UC Santa Cruz

UC Santa Cruz Electronic Theses and Dissertations

Title

A quantitative RT-PCR assay to monitor luciferase reporter mRNA levels

Permalink

https://escholarship.org/uc/item/10h1h513

Author

Thomas, Richard

Publication Date

2022

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

A quantitative RT-PCR assay to monitor luciferase reporter mRNA levels

A thesis submitted in partial satisfaction
of the requirements for the degree of

MASTER OF SCIENCE

in

MOLECULAR, CELL AND DEVELOPMENTAL BIOLOGY

by

Richard Thomas

September 2022

The thesis of Richard Thomas
is approved:

Professor Alan Zahler

Professor Jeremy Sanford

Professor Joshua Arribere

Peter Biehl
Vice Provost and Dean of Graduate Studies

Copyright © by

Richard Thomas

2022

Table of Contents

List of Figures and Tables

Chapter 2

Abstract

A quantitative RT-PCR assay to monitor luciferase reporter mRNA levels

Richard Thomas

Bioluminescent reporter assays are critical for functionally characterizing gene regulatory elements. Despite the wide-spread usage of reporter assays, the accurate interpretation often requires additional measurements beyond reporter enzyme activity. For example, when studying cis-acting RNA elements involved in post-transcripitional control it is critical to measure the steady state reporter mRNA levels. This can assist in drawing inferences about the underlying mechanism responsible for any differential reporter activity. My thesis work describes development of an RT-qPCR assay to monitor mRNA levels from a transiently transfected dual luciferase reporter system. After validating individual qPCR primer sets, this assay was applied to a comparative transcriptomics study, which previously identified orthologous mRNA isoforms with different polyribosome association profiles. It was shown that a single nucleotide variant in the GGCX and MELK orthologs of humans and chimpanzees was sufficient to modulate luciferase enzyme activity of reporters harboring single exons from these genes. The RT-qPCR results demonstrate that the reporter mRNA levels were not affected by the SNV, indicating that differential reporter activity is due to the translational control of orthologous exons.

Acknowledgements

First and foremost, I would like to thank Dr. Sanford for being kind enough to accept me into his laboratory without me having any prior research experience, especially at a time in 2020 when there was so much uncertainty. Not only has he shown me utmost kindness, but he's been an amazing teacher as well. Most people can't say that they enjoy working with their boss, but I sure can! If I had to give him a rating as an employer it would be a 10 out of 10. He's gone above and beyond to help me in any issue he possibly could and has been genuinely supportive about me taking the best path in life. I cannot ask for a better PI. I highly doubt that, "Richard Thomas," acquiring a masters degree would have ever happened if it wasn't for him.

Next, I would like to thank all my colleagues; Alexander, whose constructive criticism for my research presentation significantly helped me achieve the desired grade in a course I needed to graduate. Victor, who took significant amounts of time out of his very busy schedule to assist me in diagnosing problems with my research. Zach, not only for being a great friend, but for helping redesign primers for my qPCR experiment. If it wasn't for him, I would not have finished my project. Also, thank you to all the other faculty members here at UCSC who taught me important science in my courses, who were a part of my committee (Dr. Arribere and Dr. Zahler), and who gave me the best advice throughout the last couple of years (that includes you Grace and Dr. Ro).

Chapter 1: Introduction

1.1 Genetic Variety Stored In One Place

Individual eukaryotic organisms display a significant variety of cellular distinctions. Yet, the genetic information required to engender these distinct cellular phenotypes is often indistinguishable. To illustrate, rod and cone cells in the eyes relay messages through the bipolar cells allowing a person to visualize their surroundings. On the other hand, olfactory neuron receptors in the nasal cavity relay information to the olfactory bulb allowing a person to smell their next meal. Although each of these cell types have a distinct function, the information required to generate each of these unique cells is all derived from the same source. The nucleus contains the entirety of an organism's genome where it is present in the form of nucleic acid base pairs. These genomic sequences are tightly wrapped around histones forming nucleosomes, in a form known as chromatin. Genetic information (or a gene) is carefully unwrapped and transcribed when specific proteins are required by the cell. Although the nucleus in every cell contains identical genetic information, regulatory factors determine which genes must be utilized to produce the necessary proteins.

1.2 Proteome Expansion

In 1941, George Beadle and Edward Tatum proposed that one gene must produce only one protein and that it would take over 100,000 genes to create a human being (Nilsen et al., 2010). However, this isn't evident according to contemporary studies. Recent estimates suggest only about 20,000 genes are

responsible for generating all proteins required in the assembly of an organism, with the discrepancy explainable via alternative splicing and other regulatory pathways (Nilsen et al., 2010). There are four types of alternative splicing pathways, all of which expand protein variety. These include the cassette exon inclusion, 5' splice selection, 3' splice site selection, intron retention pathway and mutually exclusive exons (Nilsen et al., 2010). It is not known how many isoforms are created specifically for each cell type. But the rate of transcription and variables such as the structure of the transcript, the accessibility of the binding sites to the spliceosome, and splicing regulators could affect the ratio of isoforms created between cell types. It may also be possible for certain intronic sequences to promote splicing as well. Cis-elements found in the sequences of certain introns or exons may act as splicing enhancers or silencers among pre-mRNA transcripts. Additionally, the rate of transcription and, thus, the rate of alternative splicing may be influenced by dynamic chromatin modification (Nilsen et al., 2010). These variables would provide advantages or disadvantages for the differential expression of certain isoforms. All in all, these processes further contribute to the expansion of the proteome through a limited number of genes.

1.3 Post-Transcriptional Control of Gene Expression

The central dogma of biology describes the flow of genetic information from DNA to RNA (transcription), and then RNA to protein (translation). However, RNA is not immediately translated after transcription. Instead, it must

undergo processing in the form of splicing to remove intervening sequences known as introns. A messenger RNA transcript with introns in its sequence is known as a pre-mRNA. There are numerous ways in which a pre-mRNA transcript may be spliced to create a variety of differing mature messenger RNA transcripts (mature mRNA). A common pathway is alternative splicing which occurs during gene expression and involves eliminating the introns, meanwhile the exons are assembled together in several distinct combinations (Will and Lührmann 2011). These differing combinations of mRNA transcripts derived from the same corresponding pre-mRNA are called isoforms. For example, one isoform may contain exons 1, 2, and 4; meanwhile, the other consists of exons 1, 2, 3 and 4. Both of these mRNA transcripts were derived from the same pre-mRNA, but will be translated into different proteins due to their distinct sequences. This process of alternative splicing significantly contributes to variable genetic expression because it allows for a variety of mature mRNA transcripts to be translated from one gene.

1.4 Alternative Splicing Mechanism

The spliceosome is a manifold of biological components that assembles in situ to splice the non-coding introns and ligate exons together (Wilkinson et al., 2020). It consists of five snRNA's (small nuclear RNA) and numerous other proteins that, when combined, form snRNP's (Wilkinson et al., 2020). These five snRNAs are known as U1, U2, U4, U5, and U6 (named after their uridine content). Some introns contain base pair sequences at their 5' and 3' endpoints

bordering exons, which lead the spliceosome to bind the correct site. The first step in the spliceosome pathway consists of the U1 snRNA binding to the 5' splice site of the intron containing the GU (Guanine-Uridine) sequence, which forms the E complex (Wilkinson et al., 2020). Then, the U2 snRNA will bind the branchpoint adenosine nearest to the 3' end of the intron; forming the A complex (Wilkinson et al., 2020). Once the U1 and U2 snRNA are bound to the intron, the tri-snRNP complex composed of the U4, U5, and U6 is allowed to bind both the U1 and U2 snRNP; forming the B complex (Wilkinson et al., 2020). However, the spliceosome must release the U4 snRNA and U1 snRNA prior to the start of two step catalysis. This allows for the branch point adenosine and the guanine to bind together and completely break off the exon bound at the 5' splice site, forming the C complex with the lariat structure (Wilkinson et al., 2020). Next, the second exon located at the 3' splice site has its bond nearest to the AG sequence broken (Wilkinson et al., 2020). Both exons are subsequently attached to one another creating a mature mRNA transcript (or part of it).

1.5 Regulatory elements

Certain sequences on an exon may promote or inhibit the binding of the spliceosome at a specific site. Cis-acting elements are sequences adjacent to splice sites that help recruit trans-acting factors such as SR proteins and hnRNPs. SR proteins belong to the serine/arginine family and recognize both exonic splice enhancers (ESEs) and intronic splice enhancers (ISEs). These SR proteins are trans-acting factors that recognize positive cis-element sequences

(enhancers) and promote spliceosome assembly (Matlin et al., 2005). Similarly, heterogeneous nuclear ribonucleoproteins (hnRNPs) are trans-acting factors as well, yet they serve the opposite function (Matlin et al., 2005). These hnRNPs bind to exonic splicing silencers (ESSs) and intronic splicing silencers (ISSs), which bind on the pre-mRNA to prevent excision of an intron (Matlin et al., 2005). Together, both sets of enhancers and inhibitors interact with the transcript to guide the spliceosome assembly and composition of exons through alternative splicing.

1.6 RNA Transport

The transport of mature mRNA toward its destination is a process that is not entirely understood, but some general principles for this mechanism have been established. For example, it is widely accepted that mRNA transcripts created within the nucleus must traverse into the cytoplasm of the cell by exiting the nuclear pore complexes (NPCs) of the nuclear envelope (Nakielny et al., 1997). NPCs are composed of over 100 proteins forming pores in the nuclear membrane serving the purpose of transporting metabolites, ions, and small proteins. Some larger molecules or proteins may be transported through an active energy dependent mechanism, which is the pathway utilized for movement of many RNA species across the NPC (Nakielny et al., 1997). Although each type of RNA has a unique travel pathway, they seem to converge at one point or another for efficiency.

The first example of transport involves the *Chironomus tentans* salivary glands. These produce the Balbiani ring structure mRNA, which undergoes a step by step structural reorganization while approaching the NPC (Nakielny et al., 1997). It first associates with fibrous material on the NPC before its 5' end is oriented facing the pores (Nakielny et al., 1997). Meanwhile, the 3' end of the transcript gravitates toward the orifice of the pore (Nakielny et al., 1997). This relaxes the ring structure of the Balbiani transcript allowing it to move across the pore into the cytoplasm. This process may be different between cell types and has not been established as the norm for mRNA export. However, it exhibits interaction of the mRNA protein complexes with the NPC of the nucleus. The majority of pre-mRNA transcripts are not readily exported into the cytoplasm as they need to undergo splicing, 5' capping, and polyadenylation at the 3' region (Nakielny et al., 1997). Previous studies have shown that cis-elements containing certain intronic sequences may act as export inhibitors. It also seems that the 5' cap and 3' end as well as the addition of spliceosome proteins may promote the export of the transcript. There is evidence that the formation of the spliceosome complex and its altering of the trans-acting factors (SR proteins and hnRNAs) or splice sites, trigger movement across the nuclear envelope (Nakielny et al., 1997). This is what led to the idea of spliceosome retention hypothesis.

Although export inhibitors exist as pre-mRNA sequences, excision of these introns may not always be required to initiate transcript export. For

example, the overexpression of a pre-mRNA transcript with a specific intron in yeast may saturate the nucleus and has been shown to force out the pre-mRNAs into the cytoplasm (Nakielny et al., 1997). Conversely, the diffusion of certain pre-mRNAs may depend on the completion of splicing. To illustrate, it was shown by transfecting intronless DNA that B-globin mRNA is not exported into the cytoplasm unless splicing occurs (Nakielny et al., 1997). Yet, it is still unknown why splicing is required for exports of some mRNA, but not others.

Interestingly enough, when the B-globin gene has been paired with other intronless genes such as the herpes virus thymidine kinase and hepatitis B virus S transcripts; accumulation of those transcripts occurred within the nucleus without export occuring (Nakielny et al., 1997). Furthermore, in *X. laevis* oocyte nuclei, it's been shown that the 5' cap structure may also influence RNA transport. mRNA with m3G caps have been shown to export transcripts slower compared to those containing the m7G cap structure (Nakielny et al., 1997). Thus, even cap binding protein components may affect mRNA transport. Several other studies indicate that the stem loop structure on the 3' end and the poly(A) tail may influence mRNA transport; however, none of these may be specifically required for export (Nakielny et al., 1997).

Another variable shown to intervene in RNA transport are hnRNP proteins. There are 20 major proteins in this category which are usually bound to the nascent pre-mRNA. Two of these, hnRNP A1 and hnRNP K, continuously shuttle mRNA transcripts from the nucleus to the cytoplasm (Nakielny et al.,

1997). This is evident in the insect protein, Ct-hrp36, of the *Chironomus tentans*, which is a type of A1 hnRNP that is found in the cytoplasm bound to the mRNA transcript (Nakielny et al., 1997). In this example, the same hnRNP A1 protein must then return to the nucleus. Furthermore, there also exists hnRNPs found only in the nucleus that prevent mRNA export. These are known as the hnRNP U proteins. Any mRNA transcripts to be exported must not have these proteins attached to it. Lastly, the NPC proteins present at the nuclear membrane are heat sensitive and may contribute to mRNA transcript accumulation within the nucleus preventing mRNA transport (Nakielny et al., 1997).

1.7 Eukaryotic Translational Control

Translational control is a mechanism critical for the expression of protein. However, mRNA levels don't always correlate with protein levels due to this very same mechanism. The mRNA translation process is accomplished in three steps. The initiation step occurs when the met-tRNA, a GTP molecule and eIF2 (eukaryotic initiation factor 2) binds the 40S to form the 43S pre-initiation complex (Beyaert et al., 2009). The 43S then binds to the 5' end cap of the mRNA (for cap dependent translation) forming the 48S pre-initiation complex. The 48S complex begins scanning the mRNA in the 5' to 3' direction until a start codon (AUG) for methionine is read. This causes the 60S subunit to bind the 48S complex creating the 80S subunit before elongation begins. The elongation step requires tRNAs to provide the ribosome peptides to assemble according to the codons it reads. Elongation continues until a stop codon is read.

There are numerous initiation factors required for translation to occur in a 5' capped mRNA. The first is eIF2 (eukaryotic initiation factor 2), which binds to the 40S subunit simultaneously with the met-tRNA and a GTP molecule to form the 43S pre-initiation complex (Eliseev et al., 2018). This process is also aided by eIF1A, which oversees the binding of the tertiary complex (eIF2, GTP, and MET-tRNA) to the 40S complex (Eliseev et al., 2018). Once the 48S complex reads the start codon, the eIF2 and the GDP molecules are released before the eIF5B binds to the 48S complex to attract the 60S subunit. The chief purpose of eIF2 is to oversee the binding of the MET-tRNA to the 40S subunit. Once the 60S subunit is bound to the 40S subunit, together, they become the 80S subunit, which is ready for elongation. The eIF3, however, plays an important role every step of the way, especially because it is responsible for dissociating from the ribosome and preventing the mRNA from being translated entirely (Eliseev et al., 2018). Due to its capabilities, the eIF3 plays a very important role in translation efficiency.

1.8 RNA Decay

Proteins and mRNA transcripts alike have a set lifespan. However, the latter tend to accumulate defects and are not subject to repair. Numerous enzymes, quality control surveillance components, and cytoplasmic mRNA degradation pathways exist to turnover mRNA and harvest their components for future use.

Several exonucleases and endonucleases are responsible for RNA degradation. The 5' cap and a poly(A) tail is attached at the 3' end on nascent pre-mRNAs to prevent exonucleases from degrading useful RNA molecules (Perez-Ortin et al., 2013). Within the nucleus, mRNA associates with other proteins to form mRNPs before crossing the NPC into the cytoplasm. mRNPs are surveyed for defects by the nuclear exosome to degrade any defective mRNA before export into the cytoplasm. This process is called nuclear retention and the primary surveyor is the 3'-5' nuclear exosome containing Rrp6 (Perez-Ortin et al., 2013). First, the Trf4p polymerase scans the 3' poly(A) tail of mRNA transcripts to detect any errors. If any are detected, the transcript is retained and degraded by the exosome to prevent the export of an aberrant transcript (Perez-Ortin et al., 2013).

Furthermore, incorrectly spliced pre-mRNAs may also be retained and degraded within the nucleus. In fact, there are specific proteins dedicated to degradation dependent on orientation. For example, if the unspliced transcript is surveyed from the 3' to 5' end, the nuclear exosome is tasked with its degradation (Perez-Ortin et al., 2013). Conversely, if the transcript is detected from the 5' to 3' direction, a different enzyme known as the Rat1 is tasked with its degradation (Perez-Ortin et al., 2013). Still some erroneous pre-mRNAs may be transported successfully into the cytoplasm where they will undergo more quality control measures to eliminate harmful or unspliced mRNAs.

Mature mRNA degradation occurs in the cytoplasm through a variety of different pathways and begins with the degradation of the 3' end poly(A) tail. This occurs through deadenylation by any of the three complexes Pan2 and Pan3, the Ccr4-NOT complex, and PARN (Perez-Ortin et al., 2013). It is followed by the removal of the 5' cap by the decapping machinery, and then by mRNA degradation via the exoribonuclease Xrn1 (Perez-Ortin et al., 2013). In other instances, mRNA decay can occur in an endonucleolytic manner, in which the mRNA is cleaved internally before it is digested by exonucleases toward the ends. In the 5' to 3' direction, the mRNA fragment is degraded by Xrn1. In the 3' to 5' direction, the mRNA fragment is degraded by the cytoplasmic exosome, composed of 9 different subunits that can modulate catalytic activity and substrate specificity (Perez-Ortin et al., 2013).

One important reason why these aberrant mRNAs must be destroyed by the cell is so they won't be translated into abnormal proteins or waste energy and resources needed elsewhere. Nonsense-mediated decay (NMD) is a common pathway which terminates translation prematurely through the premature stop codons. NMD substrates include shifts in the open reading frames (ORF) or intronic sequences containing these premature stop codons. NMD is conserved across all eukaryotes and there are three proteins primarily responsible, which are Upf1, Upf2, and Upf3 (Perez-Ortin et al., 2013). After detection of aberrant mRNAs, these targets are decapped, deadenylated and decayed through exo- or endo-nucleolytic activity. In addition to their role in

surveillance, these pathways can also be utilized to control gene expression when needed. For example, in HeLa cells, the downregulation of Upf1 causes there to be a 5% increase in the correct mRNA transcription (Perez-Ortin et al., 2013). Two other mRNA degradation pathways include no-go decay and nonstop decay. Both of these begin when the ribosome is stalled during translation. No-go decays performs endonucleolytic cleavage during ribosome stalls before degrading the remaining mRNA with the Xrn1 and the peptides with the proteasome (Perez-Ortin et al., 2013). Similarly, non-stop decay occurs when the mRNA has no stop codons to cease translation, which eventually causes decay via ribosomal stalling, similar to no-go decay.

Although mRNA can be translated and degraded in the cytoplasm, it can also be stored to act as an mRNP. The function of these mRNPs has not been established, however, studies suggest that they may assist in regulating gene expression.

1.9 RT-PCR, qPCR, and RT-qPCR

RT-PCR refers to the synthesis of complementary DNA (cDNA) from a strand of RNA. First, a single stranded complementary DNA is synthesized from the sole piece of RNA by a transcriptase enzyme; followed by polymerase chain reaction, which generates abundant double stranded cDNA. Similarly, qPCR is used to detect the number of DNA copies of interest. However, it measures the quantification of DNA in real time using a fluorescent dye. Furthermore, qPCR can be probe-based or dye-based. The probe based qPCR assay requires an

additional sequence specific probe that covalently attaches a quencher molecule and releases a fluorophore when bound to the target sequence (Nolan et al., 2006). However, the dye based qPCR uses an intercalating dye that displays a strong fluorescent signal when bound to double stranded DNA. The difference between the dye and the probe is that the former is not sequence specific, which means that the dye based qPCR assay is susceptible to highlighting non-specific products such as primer-dimers. Both RT-PCR and qPCR steps above may be combined into one protocol to create the RT-qPCR method.

1.10 Dual Luciferase Assay

The luciferase assay consists of a bioluminescence reaction whereby luciferin is converted to oxyluciferin by the luciferase enzyme. This releases energy in the form of light. The light is then measured by a luminometer. However, in order to detect protein expression, the gene of interest must be cloned upstream of the luciferase gene and be introduced into the appropriate cells (such as through transient transfection). After allowing time for expression to take place (24 or 48 hours), the cells are lysed, which breaks free all the protein. The luciferin (from a Dual Luciferase Assay kit) is added to the samples and luminescence is measured on a luminometer. The results are determined by normalizing (or dividing) the expression of the firefly activity by the renilla activity for each sample creating expression ratios (McNabb et al., 2005). This assay assists in determining the amount of protein expressed, if any, from the

transiently transfected cells. determining the amount of protein expressed, if any, from the transiently transfected cells.

Figure 1. Schematic of reporter constructs.

Chapter 2:

2.1. Introduction

Dual luciferase assays are a standard tool for the studying mechanisms of gene regulation due to their luminescent indicators allowing us to measure protein expression. However, in order to determine whether the processes of transcription, mRNA stability, or translational control is the mechanism responsible for differential expression, it's imperative to measure the steady state mRNA levels.

Several years ago, our lab discovered an unexpected connection between alternative splicing and translational control of mRNA isoforms (Sterne-Weiler et al., 2013). The central hypothesis is that alternative splicing remodels the cis-regulatory landscape of functional elements in mRNA isoforms from the same gene. Recently, in an effort to discover examples of alternative splicing coupled translational control (AS-TC) and the regulatory elements responsible for differential translation, Julia Phillip and Jolene Draper used Subcellular Fractionation and High throughput Sequencing (Frac-Seq) to characterize isoform-specific polyribosome association in three closely related primate induced pluripotent stem cell lines (Philipp et al. Manuscript in preparation). This work revealed a series of orthologous mRNA isoforms that exhibit significant differences in polyribosome association, suggesting differential translation. To test this hypothesis reporter constructs were created using a firefly luciferase reporter fused in frame from either the human or

chimpanzee orthologous exons, which differed by just a single nucleotide polymorphism (Figure 1). In this chapter, I use a newly developed dual luciferase mRNA quantification assay to interpret the effect of single nucleotide polymorphisms on reporter activity.

2.2 Materials and Methods

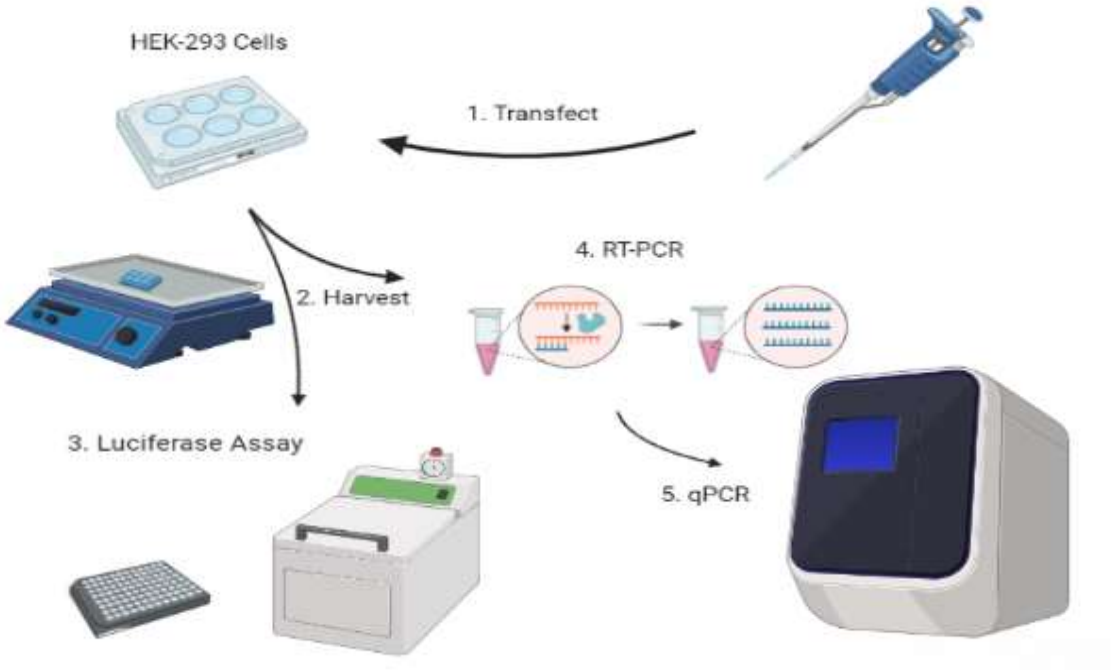

Figure 2. Experimental diagram/procedure.

Cell Culture

HEK-293 cells were utilized to perform all experiments. These were thawed from a liquid nitrogen storage tank and grown in DMEM media in a 37°C incubator until confluent. The cells were then gently washed with 1 mL of

Dulbecos Phosphate Buffered Saline (DPBS) and lysed with 1 mL of TrypLE™ Express dissociation reagent and left to incubate (at 37°C) for 5 minutes. Next, 4 mL of DPBS was used to dissociate the cells adhered on the plate along with the lysate, then centrifuged at 1200 RPM. The DPBS and remaining dissociation reagent was aspirated before the cells were resuspended and thoroughly mixed with 10 mL of DMEM media. In a 6-well plate approximately 200-250 uL of cell suspension was placed into each well containing 1.5 mL of DMEM. The cells were left in the incubator overnight.

Transfection

500 ng of each plasmid: GGCX Human, GGCX Chimp, MELK Human, MELK Chimp, SUMF2 Human and SUMF2 Chimp were co-transfected with 100 ng of the Renilla reporter empty 5' UTR. Lipofectamine L2000 in the ratio of 2 uL per microgram of DNA was used to transfect the cells before 48 hour incubation.

Cell Lysis

After 24 and 48 hours, the cells were carefully washed using 1 mL of DPBS per well. Then, 400 uL of 1x passive lysis buffer was placed into each well and the entire plate was shook for 15 minutes on an orbital shaker. Next, 300 uL of cell lysate were added to 900 uL of Trizol Tri reagent and stored at -20°C before RNA extraction.

Luciferase Assay

In a 96 well luminometer plate, 20 uL of cell lysate were placed in wells to examine three replicates of each sample. The Dual Luciferase® Reporter Assay

System kit from Promega was utilized to determine the expression of the firefly reporter within each sample. A total of 50 uL of Luciferase Assay Reagent II were added to each replicate sample before being placed in the luminometer for firefly luciferase activity measurement. Immediately after, 50 uL of a pre-made mix of Stop and Glo buffer (50 uL per reaction) and Stop and Glo Substrate (1 uL per reaction) were added to each well before Renilla activity was measured in the Victor3000 luminometer. The ratio of Firefly luciferase divided by Renilla luciferase activity was calculated for all samples.

RNA Purification

The Direct-zol RNA Miniprep Kit was utilized to purify RNA from samples frozen in Trizol Tri reagent. Samples were thawed and vortexed before 500 uL of Trizol/Cell lysate were added to 500 uL of 100% ethanol (1:1 ratio). Each of the 1 mL mixtures were added to Zymospin™ IICR Columns and Collection Tubes before being centrifuged at 15,000 RPM for 1 minute. The flow-through was discarded and a DNAse I treatment was carried out on each sample. This involved adding 400 uL of RNA Wash Buffer to each column and centrifuging for 1 minute at 15,000 RPM. Then, a mix of DNAse I (5 uL per sample) and DNA Digestion Buffer (75 uL per sample) are added to the column of each sample and left to incubate at room temperature for fifteen minutes. The RNA purification process is then continued by adding 400 uL of Direct-Zol PreWash and centrifuging twice at max speed. The liquid waste is disposed. 700 uL of Direct-Zol Wash are then added to the column and centrifuged for 2 minutes. A

Kim-wipe is used to dry the column of any residual wash solution before they are transferred into clean microcentrifuge tubes. 25 uL of water is added to the column and centrifuged before elution. RNA is then quantified and stored at -70 °C.

RT-qPCR

Reverse transcription is performed independently using the following protocol for a 1x reaction: 4.2 uL of water, 2 uL of 10x RT buffer, 0.8 uL of 25x dNTP's, 2 uL 10x random primers, and 1.0 multiscribe RT enzyme. This is added to 10 uL of water/template mix for each sample. After the reaction is carried out, cDNA is diluted 1:100. Where 4 uL of cDNA sample is used to perform mRNA analysis. The following sybr dye reagent protocol was used for a 1x reaction: 10 uL of 2x Luna Sybr dye, 0.5 uL of 10 mM Forward primer, 0.5 uL of 10 mM Reverse primer, and 4 uL of water. The 4 uL of template was then added for each sample, making for a total volume of 20 uL per reaction. The "Fast" cycling protocol on the QuantStudio 6 machine was used and the qPCR primer sequences used are listed below.

Primer Sequences

Renilla Luciferase Forward	TGG CTT CCA AGG TGT ACG AC
Renilla Luciferase Reverse	GTT CTC CGC ATG TTT CTC GC
Firefly Luciferase Forward	AAC ACC CCA ACA TCT TCG AC
Firefly Luciferase Reverse	TCG CGG TTG TTA CTT GA CTG

Table 1. Primer sequences.
Firefly (pMIR) and Renilla primer sequences.

2.3 Results

Optimization of Dual Luciferase mRNA RT-qPCR primers

To complement our dual luciferase assay system, we developed primers to amplify specific regions of firefly and renilla luciferase mRNAs (table 1). I characterized the performance of the primers by optimizing their annealing temperatures and cycling parameters. I calculated the amplification efficiency using 5 fold serial dilutions diluted 1:10 containing Firefly and Renilla luciferase plasmid.

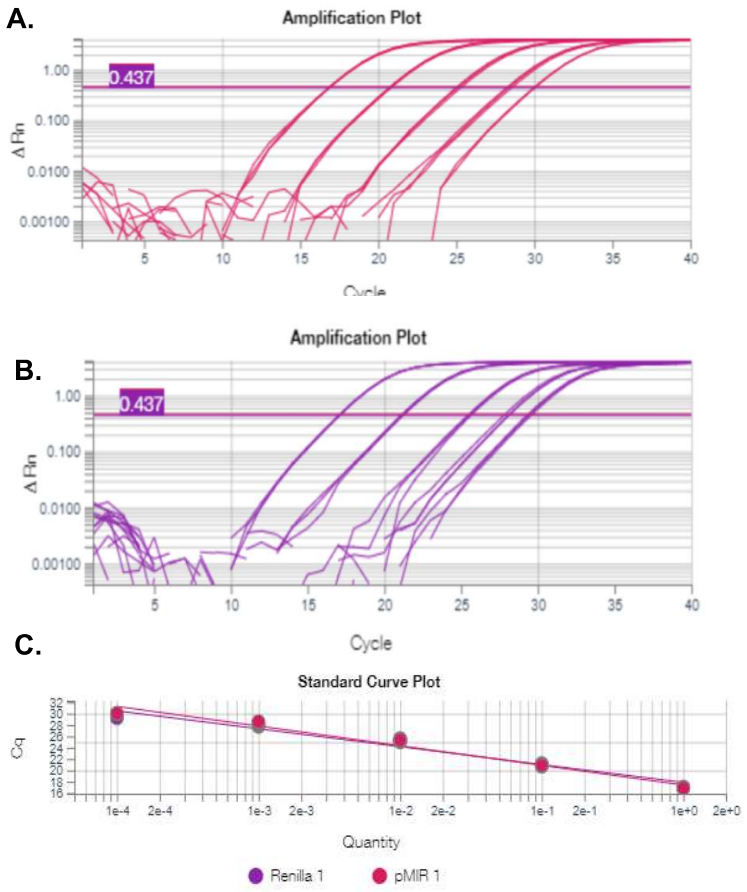

Figure 3. Standard curve data. (A) Firefly (pMIR) amplification plot. **(B)** Renilla amplification plot. **(C)** Standard Curves for Firefly and Renilla.

Using this standard curve we determined the amplification efficiency for each primer pair. Figure 3A shows the amplification plot for the Firefly (pMIR) plasmid. Figure 3B shows the amplification plot for the Renilla plasmid and Figure 3C shows the standard curve for both plasmids. Under standard conditions, Firefly and Renilla standard curves show an amplification plot with increasing Cq values and an efficiency of 107% for Firefly and 94% for Renilla.

Additionally, both Firefly and Renilla melt curves show a single peak at a high temperature indicating that there is no contamination with either experiment and that there is amplification of only one product.

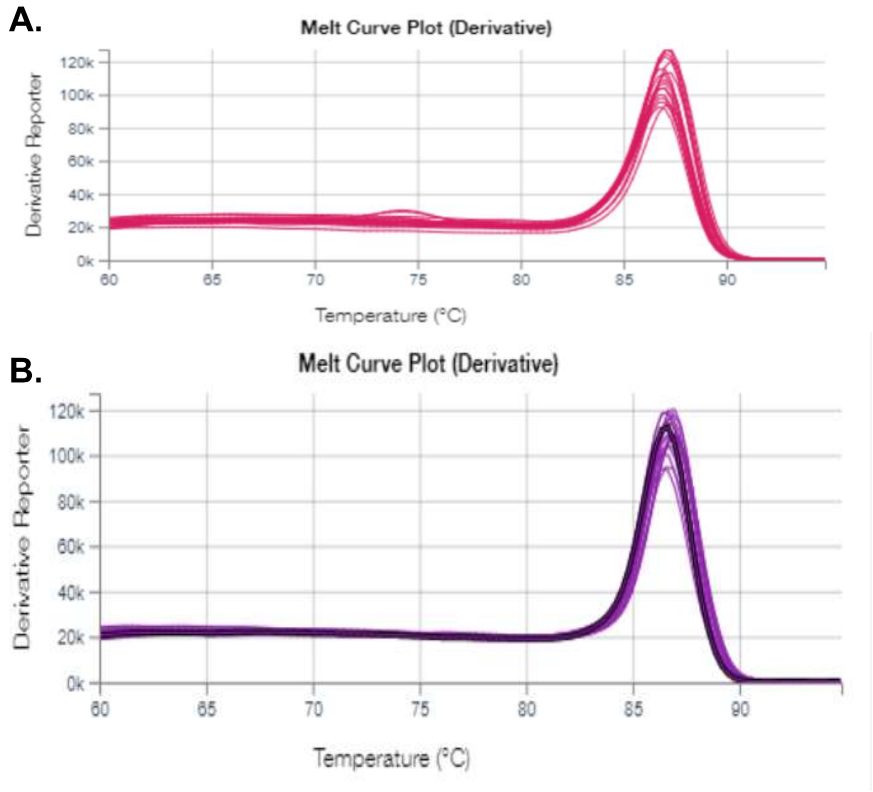

Figure 4. Melt curve data. (A) Firefly (pMIR) melt curve. **(B)** Renilla melt curve.

Measuring dual luciferase mRNA levels of orthologous reporters

In an effort to create transcriptomics tools that would predict the translation efficiencies based on the unique sequence landscapes created as a result of alternative splicing; two previous graduate students in my lab (Julia Philipp and

Jolene Draper) discovered orthologous genomic sequences that differed by only a single nucleotide. These variations in sequences are known as Single Nucleotide Polymorphisms (SNPs). The SNPs (single nucleotide polymorphisms) between GGCX and MELK orthologs show a distinction in protein expression. The following graphs (Figure 5) show that GGCX Human is expressed higher than Chimp, meanwhile, the human has a slightly lower mRNA level than the chimp. MELK expression is higher in the chimp than the human ortholog, yet the human has lower mRNA levels. The expression with the SUMF2 orthologs is negligible when compared. The SUMF2 Chimp mRNA levels seem to be higher than the human. Reverse transcription and qPCR were performed using the purified RNA. The following graph shows the steady state mRNA levels to be stable between all three orthologous exon pairs. SNPs in GGCX and MELK affect reporter activity but not mRNA levels. This analysis of reporter mRNA levels revealed allele-specific differences in translation.

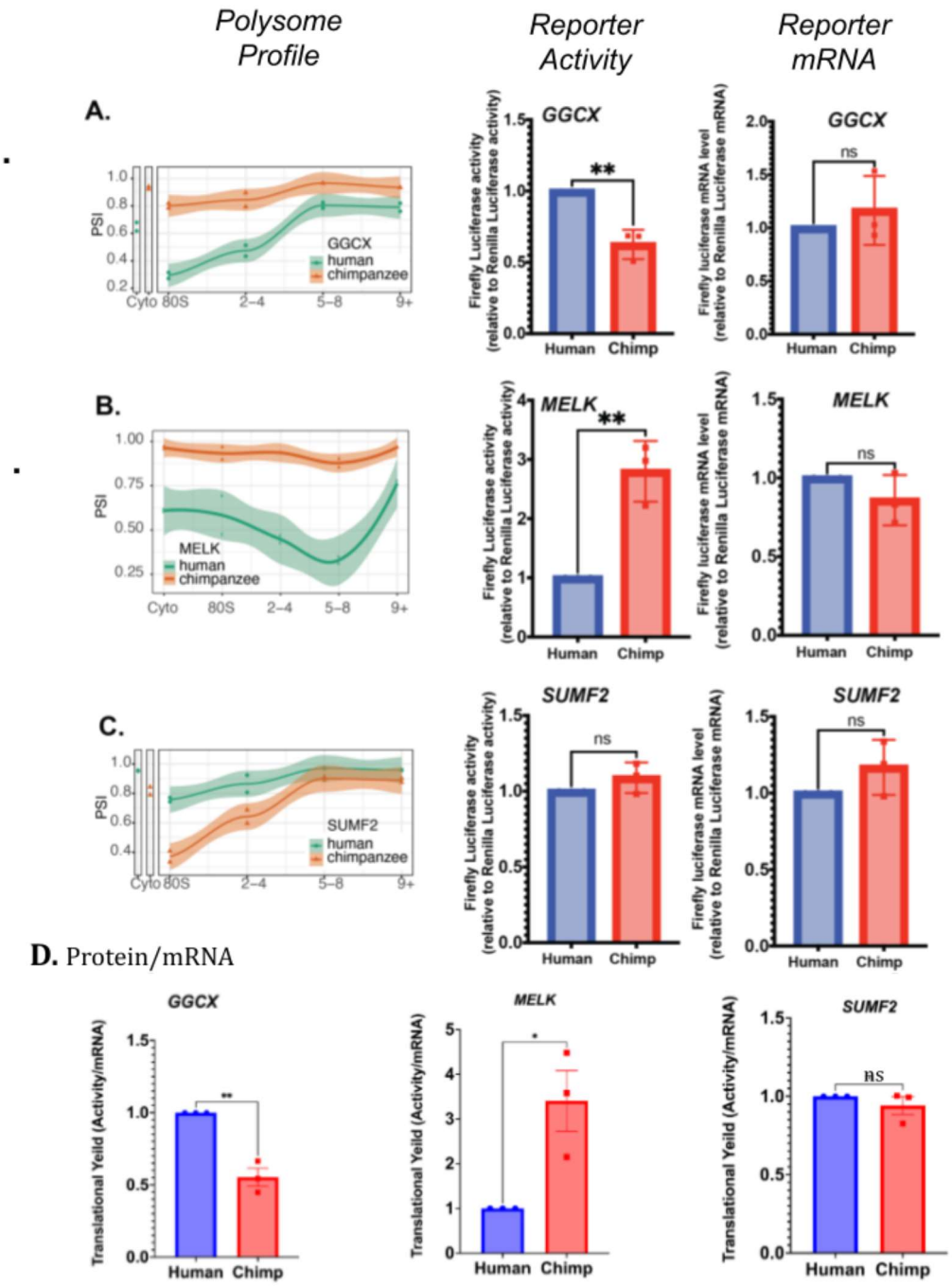

Figure 5. Reporter data. (A) Polysome profile of GGCX Human and Chimp orthologs, their protein expression, and qPCR mRNA levels. **(B)** Polysome profile of MELK Human and Chimp orthologs, their protein expression, and qPCR mRNA levels**. (C)** Polysome profile of SUMF2 Human and Chimp orthologs, their protein expression, and qPCR mRNA levels. **(D)** Protein levels/mRNA levels.

2.4 Discussion

We were able to develop a dual luciferase assay and quantify the relative amounts of protein expression between multiple orthologous exons from a human and chimpanzee. Additionally, we developed primers for a qPCR assay to ascertain that mRNA levels are steady. Our results indicate that protein expression for the GGCX and MELK orthologs differ. After using the RNA from these transfections to perform qPCR, it was determined that the mRNA levels were stable. This means that the relative differences in protein expression are likely due to translational control rather than mRNA levels.

I think it's important to take note of the sedimentation profiles of the reporters because they exhibit a pattern. For example, the graph illustrating the percent spliced in of the GGCX human exon increases at a lower point in the gradient where the polyribosomes are located. My prediction is that there must be a factor causing this increase in translation activity compared to the chimp ortholog. This factor could be a cis-element present in the human sequence that recruits a protein. This may be the reason why there are higher levels of luciferase activity for GGCX human than chimp. A similar pattern is seen with the MELK orthologs as well. Notice how the percent spliced in of the MELK human reporter decreases towards the right (bottom of sucrose gradient) where the polyribosomes reside. Meanwhile, the luciferase activity is significantly lower than that of the chimp ortholog. In my opinion, this could

mean that there is a factor preventing efficient translation. My interpretation is that these factors are the culprit for the difference in reporter-enzyme activity. During further experiments it was also determined that a previous laboratory experiment was flawed due to the former student utilizing a common housekeeping gene to normalize the Renilla luciferase data. Moreover, it was determined that the former student had mislabelled a transiently transfected plasmid, which led to the production of misleading data.

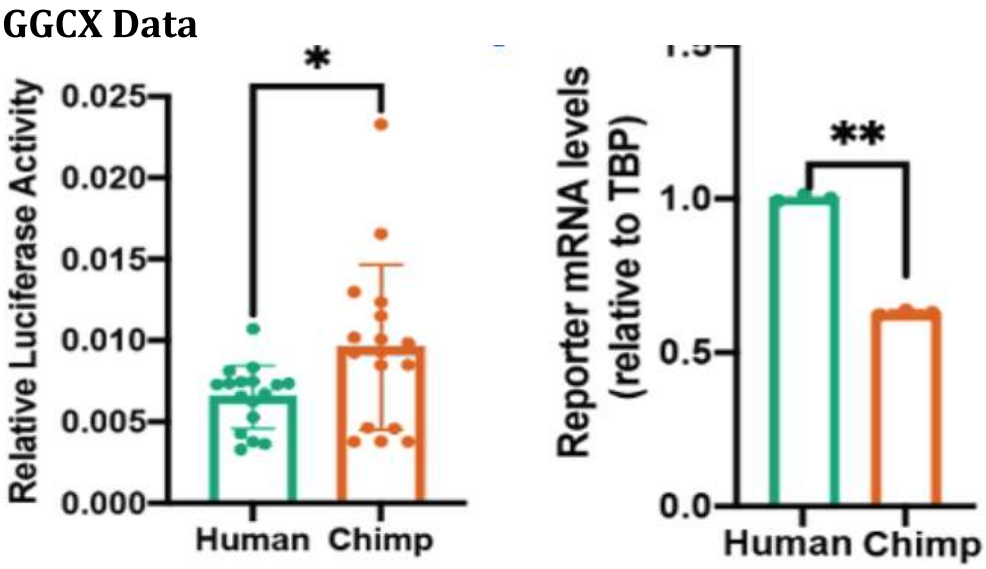

Figure 6. Flawed data. Previous students' flawed GGCX data showed human expression to be lower than chimp.

After performing a restriction digest and sequencing the plasmid it was confirmed that an incorrect plasmid was used. This discovery is important because the flawed results show that chimpanzee expression is higher than

human expression, yet, the former has relatively lower mRNA levels compared to the latter. This initially indicated that translational control was responsible for the higher rate of translation of the chimpanzee ortholog. However, after performing the experiment with the proper plasmid we found conflicting results. Instead, we found that the GGCX human ortholog was expressed significantly higher than that of the chimpanzee; meanwhile, the mRNA levels were relatively lower for the former.

The former student was unable to design primers that would properly anneal to the renilla plasmid for proper qPCR analysis, which is how this current project came to be. We spent several months optimizing and troubleshooting the primers to proper anneal to our plasmid of interest. We optimized cycling parameters, extension times, template concentrations, and used uracil deglycosylase to rid the samples of any amplicons containing deoxy uracil from previous reactions to acquire standard curves with roughly 100% efficiency. Lastly, throughout these experiments, the plasmids were continuously sequenced to confirm their authenticity and prevent production of any flawed data.

After finding that a single nucleotide variant causes a significant difference in reporter-enzyme activity, some future experiments to determine how this process occurs should be performed. It would be interesting to know whether overexpression of certain RBPs (RNA binding proteins) would affect the translation efficiencies of these reporters. These experiments could be done

by transiently co-transfecting plasmids coding for RBPs at different concentrations before measuring luminescence, mRNA levels and performing a western blot to verify protein expression of the RBPs.

2.5 References

Eliseev, B., Yeramala, L., Leitner, A., Karuppasamy, M., Raimondeau, E., Huard, K., Alkalaeva, E., Aebersold, R., & Schaffitzel, C. (2018). Structure of a human cap-dependent 48S translation pre-initiation complex. Nucleic acids research, 46(5), 2678–2689.

Matlin, A. J., Clark, F., & Smith, C. W. (2005). Understanding alternative splicing: towards a cellular code. Nature reviews. Molecular cell biology, 6(5), 386–398.

McNabb, D. S., Reed, R., & Marciniak, R. A. (2005). Dual luciferase assay system for rapid assessment of gene expression in Saccharomyces cerevisiae. Eukaryotic cell, 4(9), 1539–1549.

Nakielny, S., Fischer, U., Michael, W. M., & Dreyfuss, G. (1997). RNA transport. *Annual review of neuroscience*, *20*, 269–301.

Nilsen, T. W., & Graveley, B. R. (2010). Expansion of the eukaryotic proteome by alternative splicing. Nature, 463(7280), 457–463.

Nolan, T., Hands, R. E., & Bustin, S. A. (2006). Quantification of mRNA using real-time RT-PCR. *Nature protocols*, *1*(3), 1559–1582.

Pérez-Ortín, J. E., Alepuz, P., Chávez, S., & Choder, M. (2013). Eukaryotic mRNA decay: methodologies, pathways, and links to other stages of gene expression. Journal of molecular biology, 425(20), 3750–3775.

Sterne-Weiler, T., Martinez-Nunez, R. T., Howard, J. M., Cvitovik, I., Katzman, S., Tariq, M. A., Pourmand, N., & Sanford, J. R. (2013). Frac-seq reveals isoform-specific recruitment to polyribosomes. *Genome research*, *23*(10), 1615–1623.

Van Der Kelen, K., Beyaert, R., Inzé, D., & De Veylder, L. (2009). Translational control of eukaryotic gene expression. Critical reviews in biochemistry and molecular biology, 44(4), 143–168.

Wilkinson, M. E., Charenton, C., & Nagai, K. (2020). RNA Splicing by the Spliceosome. Annual review of biochemistry, 89, 359–388.

Will CL, Lührmann R. Spliceosome structure and function. Cold Spring Harb Perspect Biol. 2011 Jul 1;3(7):a003707.