file name for graph
ggsave("FileName.png")

## TO DO: repeat previous section for each gene/primer pair ##

# Plot normalization efficiencies
# Ct analysis DSN 1/2
geneCtData <- geneCtData %>% gather("Sample", "Ct", 2:3)
ggplot(geneCtData, aes(x = Sample, y = Ct, colour = Gene, group = Gene)) +
  geom_point() +
  geom_line() +
  ggtitle("Normalization Eefficiency DSN 1/2") +
scale_y_reverse() +
theme_bw() +
theme(plot.title = element_text(hjust = 0.5)) +
scale_x_discrete(labels = c("Before Normalization", "After Normalization")) +
xlab(""
# TO DO: Add file name for graph
ggplot("FileName.png")

# Ct analysis DSN 1/4
geneCtData2 %>% gather("Sample", "Ct", 2:3)
ggplot(geneCtData, aes(x = Sample, y = Ct, colour = Gene, group = Gene)) +
geom_point() +
geom_line() +
ggtitle("Normalization Efficiency DSN 1/4") +
scale_y_reverse() +
theme_bw() +
theme(plot.title = element_text(hjust = 0.5)) +
scale_x_discrete(labels = c("Before Normalization", "After Normalization")) +
xlab(""
# TO DO: Add file name for graph
ggplot("FileName.png")

NEBioCalculator® Linear DNA: Mass to Moles of Ends Converter Equation

\[
\text{moles of dsDNA ends} = \frac{\text{mass of dsDNA} (g) \times 2}{(\text{length of dsDNA} (bp) \times 617.96 \left( \frac{g}{\text{motbp}} \right)) + 36.04 g/mol}
\]