Methodological Considerations for the Ribosomal Profiling of Human Post-Mortem Brain Tissue

Amanda Brown

Department of Human Genetics

Faculty of Medicine and Health Sciences

McGill University, Montreal

August 2022

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

Table of Contents

Abstract

Ribosome profiling (also known as riboseq) is a sequencing technique that allows for analysis of protein synthesis by capturing and profiling ribosome protected fragments (also known as RPFs or ribosome footprints): the ~30nt fragments of mRNA that are contained under ribosomes during translation. While first pioneered to observe changes in translational output in yeast, the applications of the procedure have diversified, and expanded to study species spanning from insects to humans. However, a ribosome profiling protocol for human post-mortem brain tissue has not yet been established in the literature, despite the fact that doing so could enable valuable research into differentially translated genes (DTGs) in psychopathologies with an epigenetic basis such as Major Depressive Disorder (MDD), post-traumatic stress disorder (PTSD), bipolar disorder, Alzheimer's disease, and Schizophrenia, and in other neurological pathologies such as Huntington's Disease and neuropathic pain. The absence of an extant protocol, and research in this tissue, is likely due to the fragility of human post-mortem brain tissue and potential degradation incurred from the inevitable post-mortem interval (PMI). This study endeavored to create a novel methodology for ribosome profiling in human post-mortem brain tissue by making methodological alterations to the original protocol in tissue lysis, ribosome footprinting, monosome detachment and RPF purification, and cDNA library construction. While a replicable protocol was not established in the course of this study, there were some promising findings, particularly with respect to the use of S7/T1 RNAse cocktail for ribosome footprinting, and the use of ligation free and random nucleotide adapter cDNA

library construction methods as possible solutions to conventional ribosome profiling methodologies.

Résumé

La séquençage des ribosomes (Riboseq) permet l'analyse de la synthèse des proteins. Riboseq isole et capture les courts (~30nt) segments d'ARNm qui sont piégés sous le ribosome pendant la traduction (appelé empreintes de ribosomes ou ERS). Il a d'abord été construit pour observer les changements dans la production de traduction chez la levure, mais les applications de la procédure se sont diversifiées et élargies pour étudier des espèces allant des insectes aux humains. cependant, il n'y a pas de procédure riboseq pour le tissu cérébral humain post-mortem dans la littérature, malgré le fait que cela pourrait permettre des recherches précieuses dans les gènes à traduction différentielle en les psychopathologies à base épigénétique comme dépressif majeur (DM), trouble de stress post-traumatique, trouble bipolaire, La maladie d'Alzheimer, la schizophrénie, ainsi que des pathologies neurologiques comme la maladie de Huntington et les douleurs neuropathiques. L'absence de protocole existant et de recherche dans ce tissu est probablement due à la fragilité de le tissu cérébral humain post-mortem et la dégradation potentielle résultant de l'inévitable intervalle post-mortem. Cette étude a tenté de créer une nouvelle méthodologie pour le riboseq dans le tissu cérébral humain post-mortem en testant des modifications méthodologiques au protocole original dans lyse tissulaire, empreintes de ribosomes, détachement des monosomes et purification des ERS, et conversion d'ADNc. Bien qu'un protocole reproductible n'ait pas été établi au cours de cette étude, il y a eu des résultats prometteurs, notamment en ce qui concerne l'utilisation de S7/T1 ARNase pour les empreintes de ribosomes, et l'utilisation de méthodes de conversion d'ADNc d'adaptateur de nucléotide sans ligature et aléatoire comme solutions possibles aux méthodologies de profilage de ribosome conventionnelles.

List of figures

List of Tables

Acknowledgments

Firstly, I would like to express my sincere gratitude to my supervisor Dr. Gustavo Turecki for granting me the opportunity to work on this project, and for all of the direction, support, supervision, and mentorship that was crucial to this process. I greatly appreciate every moment he both supported and challenged me, and have learned a great deal about being innovative, proactive, and diligent from his guidance.

As well, I would like to extend this gratitude to my committee members, Dr. Naguib Mechawar and Dr. Ziv Gan-Or, and my professor Dr. Roberta Palmer for all of their helpful feedback, advice, references, and mentorship during my graduate studies as well. Thank you.

Further, I thank the entire McGill Group for Suicide Studies family for everything you all have given me during my time here, especially Dr. Corina Nagy and Dr. Laura Fiori for the assistance with tackling the laboratory procedures together, troubleshooting, guidance, and the numerous hours editing and reviewing my written work. Additionally, I have to give immense credit to Jean-François Théroux for providing the bioinformatics analyses for this study, and Dr. Gary Gang Chen, Dr. M.A Davoli, Dr. Volodymyr Yerko, Dr. Jennie Yang, Elena Leonova-Erko, and Daniel Almeida for providing assistance in the lab. And, of course, thank you to all my MGSS peers for being there for emotional support, shenanigans, and checking in on me when I was ill. It means the world to me.

I would also like to thank Sonali Uttam, Dr. Arkady Khoutorsky, and bioinformatician Dr. Ines da Silva Amorim for allowing us to use their data to test our bioinformatics pipeline.

As well I extend my gratitude to my sources of funding throughout the project which include Fonds de Recherche du Québec, Healthy Brains Healthy Lives, The McGill University Faculty of Medicine, and The Douglas Mental Health University Institute.

Last but not at all least, I would like to thank my family: my mother Darlene, my father Todd, my brother Ryan, my extended family, and chosen undergrad family Devante, Maheshi, Salam, Yeonsu, and Nicole for all of the unconditional love, emotional support (and in some cases physical support when I was ill), and ears for all the times I rambled about my project. You're the best.

Contribution of the Authors

All Chapters and sections were written by the author in full, with pre-submission editing assistance from Dr. Corina Nagy and Dr. Laura Fiori, and from thesis reader Dr. Roberta Palmour following initial thesis submission.

Preface

This thesis is traditionally formatted. This study was performed under the supervision of Dr. Gustavo Turecki and addresses the major challenges and trials of attempting to create a replicable methodology for ribosome profiling in human-post mortem brain tissue.

Chapter 1 is an introduction and literature review of the history, methodology, and applications of ribosome profiling. Chapters 2-5 present the methods, results, discussion, and conclusion and future directions of the trials conducted in pursuit of creating this methodology.

Chapter 1. Introduction and review of literature

1.1 Ribosome profiling- background

For almost 20 years, RNA sequencing technology has been available to characterize the transcriptome, and has been most commonly used to identify differentially expressed genes (DEGs) (Weber, 2015). While much can be inferred from DEG data, one crucial phenomenon it does not perfectly reflect is protein synthesis. This is due to many complex factors such as post-transcriptional regulation, noncanonical translation initiation, nonsense read-through, and variation in how much protein is produced by a single mRNA before it is degraded. While there are methodologies that examine proteins directly, these can be labour intensive, low-throughput, and less precise. However, a more high-throughput methodology that permits analysis of protein synthesis, ribosome profiling, was first described in 2009, allowing for a new method of analyzing translation (Ingolia, Ghaemmaghami, Newman, & Weissman, 2009).

Ribosome profiling (or Riboseq) is a specialized mRNA sequencing methodology developed by the Ingolia lab, that broadly functions by sequencing only the gene regions that are being actively translated (also known as the translatome). This is accomplished by isolating and then sequencing the approximately 30 nucleotide (nt) long fragments of mRNA that are contained under the ribosome during translation (also known as ribosome footprints, ribosome protected fragments, or RPFs).

Numerous applications of Riboseq have been developed, for example, Riboseq has been used for whole translatome mapping, identifying whether suspected long non-coding RNAs (lncRNAs) are truly non-coding and untranslated, identifying differentially translated genes

(DTGs) related to various pathologies, and examining how translation is either altered in, or regulated by, factors such as life stage, circadian rhythms, and various treatment conditions (Chew et al., 2013; Ingolia et al., 2009; Jang, Lahens, Hogenesch, & Sehgal, 2015; Jensen et al., 2014; Uttam et al., 2018).

1.2 The impact of ribosome profiling on our understanding of molecular biological processes

The original 2009 riboseq yeast paper was able to demonstrate an up to 100-fold difference in translation efficiency indicating an important difference between overall transcript levels, and those that specifically show evidence of becoming proteins, which varies from gene to gene(Ingolia, Ghaemmaghami, Newman, & Weissman, 2009). This level of variance reflects the need to have a methodology distinct from mRNA sequencing for quantifying protein products. Additionally, they found evidence of changes in translation efficiency between starved (amino acid deprived) versus nutrient rich conditions, with 291 genes having a strong change in translational efficiency, with 111 of these genes being down-regulated (Ingolia, Ghaemmaghami, Newman, & Weissman, 2009). This study not only demonstrated that there can be substantial differences between transcription and translation rates for specific genes but also how environmental conditions can have specific effects on these rates. This study emphasized the differences between transcription and translation rates, indicating that transcription alone is not an adequate proxy for levels of translated products.

Riboseq has also been used to assess the translational potential of RNA species that are not traditionally considered to code for proteins. Chew et al. found that after applying riboseq to zebrafish embryos, several suspected Long-noncoding RNAs were found to be translated, although they speculated that they may generate non-functional proteins (Borsani et al., 1991; Chew et al., 2013). These results represent not only an important application of the methodology, but also suggest the possibility that many supposed lncRNAs have been misclassified- a finding supported by other studies (Ji, Song, Regev, & Struhl, 2015; Zeng, Fukunaga, & Hamada, 2018). This finding has since been reinforced in a bioinformatic study of human Riboseq datasets that set the translated "lncRNA" rate at 40% . Another study that used mouse and human Riboseq datasets from the NCBI GEO data repository to find lncRNA–ribosomal associations found that 39.17% and 48.16% of "lncRNAs" interact with ribosomes in the human and mouse datasets, respectively (Zeng et al., 2018). This use of riboseq greatly expands our understanding of the complexity of lncRNA function.

The complexity of translational regulation was also demonstrated by Dunn et al. (2013), who demonstrated that eukaryotic ribosomes can read-through stop codons in a regulated manner. After performing Riboseq on *Drosophila melanogaste*r, it was found that a significant number of translated footprints exhibited elongated peptides (Dunn, Foo, Belletier, Gavis, & Weissman, 2013). The *Drosophila* findings of stop-codon read-through and elongated translation were observed in yeast and humans as well, indicating that read-through is present in eukaryotes, where it appears to be a mechanism of translational regulation of gene expression (Dunn et al., 2013). It is studies like this one that help broaden our understanding of the mechanisms of translation regulation.

The methodology was further used on several studies involving the parasites *Trypanosoma brucei* and its sister species *T.a cruzi,* where up to 100 fold differences in translation efficiency were found in two different life cycle forms (Jensen et al., 2014; Smircich et al., 2015; Vasquez, Hon, Vanselow, Schlosser, & Siegel, 2014). These studies supplied unique information about the likely role of parasitic form in regulating translation, which would not have been obtained simply using transcriptomics. Even greater complexity in eukaryotic translation was observed in mouse embryos (Ingolia, Lareau, & Weissman, 2011). Riboseq on the mouse embryo stem cells revealed thousands of instances of amino-terminal extensions and truncations, upstream open reading frames at AUG and non-AUG codons, and short polycistronic ribosome-associated coding RNAs (sprcRNAs) which code for peptides.

In a study on neuropathic pain in mice, mRNA and riboseq were used to detect differences between transcriptomic and translatomic data in Dorsal Root Ganglion (DRG) and spinal cord tissue(Uttam et al., 2018). They found differences in both up- and down-regulation when comparing the riboseq to transcriptomic data, notably in genes associated with enzymes, ion channels, and transcription regulation. (Uttam et al., 2018). Considering that neuropathic pain remains a somewhat elusive condition with varying success with current pain management therapies, and that inhibition of select mRNA translation (but not expression) has shown some promising evidence in pain alleviation in preclinical studies, studies such as this one represent early work in potentially developing translatome based treatments that could give relief to the millions of people who suffer from pathological pain states, globally(Asante, Wallace, & Dickenson, 2010; Geranton et al., 2009; Obara et al., 2011).

Riboseq has also shown an association between reduced eIF4E phosphorylation and selective translation that is associated with depressive-like behaviours in male mice including immobility, feeding latency, and reduced open field exploration (Amorim et al., 2018). This allowed researchers to distinguish actively translating ribosomes from positions of inactive ribosomes. Other discoveries include: the identification of conserved translation sites between mouse and human samples, the identification of 5'UTRs and pseudogenes that are actually being translated, dual decoding sites on the human genome, how Shine-Dalgarno motifs and a lack of elongation factor (EFP) cause translational pausing in bacteria, how translation occurs in a circadian fashion for many human genes, and how translation repression of specific genes in the hippocampus are associated with memory formation, among other applications (Amorim et al., 2018; Calviello et al., 2016; Cho et al., 2015; Fields et al., 2015; Jang et al., 2015; Ji et al., 2015; Mohammad, Woolstenhulme, Green, & Buskirk, 2016; Woolstenhulme, Guydosh, Green, & Buskirk, 2015). These studies demonstrate the potential for ribosome profiling to assist in understanding issues in molecular biology ranging from accurate categorization of genes and RNAs, translation regulation, human physical and psychological pathologies, and aspects of brain function (such as memory formation).

The studies described above demonstrate how Riboseq has been used and adapted to describe the diverse roles of translational regulation on protein synthesis. Which alterations have been made to the Ingolia methodology—and which instances they apply best to—will be explored in Section 3 of this review. However, it is first necessary to explain the general structure of the original Ingolia Riboseq methodology in order to explore how it has since been expanded to accommodate more species, tissues, and conditions of study.

1.3 Protocol structure overview

1.3.1 Lysis

The Ingolia methodology has been altered in both minute and substantial ways since its development, but the overall workflow remains the same. This starts with lysing the tissue, often in the presence of reagents such as cycloheximide, which stops translation elongation during lysis by blocking tRNA translocation. Lysing is done via application of lysis buffer (typically made from Tris-Cl, NaCl, $MgCl_2$, DTT, cycloheximide, Triton X-100 and Turbo DNase I), which is distributed in the sample through either pipetting (cell culture), or homogenization via douncer for other sample types (Ingolia, Brar, Rouskin, McGeachy, & Weissman, 2012).

Cell cultures, yeast samples, and fresh tissue have proven ideal for this procedure due to active translation at the time of collection, their ability to be experimentally manipulated (such as nutrient rich or starved conditions in the original study), and limited degradation of the tissue. However, flash frozen tissue, such as that from animal models, has also been used successfully.

1.3.2 Ribosome footprinting

Lysed cells are centrifuged to remove debris, and the RNA rich supernatant is treated with an RNase cocktail (RNase I in the original procedure) to degrade all of the RNA that is not bound by the ribosome. The ribosome-which is more resistant to RNAse action- prevents RNAses from accessing the mRNA that is bound by the ribosome, essentially protecting it from degradation at this stage. After this “footprinting” occurs, the ribosomes are isolated via sucrose cushion or S-400 size exclusion columns.

Sucrose cushion operates on the principle of a density gradient. In Riboseq, the RPF digestion is added to an ultracentrifuge tube, followed by the addition of the sucrose cushion solution at the bottom of the tube. The lysate detritus rises to the top of the sucrose in a discreet layer, and the ribosome pellet remains at the bottom. After centrifugation, the supernatant of lysate and sucrose can be removed, keeping the remaining ribosome pellet (Ingolia et al., 2012).

Alternatively, S-400 spin columns use gel filtration by applying the RPF digestion to columns with a resin with different pore sizes. With centrifugation, molecules larger than the largest pores are excluded from the gel and are eluted first, and the smallest molecules are eluted last as their penetration of the gel matrix delays movement. This is how the monosomes (a unit of one ribosome and its bound mRNA) are eluted out of the column last (Kodzius et al., 2006). The ribosome pellet is then treated with a lysing agent such as QIAzol or TRIzol to remove the ribosome from the RNA, and then the RNA is purified (with a kit such as the Qiagen miRNeasy kit, or the Zymo RNA Clean and Concentrator kit). A size selection purification for the RPFs is then performed on a polyacrylamide Tris-borate-EDTA (TBE) gel, to obtain ~30 nucleotide long purified RPF fragments, and avoid smaller, degraded fragments (Ingolia et al., 2012).

1.3.3 Library construction

From there, sequencing libraries are prepared from the RPF fragments. First, a linker is ligated to the RPF fragments, which will serve as a priming site for reverse transcription. The reverse transcription reaction begins with the addition of reverse transcription primer.

The fragments—now with attached linker ligation and reverse transcription primer—undergo an incubation reaction in the presence of a reverse transcriptase and dNTPs (as well as buffer, RNase inhibitor, and DTT) to synthesize complementary DNA.

From there, the single strand cDNA is circularized with circuligase. Circularization is conducted so that the resulting RPFs are now flanked by two priming sites, to facilitate PCR amplification. Finally, the cDNA is amplified and converted into viable RPF libraries. The two priming sites used to create the circularization serve as annealing sites for forward and reverse primers in a PCR reaction (also including buffer, dNTPs, and Phusion polymerase). The indexed reverse primers vary in sequence, and effectively barcode the amplified RPFs. They are now functional Riboseq libraries that can be sequenced for analysis. Library construction may also include a ribosomal RNA (rRNA) depletion step to remove these highly expressed molecules, although not all studies utilize this step.

1.3.4 Sequencing and analysis

Amplified and barcoded libraries are sequenced via next generation sequencing. There are a number of bioinformatic pipelines that have been used to analyse Riboseq data, including ones developed for Riboseq datasets (such as RiboFlow, RiboR RiboPy, RiboStreamR, and RiboProfiling), as well as standard RNAseq pipelines (the 2009 Ingolia paper uses GAPipeline v0.3.0 and SOAP v1.10) (Ingolia et al., 2009; Ozadam, Geng, & Cenik, 2020; Perkins, Mazzoni-Putman, Stepanova, Alonso, & Heber, 2019; Popa et al., 2016).

1.4 Alterations to the original protocol

Both minute and major alterations to the original procedure have been applied to Ribosome profiling studies over the years, and will be outlined in the upcoming section. Supplementary Table 1 and Supplementary Table 2 demonstrate the extent of diversity in altering the procedure.

1.4.1 Lysis

The first of the minor alterations that can be found across the procedure is the specific reagents and centrifuge times used for the lysis buffer. While there is minute variance in this aspect of the procedure, it does not seem to vary in any substantial way with regard to species or tissue being studied, or have obvious effects on the results.

1.4.2 RNase cocktail

An early and crucial step of the Riboseq procedure is the treatment with RNases, which degrade the non-ribosome bound mRNA. The standard method uses RNase I, which is nucleotide indiscriminate in hydrolyzing single-stranded RNA into 80S ribosome monomers (Ingolia et al., 2009). The majority of Riboseq studies use this RNase; however, RNases A, T1, S7, and MNase are also used (Gerashchenko & Gladyshev, 2017; Liu et al., 2019). RNase selection is an important consideration for successful Riboseq, given that RNases have varying efficacies by both species and tissue type. RNAse I and A successfully generate high resolution data in yeast and a variety of other species, but

degrade rRNA and mRNA more aggressively than some other RNases. This can make them less ideal for Riboseq in tissues that are more susceptible to degradation, especially if they erode the rRNA of ribosomes so severely it renders the RPFs themselves vulnerable to degradation.

RNase I, while capable of generating robust data in yeast experiments, cannot be assumed as a "one size fits all" footprinting treatment across all Riboseq studies and tissues. It is on this basis that Gerashchenko and Gladyshev sought to investigate the efficacy of common Riboseq RNases across model organisms (Gerashchenko & Gladyshev, 2017). They performed ribosome footprinting with RNases I, A, S7 and T1. One of the main sources of difference among these RNases is their cutting preference: RNases I and S7 cut at all four nucleotides, T1 cuts only at guanines, and RNase A cuts after cytosine and uridine.

Significant differences in efficacy of these RNases were found across trials with *E. coli* bacteria, budding yeast, *C. elegans*, *D. melanogaster*, and house mice. All of the RNases could be used to digest *E. coli* RNA, except for RNase I, which is inactive in the *E. coli* lysates. By contrast, only RNase T1 proved able isolate intact RPFs in the *Drosophila* samples. Mouse liver samples also demonstrated an RNase specificity, with RNases S7 and T1 successfully producing RPFs (Gerashchenko & Gladyshev, 2017).

RNase I has also been found to disrupt the integrity of monosomes in various mammalian tissues, to the point that RPF preservation can be compromised. For this reason, certain mammalian Riboseq studies (such as in mouse cortical tissue) have depended on an RNase A + RNase T1 cocktail. This cocktail produces higher yields for libraries (especially in cortical and neural tissues), but at the cost of lower reading frame resolution due to base cutting preferences. It is suited to the study of brain tissues (such as in mouse models)

which exhibit more modest translation and gene expression changes compared to many other cell types. One study that used this cocktail studied the dysregulation of protein synthesis in fragile X syndrome. (Liu et al., 2019).

Drosophila samples have also proved too sensitive for RNase I treatment, likely due to the uniqueness of their rRNA sequences and structures. As a substitute, studies such as those by Dunn et al. (2013) have used micrococcal nuclease (MNase), which sufficiently measures translation rate but is not a standard RNase used in Riboseq experiments due to its strong 3′ A/T cutting bias. This bias can create ambiguity in where the RPF begins, and uncertainty in determining the first P-site (the site that holds the tRNA that is associated with the growing polypeptide chain during translation). Accurate P-site mapping is an important part of establishing triplicate periodicity, and deeper analysis of RPFs (Dunn et al., 2013). It has also been used in Riboseq studies involving *E. coli* (Zhang et al., 2017).

After an RNase cocktail has been selected, the incubation time may also vary. While 45 minutes at room temperature is the standard, 30, 40, or 60 minutes have all been used as well, depending on the fragility of the tissue, and how long it will take the selected RNase to degrade the non-ribosome bound mRNA (Chew et al., 2013; Dunn et al., 2013; Guo, Ingolia, Weissman, & Bartel, 2010; Hsu et al., 2016; Jensen et al., 2014).

Clearly, these studies demonstrate the importance of RNase selection in successfully capturing ribosome footprints. Depending on the tissue or species of study, use of RNase I may not be appropriate. In addition to the RNase itself and incubation time, the concentration of the RNase, and incubation temperature during footprinting have all been found to significantly impact the quality of RPF samples, especially in low input sample

types that typically generate less useable RNA, and experimental conditions in which differential translation is less dramatic (Liu et al., 2019).

The use of various RNase treatments has allowed for translatome sequencing to be feasible in a variety of species and tissue types, which has greatly expanded the scope of this method.

1.4.3 Monosome purification

Post purification, the remaining single-stranded RNA footprint fragments are converted into cDNA libraries for sequencing. While many studies still rely on the original method, a number of alternative library preparation methods have been developed. Some differ only in omitting the circularization step during reverse transcription, or using other pre-existing RNA library preparation techniques (Bazzini, Lee, & Giraldez, 2012; Calviello et al., 2016; Guo et al., 2010).

Conversely, one procedure used for Riboseq library preparation differs substantially from the original methods paper. Hornstein et al. (2016) used a ligation free method to generate Riboseq libraries using Illumina's SMARTer technology (Hornstein et al., 2016). The conventional methods are relatively intricate and take days to prepare with multiple ligation, circularization, and gel purification steps. By contrast, the ligation free SMARTer approach can be completed in one day with only one initial gel purification step. It works by utilizing a template-switching reverse transcriptase, which adds sequences to both the 5' and 3' ends of the cDNA that allow for both 5' and 3' universal adapters to be added to synthesized cDNA within a single reaction (Hornstein et al., 2016). This procedure, in addition to being efficient is also relatively sensitive, with only ~1ng of input required for successful library

construction (compared to 2µg required for the original library method) (Hornstein et al., 2016; Ingolia et al., 2012). This makes this option particularly useful for low input samples that may be prone to degradation, such as post-mortem brain tissue.

1.4.4 Alternative library preparation methods

Since the original Riboseq methodology was published, a number of commercial Riboseq kits became available (although some of them have since been commercially discontinued). These kits typically involve the use of their own lysis buffers, RNase cocktail, and library preparation PCR primers. They therefore encompass alterations to several major steps of the methodology which warrant their own section of discussion. A typical Riboseq Kit performs the steps from cell lysis to library preparation. Some of these kits cater to specific subjects of study (such as mammals or yeast), making them ideal for undertaking an efficient species-specific Riboseq procedure. Their efficiency in general, in providing pre-prepared reagents and optimization guidelines, has made them a common selection in various Riboseq studies.

The Illumina ARTseq™ Ribosome Profiling Kit (which has a mammalian and yeast protocol), follows many of the standard steps of the original library preparation procedure exactly: specifically the 3'ligation, reverse transcription, and circularization steps. The main differences of the ARTseq™ kit involve the proprietary RNase and forward PCR primers, as well as the use of ScriptMiner™ Index PCR Primers instead of the index primers used by the Ingolia method. The ARTseq™ Ribosome Profiling Kit has been used in additional species, including zebrafish and *E. coli* (Bazzini et al., 2014; Bazzini, Lee, & Giraldez, 2012; Zhang et al., 2017).

Another Riboseq kit that was available for streamlined translatomics is the Illumina TruSeq Riboprofile Kit (although it has been commercially discontinued). Like the ARTseq™ kit, the library preparation steps also function on the same basic 3'ligation, reverse transcription, and circularization principles of the original method. Like the ARTseq™ kit, it produced its own branded RNase cocktail, index PCR primers (TruSeq Riboprofile Index PCR primers 1-12, and forward primer), and their own polynucleotide kinase (PNK) and reverse transcription mix.

1.4.5 Ribosomal RNA (rRNA) depletion

When included, the rRNA depletion step of the standard procedure occurs after circularization of the cDNA and before PCR amplification of the libraries. Depending on the Riboseq procedure, and relative integrity of the monosomes throughout the procedure, residual rRNA can account for up to 84% of the RNA in a Riboseq sample (Ingolia et al., 2009). Some Riboseq studies do not specify this step in their methodology, but if the rRNA fraction is substantial, its removal is imperative for successful RPF amplification and sequencing. In the original method, this depletion is accomplished via hybridizing samples to biotinylated rDNA-complementary oligonucleotides, and then removing them using MyOne Streptavidin C1 DynaBeads, a form of subtractive hybridization (Ingolia et al., 2009).

Other methods have implemented commercial rRNA depletion kits such as the Ribo-Zero™ kit from Illumina (which removes rRNA via an enzymatic reaction pre-cDNA synthesis), and post sequencing bioinformatic methods such as Bowtie or SILVA. (Amorim et al.,

2018; Bazzini et al., 2012; Calviello et al., 2016; Chew et al., 2013; Cho et al., 2015; Fisunov et al., 2017; Ingolia et al., 2011; Zhang et al., 2017). Similar to library generation methodologies, differences in the selection of methodology have been largely at the discretion of the researcher. There does not seem to be a bias towards selecting depletion method by species, as can be observed in (Supplementary Table 1, Supplementary Table 2).

1.5 Ribosome profiling of human post-mortem brain tissue as a future direction

Since the original Ingolia methodology was first conceived to examine changes in translation in nutrient starved yeast, the ways in which ribosome profiling can be conducted has substantially diversified. It has since evolved to be useable in a multitude of species—including mammalian tissue—to examination of a wide array of conditions (Supplementary Table 1, Supplementary Table 2). Both minute and more substantial methodological changes have been made in order to accommodate these tissues, conditions, and pathologies. The most important changes to this methodology involve the selection of RNase cocktail and the development of lower-input RNA library preparation methods such as SMARTer® (Illumina®). Both of these advancements can be universally applied across fields of study and expand the array of samples that can be investigated through ribosome profiling. Ribosome profiling has already allowed the discovery of some of the complexities of translatomic regulation, and its scope and potential will continue to be expanded in the future.

One way in which it had not yet been expanded at the onset of this study was through the creation of a ribosome profiling procedure for more delicate tissues, such as human post

mortem brain tissue. This tissue, while more fragile than cell cultures and peripheral tissue, can be an effective way to directly study transcriptomic alterations- like differentially expressed genes (DEGS) - in the brains of individuals with psychopathologies with an epigenetic basis such as Major Depressive Disorder (MDD), post-traumatic stress disorder (PTSD), bipolar disorder, Alzheimer's disease, and Schizophrenia (Abdolmaleky et al., 2011; Lutz et al., 2017; Rao, Keleshian, Klein, & Rapoport, 2012; Zannas, Provencal, & Binder, 2015).

While there have been a great deal of transcriptomic differences identified in these psychopathologies, translatomic differences have remained more elusive. It cannot be assumed that all differentially expressed genes (DEGs) in a given cell will also be differentially translated genes (DTGs) due to mRNAs that are exported for synthesis and variance in mRNA degradation - which is a key component to the regulation of gene expression and highly interconnected to translational output (Huch & Nissan, 2014). DTGs which likely have an impact on phenotype due to their involvement in protein synthesis, therefore merit their own assessment that cannot be assumed via transcriptomics and characterization of DEGs. In order to gain a more complete picture of these frameworks, ribosome profiling of human-post mortem brain tissue would be a great asset.

Yet, ribosome profiling typically makes use of *in vivo,* fresh, or flash frozen samples (including brain tissue from mouse models) to facilitate the capture of the translation process before significant degradation to the ribosomes and RPFs occur (Bartholomaus, Del Campo, & Ignatova, 2016; Liu et al., 2019; Weinberg et al., 2016). Human post-mortem brain tissue presents with a significant challenge as it is often difficult controlling for post mortem interval (the period between death of the individual and collection of the brain

tissue), and limiting post mortem cell degradation in this time. Factors including location of death (such as distance from hospital/facilities), the interval between time of death and discovery of death (often significant if someone dies alone), and expediency of collecting brain tissue after death can greatly extend the post mortem interval and affect the ability to collect usable RPFs. As well, limiting post mortem cell degradation once tissue is collected via proper storage and freezing protocol is crucial to assist preserving intact ribosomes and RPFs (Bartholomaus et al., 2016). Prolonged freezing, however, could have a similar effect on reducing tissue quality. Tissues stored over longer periods of time – such as human post mortem tissue stored in brain banks for future use- may be particularly susceptible to these effects.

These factors significantly complicate the endeavour to find a functional ribosome profiling protocol for human post-mortem brain tissue. However, if found, it could greatly elucidate the role of DTGs in many human psychopathologies.

1.6 Aims

The primary goal of this project was to determine the feasibility of, and optimize, a protocol for ribosome profiling of human post-mortem brain tissue. I tried multiple alterations in most of the major methodology steps, such as RNase use in ribosome footprinting, cDNA library construction, sample purifications, and bioinformatics analysis, in order to try to accommodate the potential fragility of the tissue. It was originally hypothesized that this could be achieved within the scope of this study, with a novel procedure being constructed from tried RNase selections, purification methods, and cDNA library construction. If this

had been achieved, this methodology would have been used to examine DTGs in the orbital frontal cortex (OFC) of individuals who experienced early life adversity (ELA) and later developed MDD and died by suicide.

Chapter 2. Methods

2.1 Overview of major trials in the methodology

Table 1. Overview of major trials in the methodology.

Major Trials	Samples	Relevant sections
RNAse Condition for Footprinting	2 BA 8/9 samples	2.3
	Aliquots of same BA 8/9 sample (described in Figure 2.1)	2.3,3.3
Illumina® TruSeq Ribo Profile (Mammalian) Library Prep Kit	2 BA 8/9 samples	2.4, 3.2
	Aliquots of same BA 8/9 sample (described in Figure 2.1)	3.3
SMARTer® smRNA-Seq Kit (Illumina®) RNase trial	Aliquots of same BA 8/9 sample (described in Figure 2.1)	3.3
	9 Human post-mortem Orbital Frontal Cortex (OFC) samples	3.4
	3 cell culture samples (Be(2)-C neuroblastoma)	3.5
	32 mouse frontal cortex samples	3.7
NEBNext® Multiplex Small RNA Library Prep Set for Illumina®	3 cell culture samples (Be(2)-C neuroblastoma)	3.5
Galas Lab 4N RNA library preparation	3 cell culture samples (Be(2)-C neuroblastoma)	3.5

2.2 Samples

Human post-mortem brain tissue used in study came from the Douglas-Bell Canada Brain Bank, situated within the Douglas Mental Health University Institute. This study was

approved by the Douglas Hospital Research Ethics Board, and written informed consent was obtained from the next-of-kin for each subject through an agreement with the Quebec Coroner's Office. The two primary regions of human post-mortem brain tissue utilized throughout the trials were Brodmann area 8/9, and the orbital Frontal Cortex (OFC). In the initial trial, which used the Illumina® TruSeq Ribo profile (Mammalian) Kit methodology, Brodmann 8/9 tissue was used.

In subsequent trials with human post-mortem brain tissue, OFC tissue was used. OFC tissue came from 3 groups: control, ELA/MDD/suicide, and MDD/suicide/no history of ELA. The rationale behind these groups was that, if ribosome profiling and sequencing of the RPFs proved successful, DEGS and DTGs could be analyzed across groups for the study of the ELA to MDD and suicide epigenetic framework (Bremner et al., 2002; Hanson et al., 2010; Tonelli et al., 2008).

All human post-mortem brain tissue was excised from frozen whole brain, into 100mg frozen and unfixed sub-samples, which were used for ribosome profiling.

Within this larger study, a smaller study was also performed to determine the effect of PMI on efficacy of ribosome profiling. To control for PMI and freezing conditions, frontal lobe tissue (analogous to the human brain tissue used in this study) from a mouse model was used. Mice brain tissue was divided into the following experimental groups: control (immediately flash frozen), and then 4 experimental conditions which included spending 0,3,6, and 12 hours respectively at room temperature before being refrigerated for 24 hours and then flash frozen.

The final type of tissue used in this study was the use of Be(2)-C Neuroblastoma cell culture grown on a 1:1 mixture of ATCC-formulated Eagle's Minimum Essential Medium and F12 Medium, supplement with 10% FBS. Cells were cryopreserved in the vapor phase of a liquid nitrogen tank and then thawed for growth. For harvest, the cells were washed with PBS, detached with trypsin-EDTA, and trypsin was deactivated with growth media. Cells were then pelleted by spinning at 1000rpm for 5min at ~24°C, washed with PBS, and centrifuged at max speed for 2min at 4°C. The pellets were stored at -80°C between use. These samples were used as a control for tissue type in conjunction with human post-mortem brain tissue and mouse brain tissue.

2.3 Lysis

In all trials, the ribosome profiling procedure began with adding 100mg of human post-mortem brain tissue to of 1ml of mammalian lysis buffer (200µL 5X Mammalian Polysome Buffer, 100 µL 10% Triton X-100 10 µL 10% NP-40, 10 µL 100 mM DTT, 10 µL DNase I (1U/µl), 1 µL 100 mg/ml cycloheximide, 678 µL Nuclease-free water). The samples were manually homogenized in a glass douncer. The homogenized lysate was centrifuged at 14000g for 10 minutes, and then divided into 2 aliquots: 200µl (with 20µl SDS) to prepare total RNA libraries, and 100µl RPF samples. Total RNA samples were used as a control to gauge functionality of the procedure irrespective of RPF capture quality. This procedure was established in the preliminary trial and applied throughout subsequent ribosome profiling trials.

2.4 Ribosome footprinting and purification

In the original preliminary trial, ribosome footprinting was performed with 2 (male) BA 8/9 samples, each prepared with 2 different Rnases: the TruSeq Ribo Profile (Mammalian) Library Prep Kit (Illumina®) Rnase, and a cocktail of equal parts RNAse S7 and T1 (totalling 10U/ul). A cell lysate control was also prepared. All samples were incubated with the Rnase at 25°C for 45 minutes. The footprinting reaction was stopped with the addition of Rnase inhibitor (ABI).

The RPF samples were then purified via microspin S-400 HR columns (Illustra). An abundance of ~30nt fragments in several Rnase conditions in the Rnase trial, and in ~30nt bands during Urea PAGE gel extraction seemed to indicate success with this S-400 HR columns purification method. It was continued for subsequent trials.

Following S-400 spin column application, the total RNA and the RPF samples were purified via a RNA clean and concentrator-5 kit (Zymo). Quantification of RPF RNA was regularly performed via NanoDrop spectrophotometer once this stage of purification was complete and volume was consistently above the 1-5µg of total and RPF RNA required as input.

Success at capturing the ~30nt RPFs was measured using an Agilent Small RNA kit on an Agilent 2100 Bioanalyzer. Samples were prepared according to protocol.

To more explicitly determine the most effective Rnase or Rnase cocktail for human post-mortem brain tissue, a subsequent trial with 8 RPF samples from the same Brodmann area 8/9 human post-mortem brain sample was created. Samples contained 5 units of (10 U/µl) nuclease mixture per 1 OD_{260}, with the exception of one condition that tested efficacy at

1/10th the concentration of RNAse. Sample conditions were: S7 Rnase at 45 minute incubation, T1 Rnase 45 minute incubation, Rnases S7/T1 at 45 minute incubation, (Illumina®)TruSeq Ribo Profile (Mammalian) kit Rnase at 45 minute incubation, Rnases S7/T1 at 20 minute incubation, Rnases S7/T1 at 90 minute incubation, Rnases S7/T1 at 45 minute incubation and 1/10th the amount of enzyme, and undigested total RNA (control). All samples were incubated at 25 ° C for 45 minutes. The footprinting reaction was stopped with the addition of Rnase inhibitor (ABI). Purification of the RPFs was completed using S-400 spin columns as described above and RPF capture was measured with an Agilent Small RNA kit on an Agilent 2100 Bioanalyzer.

Produced measures of average RNA fragment size were used to select the optimal Rnase conditions for subsequent ribosome profiling endeavors. Additionally, to validate these results, the S7, T1, S7/T1, S7/T1 with 1/10th the amount of enzyme, and Kit Rnase samples (all incubated for 45 minutes) were converted into cDNA libraries using 2 small RNA library preparation protocols (Illumina®TruSeq Ribo Profile (Mammalian) Library Prep Kit, and SMARTer® smRNA-Seq Kit for Illumina®) and then sequenced using a Miseq V3 150 cycle kit.

In all successive ribosome profiling trials, a cocktail of Rnases S7/T1 incubated at 25°C for 45 minutes was used.

To purify and size select the RPFs from remaining monosome rRNA and debris in the reaction, the RPF sample was run on a 15% Urea PAGE gel (Novex). Urea PAGE gels wells were loaded with 10µl RNA control (5µl TruSeq Ribo Profile RNA Control, 5 µl Gel Loading Buffer II), 10µl RPF samples (5µl sample with 5µl Gel Loading Buffer II), and ladder (4µl 20/100 Oligo Ladder, 5µl Gel Loading Buffer II, 1 µl nuclease free water).

Reactions and ladder were placed on a preheated thermal cycler set at 95°C, and incubated for 5 minutes to denature. Samples and ladder were subsequently removed and placed on ice before being loaded into a 15% urea-polyacrylamide gel and run at 180 V for 70 minutes or 210 V for 1 hour.

The urea PAGE gel was then removed from the casing, and stained in pre-chilled SYBR Gold, which was kept at 4°C to avoid diffusion of libraries across the gel, and shaken for ~40 minutes to ensure adequate staining and even distribution across the gel. The gel was then viewed under a blue light transilluminator and excised at the 25-40nt range to capture RPFs.

The excised gel slices were transferred to a 0.5 ml tube with a hole punched in the bottom by sterile 20-gauge needle, which was placed in a larger 1.5 ml micro-centrifuge tube. They were then centrifuged at 12,000 × g for 2 minutes. Nuclease-free water (400 µl), 5 M ammonium acetate (40 µl), and 10% SDS (2 µl) were then added to the gel slurry occupying the bottom tube. The reactions were shaken at 2-8°C for 24 hours to elute the RNA. The eluted slurry was transferred to 1.5 ml filter tubes and centrifuged at 2000 × g for ~3 minutes. Once transferred to fresh tubes, 2 µl glycogen and 700 µl 100% isopropanol were added to each tube, and then Set aside at -25°C to -15°C for up to 1 hour. Reactions were centrifuged again at 12,000 × g for 20 minutes at 2- 8°C, and the resulting RPF pellet was washed with ice-cold 80% ethanol and then air dried. The resulting pellets were re-suspended in 20 µl nuclease-free water. The reactions were then prepared for conversion into cDNA libraries.

2.5 Library preparation: Illumina® TruSeq Ribo Profile (Mammalian) Library Prep Kit

Once protocols for ribosome footprinting and purification were established, trials to identify a successful method of library construction for human-post mortem brain tissue RPFs were conducted. The Illumina® TruSeq Ribo Profile (Mammalian) Library Prep Kit very closely resembles the Ingolia method for the ribosome profiling of mammalian cells, and was used in initial trials as a means to test the efficacy of the standard method before attempting optimization trials.

Two samples from Brodmann 8/9 were used in the initial trial of this methodology, with each sample being separated into both total RNA and RPF aliquots following lysis (N=4). Samples were prepared with the lysis, footprinting, and purification methods previously described.

From here, both RPF and total RNA samples were subjected to T4 PNK treatment following the TruSeq Ribo Profile PNK kit (Illumina®) and protocol, to phosphorylate the 5’ fragment ends, remove the 3’ phosphoryl groups, and fragment the total RNA sample so that it could be processed with the same small RNA library preparation as the RPF samples. PNK treatment was performed according to procedure, as described in the TruSeq Ribo Profile (Mammalian) Kit (Illumina®) User Guide. After PNK treatment, 3’ adapter ligation, reverse transcription, and cDNA circularization were performed according to procedure, and a second round of 15% Urea PAGE purification was performed as described above.

Following the second purification, PCR amplification of the circularized cDNA products was performed. On initial attempt, PCR amplification was performed exactly as described in the procedure.

After completion of this procedure, the PCR products were visualized on a high resolution agarose gel to evaluate whether amplification was adequate. After 3' adapter ligation, reverse transcription, cDNA circularization, and PCR primer attachment the resultant fragments should be in the ~140-160 bp range. Amplification was inadequate for sequencing after having run 9 PCR cycles, as per the original procedure. The procedure was re-performed with the sole alteration of augmenting the PCR cycles to 25. With the PCR cycles increased, visualization on high resolution agarose gel revealed substantial bands in the ~140-160bp range.

To perform a final purification of the amplified samples prior to sequencing, both NEBNext® AMPURE XP beads (Illumina®) and excision from high resolution agarose gel were attempted. An excess of sample was lost during the AMPure XP bead (Illumina®) attempt, rendering it impossible to retain enough cDNA for sequencing. High resolution agarose excision was selected as the final purification method for the remainder of the TruSeq Ribo Profile (Mammalian) Library Prep Kit (Illumina®) trial. Samples were sequenced using the Miseq V3 150 cycle kit.

After the original TruSeq Ribo Profile (Mammalian) Library Prep Kit (Illumina®) trial, another was run using some of the BA 8/9 Rnase optimization aliquots: S7, T1, S7/T1, S7/T1 ($1/10^{th}$ concentration of Rnase), and Kit Rnase (N=5). These samples were later prepared for sequencing and underwent the same after library construction purification with the first SMARTer® smRNA-Seq Kit for Illumina® samples (from the same samples) for

comparison of library creation feasibility. However, too much material was lost throughout the process to proceed with sequencing.

The Illumina® TruSeq Ribo Profile (Mammalian) Library Prep Kit was discontinued during the undertaking of this study. In the absence of another commercially available ribosome profiling kit adapted for mammalian tissue, three non-ribosome profiling specific methods for small RNA library preparation were tried. These methods were appropriate for consideration given that post ribosome footprinting and purification, the RPF samples can be treated as any other small RNA sample with regards to library preparation and sequencing. Subsequent alterations tried after this initial attempt are outlined in the results section.

2.6 Bioinformatic analysis

Bioinformatic analysis has largely followed the procedure laid out by Ingolia et al (2012), with emphasis on the hallmarks of successful RPF capture, such enrichment for reads mapping to coding regions, and triplicate periodicity (Ingolia et al., 2012). High coding region enrichment is a hallmark of successful RPF capture due to the fact that the majority of mRNA fragments bound by the ribosome should be so for the purpose of protein synthesis. Another hallmark, triplicate periodicity, is defined as a characteristic peak in number of fragments with a specific reading frame (defined as every third nucleotide) per fragment size, centered at the RPF fragment size range. The implication of this is, as RPFs should be synthesizing amino acids to build proteins and all amino acids are encoded by 3

nucleotides, that fragments encased by ribosomes to build proteins would start on every third nucleotide to properly synthesize the correct amino acids. This “triplicate periodicity” pattern resembles an ascending peak in reading frame specificity in the 20-35nt range, reaching its peak at ~25-27 nt (Hornstein et al., 2016; Michel et al., 2016).

The bioinformatics analysis pipeline specific to the Illumina® TruSeq Ribo Profile (Mammalian) Library Prep kit was used in all trials that used this library preparation method, and with the Rnase trial samples that were prepared with the SMARTer® smRNA-Seq Kit (Illumina®). First the reads were generated as .fastq files, and the remaining 3’ adapter sequences are removed with the FastX toolkit so that generated reads will be compatible with Bowtie and TopHat aligners. Then rRNA alignment and other contaminant identification was performed with Bowtie. The reads aligned to the rRNA-ome were minimal and successfully removed during this step, which accounts for a lack of rRNA depletion step during the ribosome profiling protocol.

Alignment of remaining RPF sequences was then performed with the TopHat aligner. The TopHat generated reads were then assembled into transcripts using the program CuffLinks, and differential expression was quantified using Cuffdiff. Final visualization of these output files that allow for observation of coverage, aligned reads, junction counts, comparison of the data to known transcripts is performed with the Integrative Genomics Viewer. The primary ribosome profiling quality metrics we examined with this pipeline were length distribution, and read density distribution (to examine coding enrichment). Since there is no step in this pipeline dedicated to triplicate periodicity, this analysis was performed with the program RiboGalaxy.

For the Rnase trial samples that were prepared with the SMARTer® smRNA-Seq Kit (Illumina®) , the coding enrichment analysis was still performed according to the protocol, but the poly(A) tails were removed manually first in an attempt to increase the coding enrichment profile and more accurately reflect the actual RPF fragments.

This bioinformatics pipeline was continued with the SMARTer® smRNA-Seq Kit (Illumina®) prepared samples (OFC and mouse brain trials), until the cell culture trial was performed. To determine if consistently low coding enrichment was attributable to the bioinformatics pipeline, an alternative pipeline was used. This new pipeline was derived from the Khoutorsky lab, and was followed according to the procedure performed in Uttam et al. (2018). Three alternate programs for coding enrichment were also evaluated for the samples from the Galas cell culture trial: PICARD, RseQC, and BROAD Institute.

Chapter 3. Results

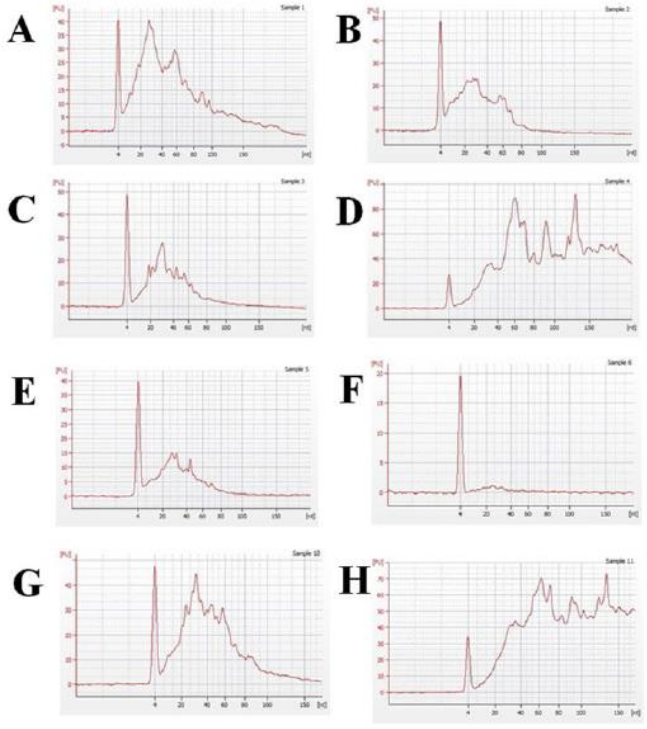
3.1 Lysis and footprinting

Success in lysis and ribosome footprinting was determined via Agilent Small RNA kit on an Agilent 2100 Bioanalyzer system. Success was measured as the highest peaks of fragments per sample occurring in the RPF range (~28-40nt).

In the first trial that included the Truseq (Illumina®) kit Rnase, and S7/T1 Rnase aliquots from 2 BA 8/9 samples, the expected peaks in the ~28-40 nt RPF range were only observed in the S7/T1 Rnase samples, and not the Truseq (Illumina®) kit Rnase samples which were derived from the same brain samples, or the cell lysate control. However, the peaks of the S7/T1 samples were much closer to 40 nt than the expected 30nt (Figure 1.1).

The subsequent Rnase trial included 8 aliquots from the same BA 8/9 sample, as described in (Figure 2).Bioanalyzer visualization of this trial indicated a single large peak in the expected ~28-40nt range, and at almost precisely the 30nt mark, in the S7, 45 minute incubation T1, 45 minute incubation, S7/T1, 45 minute incubation, S7/T1, 20 minute incubation, and S7/T1, 45 minute incubation, $1/10^{th}$ amount of enzyme samples. These samples start to decrease gradually in fragment abundance from the ~ 40-100nt range (Figure 1.1).

I



II

Electropherogram	RNAse Condition
A	S7 RNAse (10 U/µl), 45 minute incubation
B	T1 RNAse (10 U/µl), 45 minute incubation
C	S7/T1 RNase (10 U/µl), 45 minute incubation
D	Truseq (Illumina®) kit RNase (10 U/µl), 45 minute incubation
E	S7/T1 RNase (10 U/µl), 20 minute incubation
F	S7/T1 RNase (10 U/µl), 90 minute incubation
G	S7/T1 RNase (1 U/µl), 45 minute incubation
H	Undigested (total RNA)

Figure 1.1 RNAse trial fragment-length abundance **I.** Agilent Bioanalyzer small RNA electropherogram of RNAse trial samples (from aliquots of the same PFC BA 8/9 tissue) showing nucleotide length (x-axis) versus fluorescence intensity (y-axis) **II**. Table describing RNase conditions of each electropherogram.

The Truseq (Illumina®) kit RNase sample was remarkably identical to the undigested total RNA control sample in peak pattern, including the absence of a peak in the ~28-40nt RPF range, and several large peaks occurring in the 50-100nt range, likely indicating a failure to digest the total RNA surrounding the monosomes. The S7/T1, 90 minute incubation sample showed a complete absence of peaks, likely indicating an excess of RNase incubation time that degraded all of the RNA, including the monosome protected RPFs. While the S7/T1 20 minute incubation sample did indicate the expected ~30nt peak, the fluorescence intensity was much lower than the other samples with the expected RPF range peak. All of these conditions were eliminated as feasible RNase options

Based on the bioanalyzer results, all of the remaining conditions were judged as approximately equal in RPF capture quality, and were divided into aliquots and prepared for sequencing using both the Illumina® TruSeq Ribo Profile (Mammalian) Library Prep Kit, and SMARTer® smRNA-Seq Kit (Illumina®) small RNA cDNA library construction methods. The data was processed as described in section 2.6, and measures of coding region enrichment, triplicate periodicity, and RPF fragment length were all calculated. There were no substantial advantages with respect to any of these measures observed for any of the remaining RNase conditions (Figure 1.2). In light of this, the S7/T1 45 minute incubation condition was selected for all subsequent experimental use due to the potential assist in the diversity of RNases and cleaving sites, and the possibility that $1/10^{th}$ of the RNase concentration may not be adequate for complete ribosome footprinting of all samples.

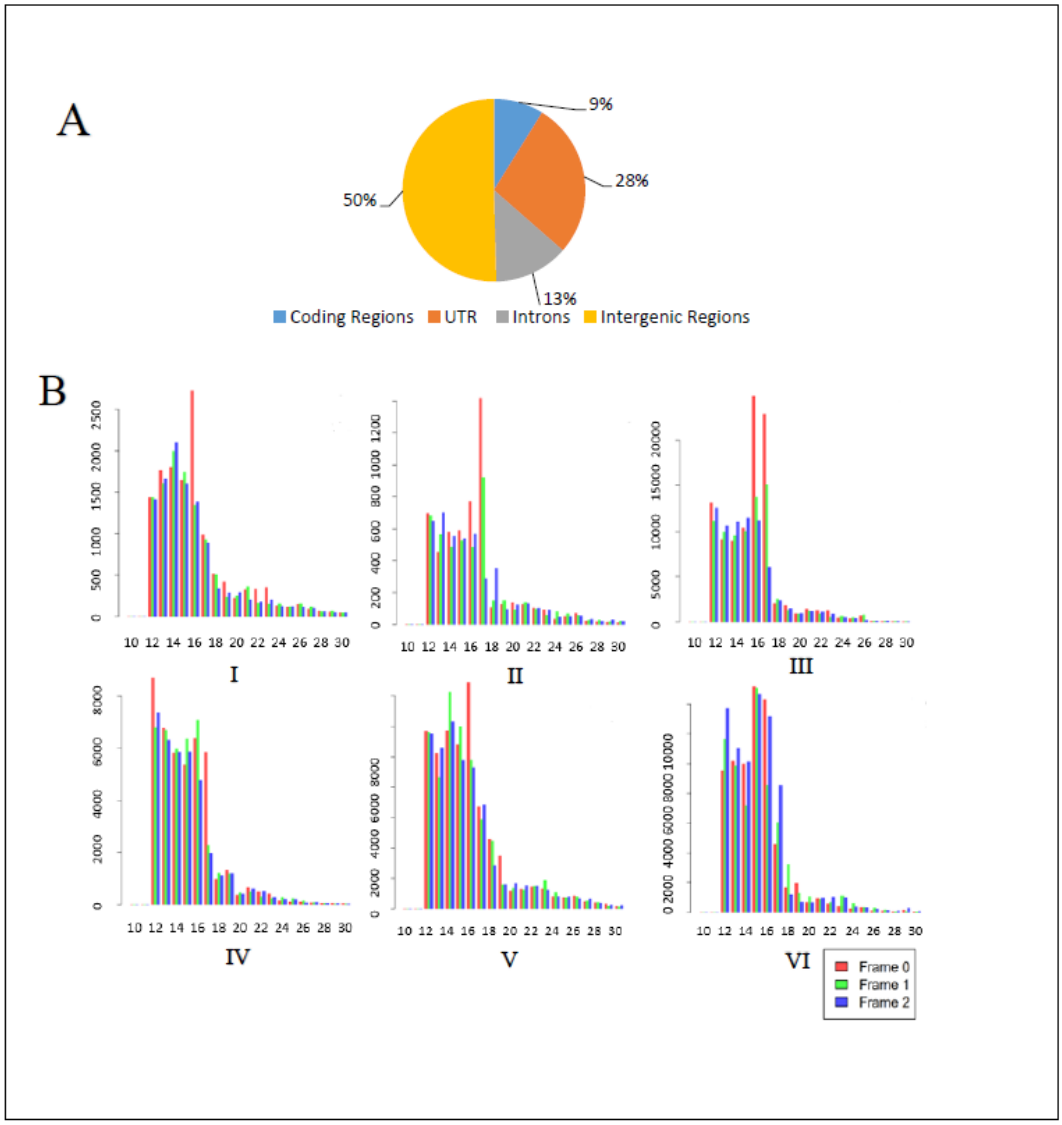


Figure 1.2 Hallmarks of RPF capture for the RNAse trial. **A.** Coding enrichment for S7/T1 RNAse (10 U/µl), 45 minute incubation **B.** Triplicate periodicity profiles showing fragment length in nt (x-axis) versus frequency of 5'ends mapping to each frame for **I.** S7 RNAse (10 U/µl), 45 minute incubation **II.** T1 RNAse (10 U/µl), 45 minute incubation **III.** S7/T1 RNase (10 U/µl), 45 minute incubation **IV.** Truseq (Illumina®) kit RNase (10 U/µl), 45 minute incubation **V.** S7/T1 RNase (1 U/µl), 45 minute incubation **VI.** Undigested (total RNA).

3.2 Illumina® TruSeq Ribo Profile (Mammalian) Library Prep Kit

The 2 originally tried BA 8/9 samples, which were footprinted using the S7/T1 RNase cocktail, were sequenced with this library preparation protocol to determine its feasibility for human post-mortem brain tissue without alteration to the protocol. Sample 1, and sample 2, were each subdivided into 1 RPF and 1 total RNA (control) aliquot and prepared for sequencing (N=4).

Miseq V3 150 cycle kit sequencing revealed that the percent of reads identified for each sample were: 33.915% for the sample 1 total RNA, 26.225% for sample 1 RPF, 9.497% sample 2 total RNA, 15.912% sample 2 RPF.

Principal component analysis (PCA) of the samples revealed greater overall similarity within the reads identified between the total RNA and RPF samples derived from the same original BA 8/9 sample (Figure 2.1). This defied the expectation that the RPF samples would show more similarity due to selection of actively translating sites, although differences between the samples and the presence of more reads identified overall (and therefore likely more genes to match) in sample 1 may have been relevant factors.

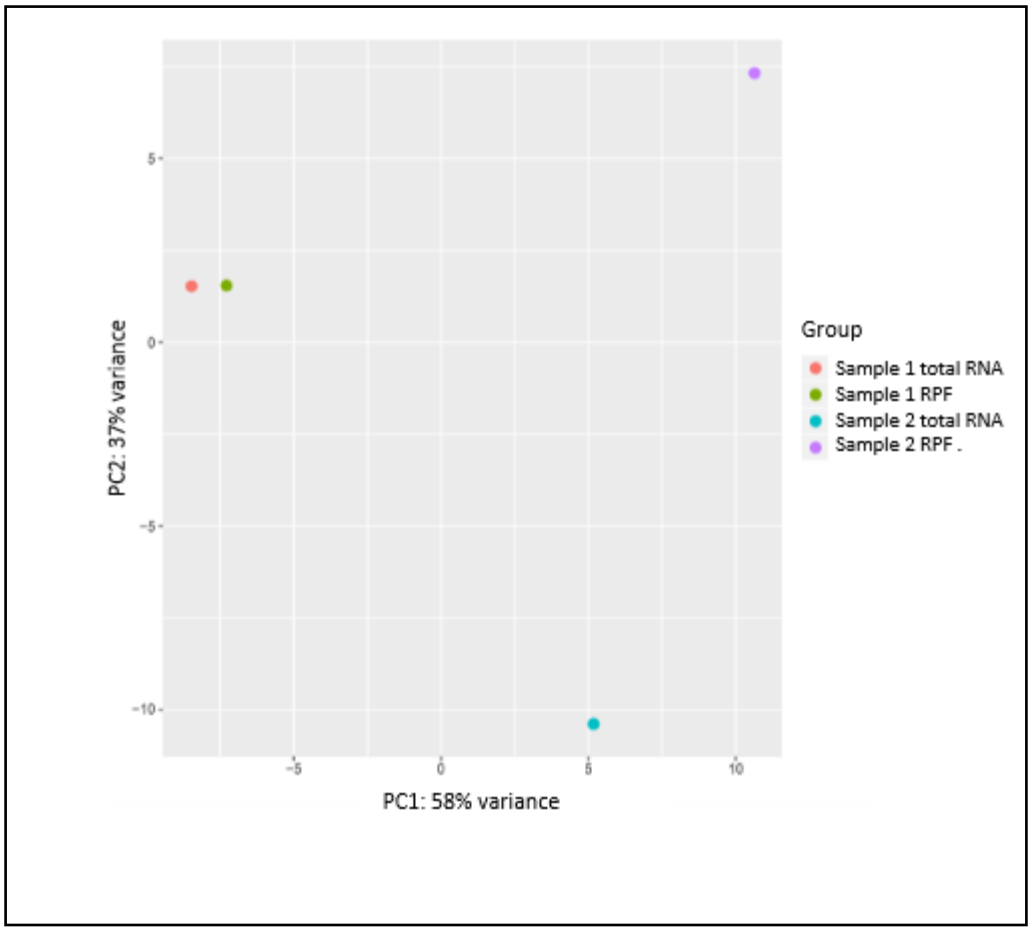


Figure 2.1 PCA analysis of TruSeq Ribo Profile (Mammalian) Library Prep Kit samples.

After bioinformatics analysis, the RPF samples showed a coding enrichment that was substantially below expected levels. When using S-400 size exclusion columns, the read density distribution for RPF samples is expected to be as follows: ~65.7% coding regions, 22.1% Untranslated regions (UTR), 1.5% introns, and 10.7% intergenic regions. Across the RPF samples the average read distribution was: 12.5% coding regions, 24.98% UTR, 17.76% introns, and 44.69% intergenic regions (Figure 2.2A). The total RNA samples met the expectations for coding enrichment: 20.17% coding regions, 31.10% UTR, 15.06% intron, 33.67% intergenic regions (Figure 2.2A).

Triplicate periodicity, which was analyzed with RiboGalaxy, also did not meet expectations for successful RPF capture. Across the 26-30nt range, there were no obvious peaks for any of the reading frames, likely indicating a lack of triplicate periodicity (Figure 2.2B).

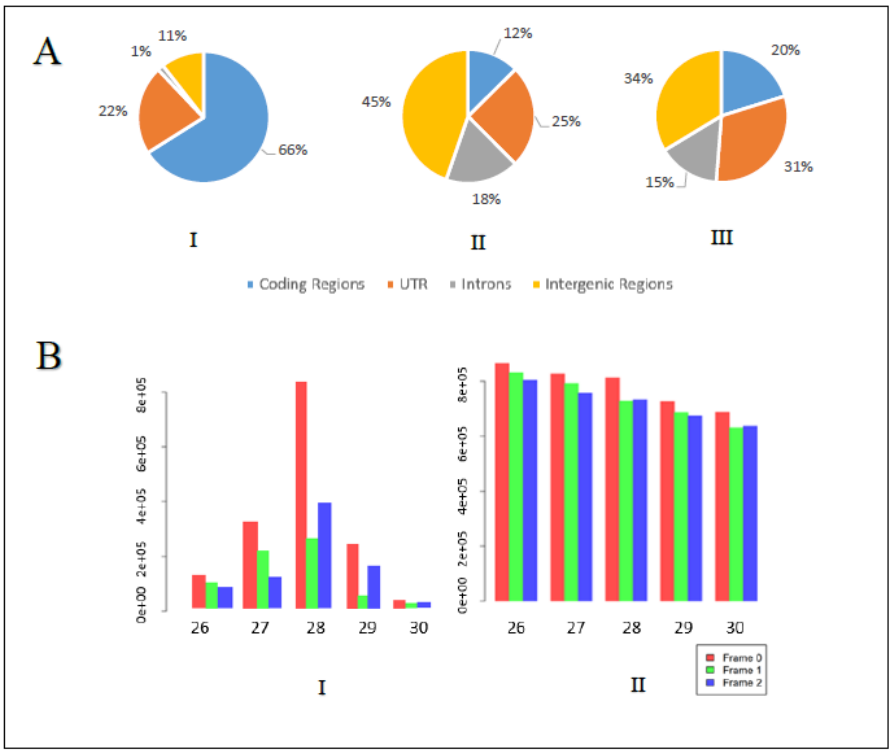


Figure 2.2 Hallmarks of RPF capture for the initial PFC BA 8/9 trial (Illumina® TruSeq Ribo Profile (Mammalian) Library Prep Kit) **A.** Coding enrichment profiles by percent reads aligned to each region in the legend **I.** Expected profile for RPF samples purified and isolated by size exclusion columns (such as S-400 spin columns). **II.** Average coding enrichment for both RPF samples sequenced in this trial **III.** Average coding enrichment for both total RNA samples sequenced in this trial. **B.** triplicate periodicity for **I.** The expected RPF profile **II.** One of the RPF samples sequenced in this trial.

The second and final trial using the Illumina® TruSeq Ribo Profile (Mammalian) Library Prep Kit which included different RNase conditions across the same BA 8/9 sample did not amplify sufficiently in the expected RPF range (~140-160bp), as verified via high resolution agarose gel (where excision was attempted above the visible ~113bp dimer range) and via TapeStation with Agilent D1000 ScreenTape, and was not sequenced.

3.3 SMARTer® smRNA-Seq Kit (Illumina®) RNase trial

Before the discontinuation of the Illumina® TruSeq Ribo Profile (Mammalian) Library Prep Kit, the SMARTer® smRNA-Seq Kit for Illumina® was tried consecutively with the afore mentioned protocol for sample loss, efficiency, and sequencing success comparisons. Samples from the same BA 8/9 RNase optimization aliquots: S7, T1, S7/T1, S7/T1 (1/10th concentration of RNase), and Kit RNase (N=5) were used.

Operating with a ligation-free workflow, this protocol was selected to simplify the multiple ligation and gel purification rounds required in the previous protocol. It was theorized that this simplification could potentially assist in preserving useable material for sequencing that could be lost in the purification steps.

It works by adding poly(A) tails to the 3’ end using poly(A) polymerase , then reverse transcribes the RPF using an enzyme with template-switching activity. In this reaction, a reverse transcriptase (RT), extends a primer (universal sequence on 5’ end) to produce cDNA, and adds a low complexity sequence to the 3’ end of the cDNA in a non-template directed fashion. In this reaction, there is another universal sequence adapter that is 3’ terminated with a low complexity sequence, which hybridizes to the tail added to the cDNA

by the RT. After hybridization of this second adapter, the RT switches templates and copies the second adapter onto the 3' end of the cDNA. So, the 5' and 3' universal adapters are simultaneously added to the cDNA in a single reaction without single stranded ligation or intermediate purification steps.

Prior to small RNA library preparation, T4 PNK treatment (New England Biolabs) was performed on all samples using their Non-radioactive Phosphorylation with T4 PNK or T4 PNK protocol. This procedure was completed according to protocol without alteration, with the maximum suggested 17 PCR cycles being selected for use. PCR products were purified with the NucleoSpin® Gel and PCR Clean-Up kit (Takara).

Library construction was validated via the Agilent D1000 ScreenTape System. Expected library peaks in the SMARTer® (Illumina®) protocol are in the range of 153 + input RNA size (nt), so 153+30nt= 183bp for RPF libraries. Bands observed in the ~165-190bp range were considered worth excising for final purification, as it was ~10bp above the expected primer dimer range, and less than 20bp above the expected RPF range.

The prepared SMARTer® (Illumina®) samples were pooled, as were the TruSeq (Illumina®) samples, and were run on the same high resolution agarose gel for final purification, and excised in both of their expected bp ranges respectively: 170-190bp, and 140-160bp. cDNA was extracted from the gel slices using the Qiagen QIAquick® Gel Extraction Kit. All expected bands were present but fainter than desired for sequencing. The pooled samples were mixed into a P5P7 primer cocktail (5µl P5P7, 45 µl library sample, 50µl Kapa) and amplified for an additional 5 rounds of PCR in total. Samples were then subjected to a second round of high resolution agarose excision and cleanup, and visualized with Agilent D1000 ScreenTape tapestation. Tapestation visualization revealed peaks in the

expected 170-190bp range for the SMARTer® (Illumina®) prepared samples, but peaks in the ≤120nt dimer range for the TruSeq (Illumina®) prepared samples. Miseq V3 150 cycle kit sequencing proceeding with only the SMARTer® (Illumina®) prepared samples.

The total percentage of reads identified for the SMARTer® smRNA-Seq Kit (Illumina®) RNase trial samples was 70.01%, with reads identified from each sample ranging from 8.827% -14.804%.

The coding enrichment of the RPF samples in this trial was substantially lower than expected. In each sample, the coding regions only accounted for 1% of the total reads identified. The average percentage for the other reads were: 9% UTR, 54% introns, and 36% intergenic regions. Overall volume of reads was also lower than the previous trial, potentially due to the poly(A) manual trimming that had not been done in the previous analysis (Figure 1.2B).

Triplicate periodicity results were mixed with respect to indicating successful ribosome profiling. The S7, T1, and S7/T1 (standard enzyme concentration) conditions did show a characteristic peak indicating reading frame preference for reading frame 0 across samples. However, these peaks occurred in the 16-17nt range, and not in the expected ~30nt range (Figure 1.2A).

3.4 SMARTer® smRNA-Seq Kit (Illumina®) OFC trial

The SMARTER® procedure was repeated for what was intended to be an OFC tissue pilot study. OFC tissue came from 3 groups (PMI <20-64 hours, Male): control (No MDD/suicide/ELA) (N=3), MDD/suicide/ELA (N=3), and MDD/suicide/no ELA (N=3).

The protocol was performed exactly as described above, with the only exception being an attempt to perform the final pre-sequencing purification with NEBNext® AMPURE XP beads (Illumina®). After both the AMPURE XP beads (Illumina®) and the previously performed high resolution agarose gel excision methods were tested on these samples, the gel excision method was selected to prepare the samples for sequencing. Excess material was lost during the AMPURE XP beads (Illumina®) method. Only one total RNA sample from the control cohort did not adequately amplify for sequencing. Samples were once again sequenced via Miseq V3 150 cycle kit.

After sequencing, there was a total of 22.316% reads identified in the control group samples (N=3). The MDD/suicide/ELA (N=3), and MDD/suicide/no ELA (N=3) samples were received at a later date for processing and sequencing, with total aligned reads equaling 46.81%.

The coding enrichment for the RPF control samples still remained below expected levels despite the variety of parameters it was examined under. Coding enrichment was calculated for 20-35 bp reads and > 8bp reads, with perfect match and primary alignment parameters both used to give 4 calculations of coding enrichment.

In the 20-35bp perfect match analysis the matched reads only reached a maximum of 10% for coding regions, with this fraction rising slightly to 14% coding regions under 20-35bp primary alignment analysis. This is substantially lower than the 41% observed in the >8bp perfect match reads and 32% observed in the >8bp primary alignment. There was no clear increase in coding enrichment between primary alignment and perfect match parameters,

but there was in in the >8bp reads compared to the 20-35bp parameter despite the fact that in successful ribosome footprinting coding regions should be primarily accounted for in the 20-25bp range (Figure 3.1A).

To reflect the likely contribution of actual RPFs to the coding enrichment without being overly stringent with the read matching, 20-35bp reads using primary alignment was selected as the coding enrichment parameter for the subsequent samples.

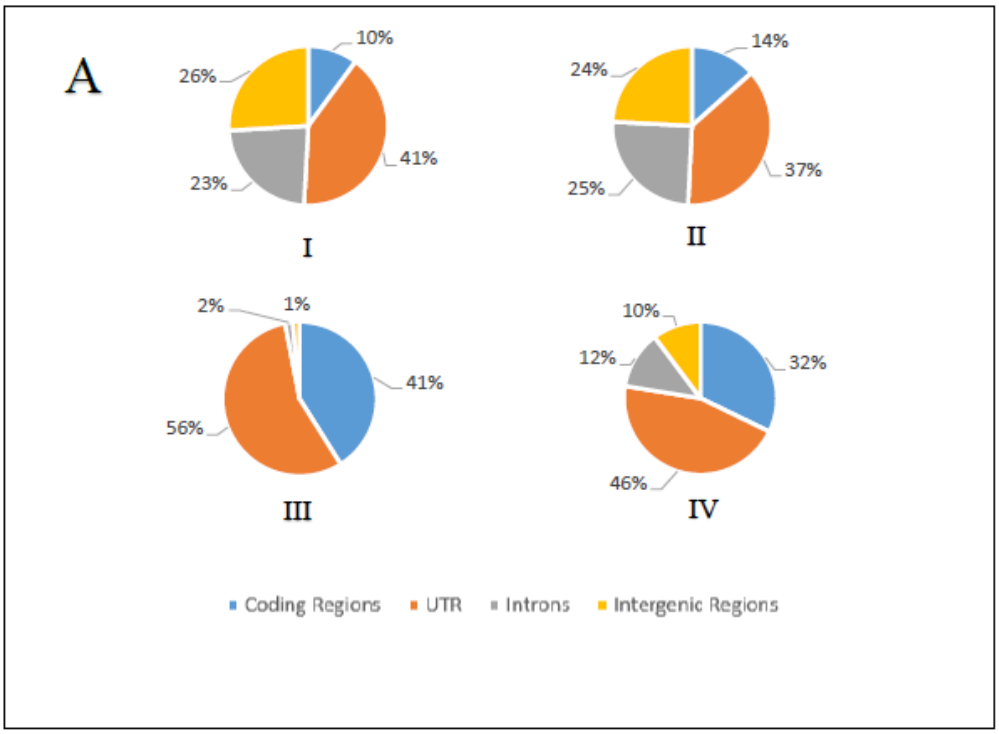
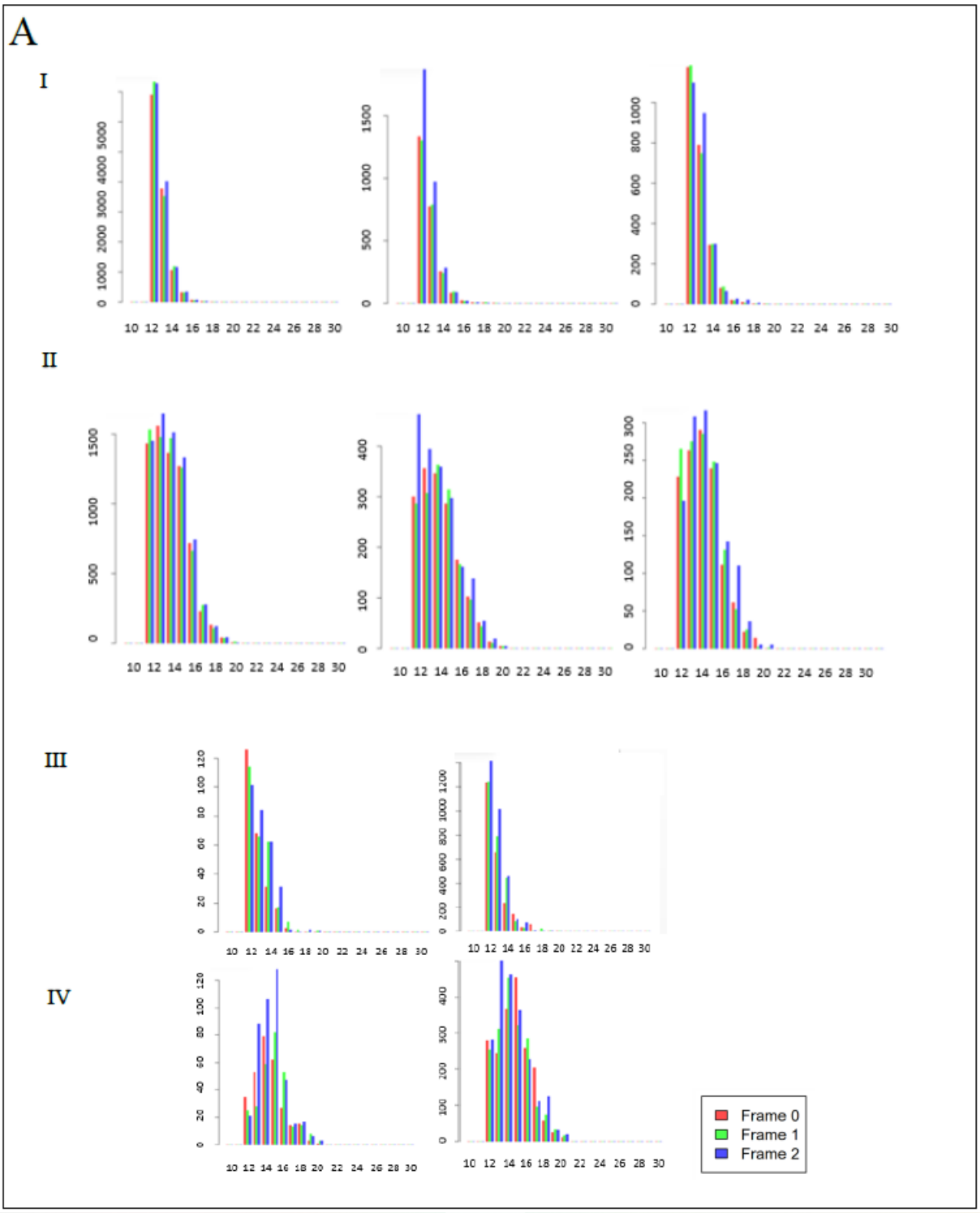


Figure 3.1 Coding enrichment profiles for SMARTer® OFC trial **A.** Average coding enrichment for control cohort (no suicide/no MDD/no ELA) samples **I**. Perfect match reads from 20-35 bp fragments **II.** Primary alignment reads from 20-35bp fragments **III.** Perfect match reads from >8 bp fragments **IV**. Primary alignment reads from >8 bp fragments.

Triplicate periodicity for the control cohort samples was conducted using both the perfect match and primary alignment parameters as well. Both parameters showed a peak shape that is characteristic of successful ribosome profiling, however, since there was a significant reduction in the number of reads aligned in the perfect match analysis, the primary alignment parameter was used in subsequent analyses (2-way repeated measures ANOVA: $F(1,4)=10.58$, $p=0.031$). In the control cohort samples, while the characteristic peak shape and a preference for a particular reading frame (reading frame 2) was observed in the samples, there were two other factors observed that are not characteristic of successful ribosome profiling. Like the previous trials, the peaks were observed in the 14-16nt range and not the expected 28-30nt range. As well, in this analysis, the total RNA samples seemed to demonstrate the characteristic RPF peak and reading frame reference as well. The MDD/suicide/ELA and MDD/suicide/no ELA samples showed a more ambiguous potential peak and reading frame reference overall, also in the ~14-16nt range, for both RPF and total RNA samples (Figure 3.2).


A
I
II
III
IV
Frame 0
Frame 1
Frame 2

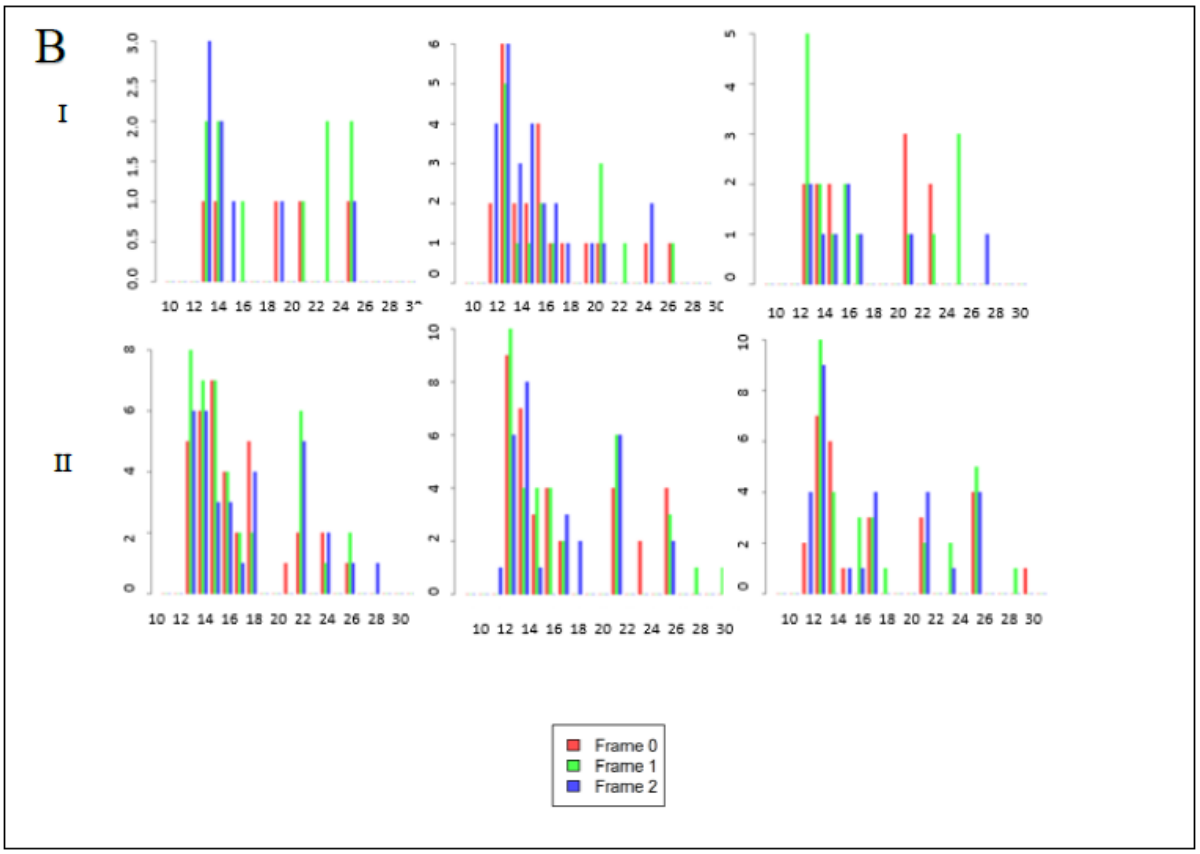


Figure 3.2 Triplicate periodicity for SMARTer® OFC trial. **A.** Triplicate periodicity for control cohort (no suicide/no MDD/no ELA) samples. **I**. Perfect match reads for RPF samples. **II**. Primary alignment reads for RPF samples. **III**. Perfect match reads for total RNA samples. **IV**. Primary alignment for total RNA samples. **B.** Triplicate periodicity for non-control cohorts (suicide/MDD/ELA and suicide/MDD/no ELA). **I**. Select RPF samples **II**. Select total RNA samples.

A PCA analysis looking at gene expression was also conducted with the control cohort samples, and revealed that the three RPF samples strongly clustered together, while the 2 total RNA samples did not cluster with either the RPF samples, or each other. This

clustering of the RPF samples and separation of total RNA samples remained true for both genome and transcriptome aligned reads (Figure 3.3).

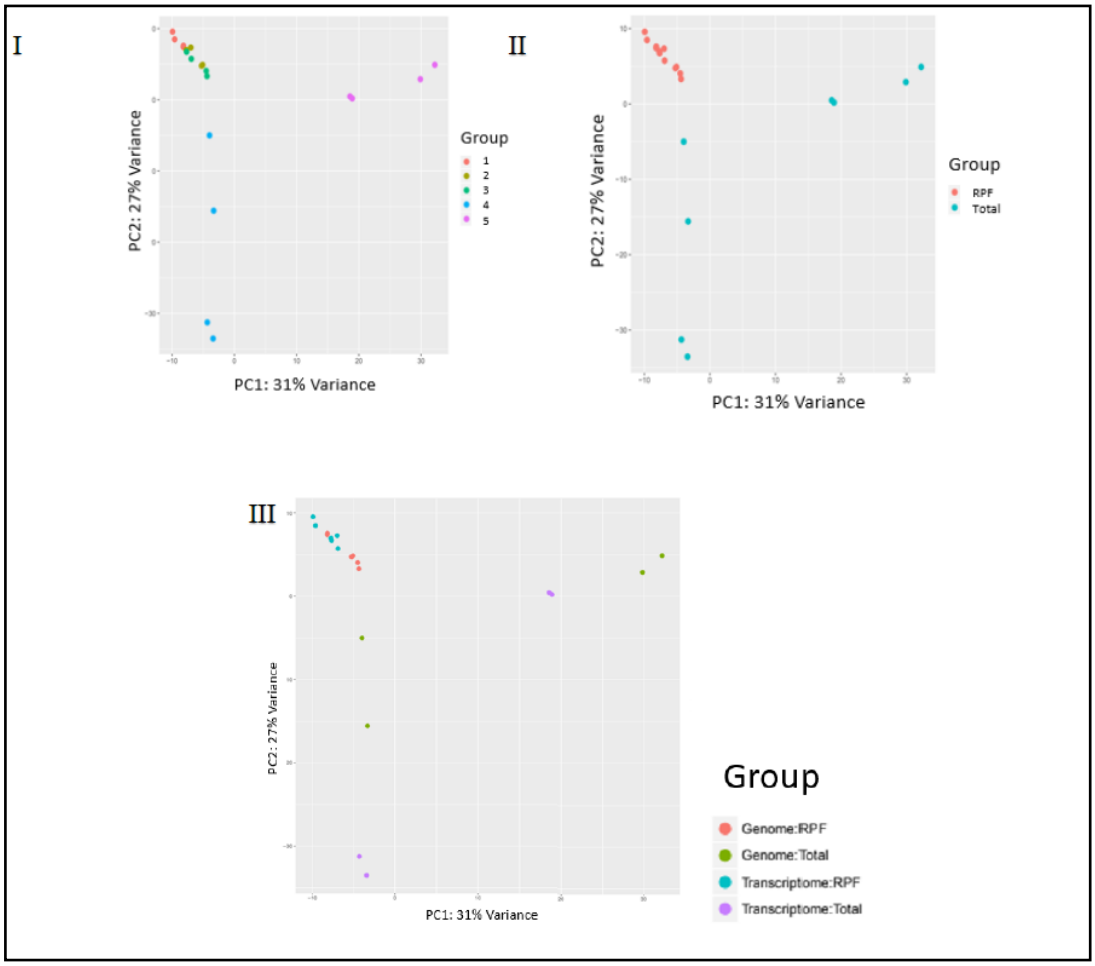


Figure 3.3 PCA analyses for the SMARTer® OFC trial control cohort samples. **I.** Data points grouped according to individual samples they derive from. **II.** Data points group by RPF versus total RNA samples **III.** Data points grouped according to whether RPF and total RNA reads aligned to the genome or transcriptome.

Distribution of fragment lengths of samples in this trial varied by sample, with some RPF samples showing predominate peaks in the 12-20bp range, and others showing a predominate peak in the expected 25-30bp range. However, in the control cohort, after manually removing and poly(A) tails and remaining adapter, both methods showed an obvious and singular peak in the 12-25bp range. The total samples in the control cohort actually showed initial peaks in the >70bp range, which disappeared after poly(A) tail and adapter trimming, leaving a new peak in the same 12-25bp range as the other samples.

3.5 Cell culture trial

Following the discontinuation of the Illumina® TruSeq Ribo Profile (Mammalian) Library Prep Kit, and the mixed sequencing results the SMARTer® smRNA-Seq Kit (Illumina®) had produced thus far in the experiment, it was proposed that multiple small RNA library construction protocols should be tried with the samples to test for efficacy of use with ribosome profiling, particularly for that of human post-mortem brain tissue. This trial was conducted with one RPF and one total RNA sample from a neuroblastoma cell culture population, and one RPF human post-mortem brain tissue sample (from the OFC). These samples were subsampled into 9 aliquots so that all 3 samples could be tried with the SMARTer® smRNA-Seq Kit for Illumina®, NEBNext® Multiplex Small RNA Library Prep Set for Illumina®, and the Galas Lab 4N RNA library preparation method (N=9). The inclusion of cell culture samples served the purpose of examining whether the method itself did not accommodate our ribosome profiling samples, or if the issue in producing adequate sequencing material was related to the use of human post-mortem brain tissue specifically.

Cell culture RPF samples were prepared according to the same procedure as previously described for the human-post mortem brain tissue, including manual douncing during the lysis stage.

Prior to small RNA library preparation, T4 PNK treatment (New England Biolabs) was performed on all samples using their Non-radioactive Phosphorylation with T4 PNK or T4 PNK protocol.

The SMARTer® smRNA-Seq Kit (Illumina®) samples were prepared as described in the previous section with the exception of final purification. Using a BluePippin system for final purification and size selection became a viable option and was successfully used for pre-sequencing purification of these samples.

The NEBNext® Multiplex Small RNA Library Prep Set for Illumina® was attempted with 3 samples: an RPF cell culture sample, a total RNA cell sample from the same source, and one RPF human post-mortem brain tissue sample (from the OFC). It was selected on the basis of being tested and appropriate for small RNA library preparation from human brain samples, as well as being intended to produce higher yields and lower adaptor-dimer contamination.

The NEBNext® Multiplex Small RNA Library (Illumina®) protocol involves the ligation of a 3’ adapter, hybridization of a reverse transcription primer to the excess of 3´ SR Adaptor that remains after the 3´ ligation reaction to create double-stranded DNAs on the 3’ end, and 5’ adapter ligation, reverse transcription, then PCR amplification. Unlike the Illumina® TruSeq Ribo Profile (Mammalian) Library Prep Kit, this methodology does not involve circularization, or the multiple PAGE purification steps. The protocol was followed

according to the manual, with the recommended 12 cycles being used during PCR amplification of human brain samples.

Final purification before sequencing was attempted via both high resolution agarose gel (with excision occurring in the ~140-150bp range) and NEBNext® AMPURE XP beads (Illumina®). Samples were evaluated for sequencing readiness via the Agilent D1000 ScreenTape System. Due to poor resolution of any final product in the expected ~140-150bp range across purification methods, this library preparation protocol was deemed unsuccessful and not sequenced. Therefore, only the SMARTer® (Illumina®) and Galas samples were sequenced and analyzed for measures of coding enrichment and triplicate periodicity.

The final protocol tried for small RNA library construction was a 4N RNA library preparation protocol developed by the Galas Lab at the Pacific Northwest Research Institute (Galas, 2016). Originally developed to construct libraries for Extracellular RNAs (exRNAs) in biofluids, this methodology is unique in that it makes use of 4N adapters- adapters with 4 random nucleotides on their ends that ligate to the small RNA fragments. It also makes use of higher than average adapter concentrations, and the addition to larger than average concentrations of polyethylene glycol (PEG) in the ligation steps. These alterations provide the benefit of reducing biases in select RNA fragment overrepresentation by driving complete ligation of the sample fragments due to increased ease of adapter annealing, increased adapter availability, and increased use of a crowding agent (PEG) to increase effective concentration of both sample and adapter to increase interaction. One consequence of this method is an increase in adapter dimers, necessitating the use of the purification of ligation products via size selection. This methodology was selected for trial to determine if

this 4N adapter use would increase the final concentration of what had been low input samples thus far.

Prior to small RNA library preparation, T4 PNK treatment (New England Biolabs) was performed on all samples using their Non-radioactive Phosphorylation with T4 PNK or T4 PNK protocol.

This protocol begins with the preparation of dried PEG strip tubes. First, 3µl of 50% PEG 8000 was added to each tube, which was then dried in a speedvac at low heat (37°C) and stored at room temperature. It is in these tubes that the 3' ligation step was completed according to protocol, followed by 5'ligation, and reverse transcription. Reverse transcription was followed according to protocol with the exception of omitting the final step that calls for 1µl of RiboShredder RNAse blend to be added to the samples and then incubated. The first PCR amplification was then performed according to protocol.

The first purification following PCR amplification is intended to be performed with either a PippinHT or BluePippin according to the original protocol. For accessibility reasons, NEBNext® AMPURE XP beads (Illumina®) were used as a substitute in the first attempt, and the procedure was finished as described in the protocol. The 10µl samples were diluted to 100µl and added to 120µl of AMPURE XP beads (Illumina®) for this purification. Tapestation visualization of these samples revealed the presence of dimers but an absence of sequencing library material.

The protocol was re-attempted from the beginning as previously described when a BluePippin became available for use. The first BluePippin gel purification was performed as described by the protocol, with adding 10µl of marker Q2 to the samples, and size

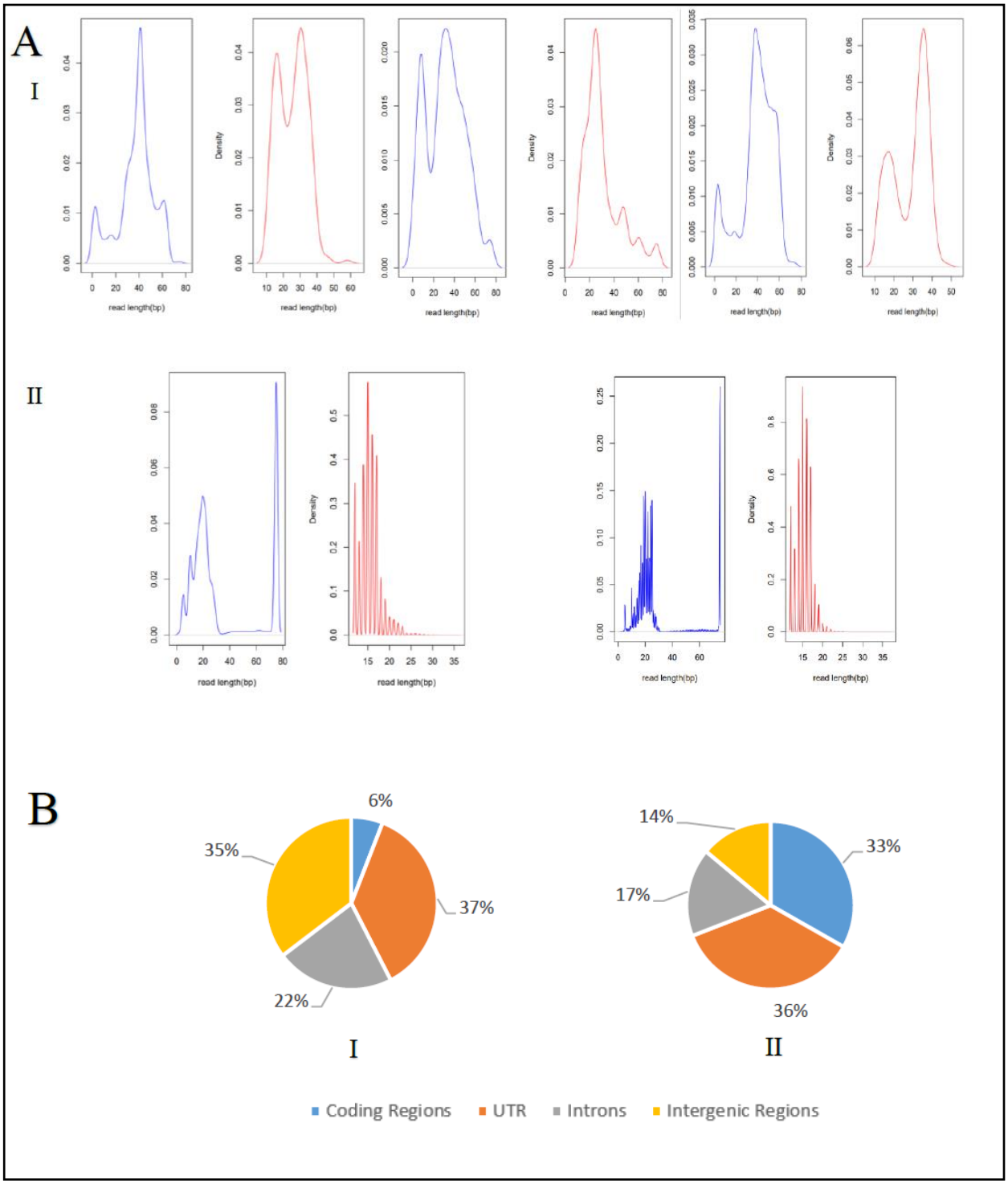
selecting in the 126-150bp range. Eluted samples were then concentrated to 22.5µl with a speedvac, and subjected to the second round of PCR amplification with the minimum suggested number of cycles (10) used. Following second amplification, BluePippin size selection purification was once again performed as described above.

Once again, only adapter dimers could be visualized after following the procedure. It was hypothesized that 5'ligation was not successfully occurring. poly(A) tails were added to the samples pre-library construction to attempt to facilitate 5'ligation. The library protocol re-started with poly(A) prepared samples this time, and followed as previously described, with the exception of performing 12 cycles in the second PCR amplification.

After the second BluePippin size selection purification was completed, the samples were visualized via tapestation, with peaks finally being present in the expected 126-151bp range.

These samples and the previously described SMARTer® smRNA-Seq Kit (Illumina®) samples were sequenced via Miseq V3 150 cycle kit.

The fragment density by read-length (bp) revealed a maximum density of fragments at ~30-40bp in all SMARTer® (Illumina®) samples, prior to read alignment, with a secondary density peak in the <20 bp range, which was typically maintained post alignment. In the Galas samples, the pre-alignment fragment density showed a peak in the <30bp range, which shifted to the <20bp range following alignment (Figure 4.1A).


A
I
II
read length(bp)
Density
B
6%
35%
37%
22%
I
14%
33%
17%
36%
II
Coding Regions
UTR
Introns
Intergenic Regions

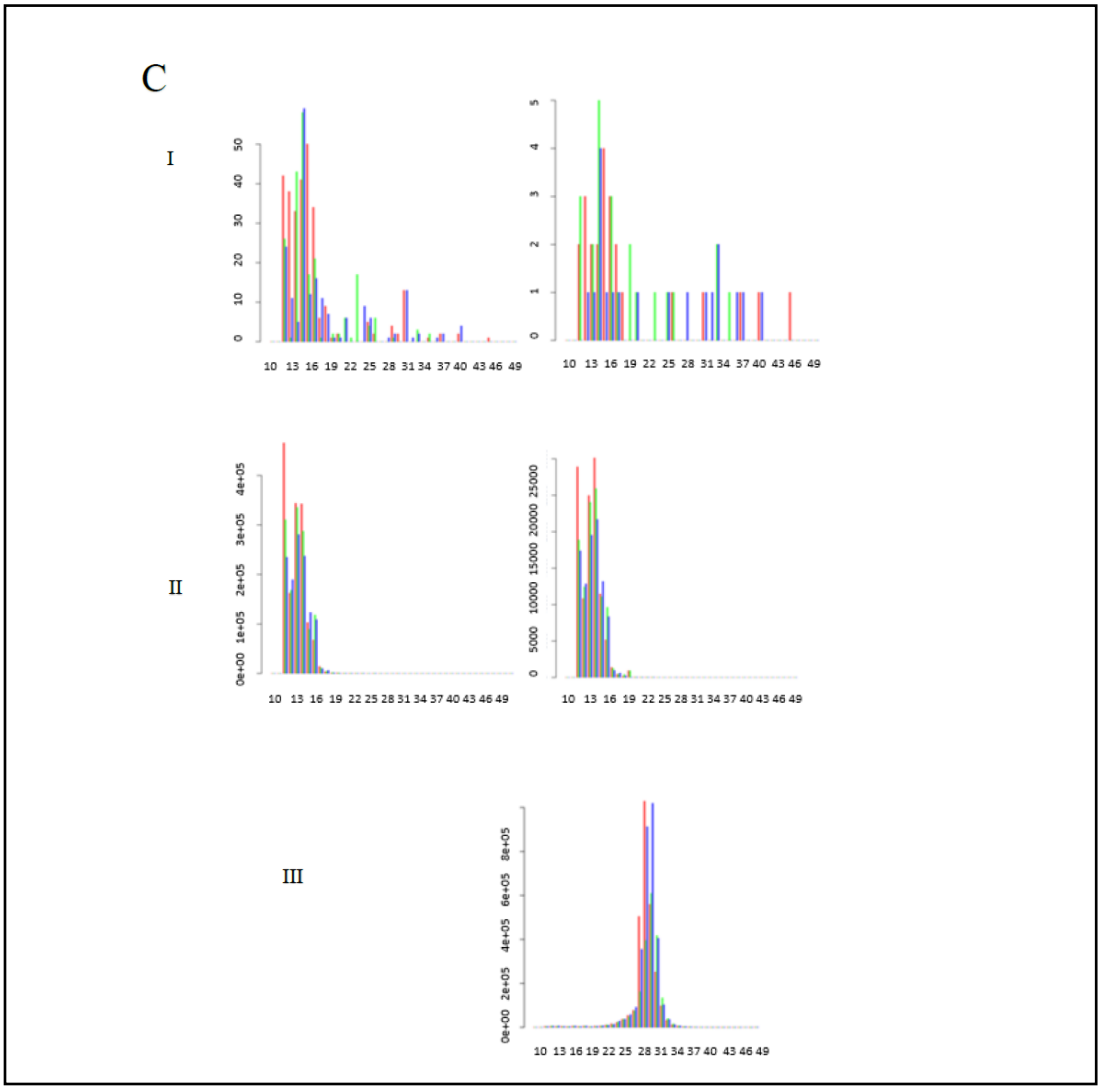


Figure 4.1 Metrics of RPF capture for the cell culture trial **A**. Fragment size distribution (density versus read length) pre-realignment (blue) and post realignment (red) **I.** SMARTer® prepared samples **II.** Galas prepared samples. **B**. Coding enrichment **I**. Average coding enrichment across SMARTer® prepared samples **II**. Average coding enrichment across Galas prepared samples **C**. Triplicate Periodicity for all aligned reads (right) and primary alignment reads (left) **I.** Select SMARTer® prepared samples **II**. Galas prepared samples **III.** Sample from the Khoutorsky lab analyzed with our pipeline.

Using the same parameters and pipeline of the previous SMARTer® (Illumina®) OFC trial, coding enrichment revealed a continued lower than expected coding enrichment. For the SMARTer® (Illumina®) samples, the average reads were 6% coding regions, 37% UTR, 22% introns, and 35% intergenic regions. Whereas in the Galas samples, even though coding enrichment was still below expected it was substantially higher than what was observed in the SMARTer® (Illumina®) samples. The average reads were 33% coding regions, 36%UTR, 17% introns, and 14% intergenic regions (Table 2, Fig 4.1B).

Table 2. Cell culture trial SMARTer® versus Galas prepared samples total reads aligned to each region type.

	Coding Regions	**UTRs**	**Introns**	**Intergenic Regions**
SMARTer Sample 1	302	3041	1274	2089
SMARTer Sample 2	1667	8461	3348	2938
SMARTer Sample 3	843	5855	5928	11770
Galas Sample 1	2832379	3655125	1358498	1144475

Triplicate periodicity was performed twice for each sample: one accounting for all alignments, and the other using only primary alignments. In the all alignment parameter, 2 of the 3 SMARTer® (Illumina®) samples showed the characteristic RPF peak (however, it was still in the lower than expected ~13-16nt range), and in the primary alignment analysis only 1 of the samples did (and still in the ~13-16nt range). In the Galas samples, both alignment parameters produced the characteristic peak with a clear reading frame preference, but it was also in the ~13-16nt range (Figure 4.1C).

Data from a previous mammalian brain tissue ribosome profiling study was analyzed through our triplicate periodicity pipeline to identify if the process was responsible for the lower than expected nt peaks (Uttam et al., 2018). Triplicate periodicity from this data was as expected with the characteristic reading frame preference and peak in the ~28-31nt range (Figure 4.1C).

Given the substantially higher coding enrichment, number of aligned reads, and more characteristic triplicate periodicity produced by the Galas protocol, it was selected for subsequent trials (Table 2, Figure 4.1A, Figure 4.1B).

The coding enrichment for the Galas protocol samples did vary when subjected to 3 different novel coding enrichment programs. PICARD and RseQC both calculated total reads for Coding regions, UTRs, introns, and intergenic regions, like the previous coding enrichment protocol (although the RseQC protocol divides the UTRs into 5' and 3' UTR exons). The BROAD Institute protocol calculated percent reads attributable to each category, which were: exons, introns, intergenic regions, and then it also gives an overall intragenic region rate. The PICARD analysis described the percentage of aligned reads as 35% coding regions, 32% UTR, 18% introns, and 15% intergenic regions. The RseQC analysis identified reads as being 37% coding regions, 9% 5' UTR, 11% 3'UTR, 24% introns, and 19% intergenic regions. The BROAD Institute analysis showed 65% exons, 21% introns, and 14% intergenic regions, with an overall intragenic rate of 81% (Figure 5.2).

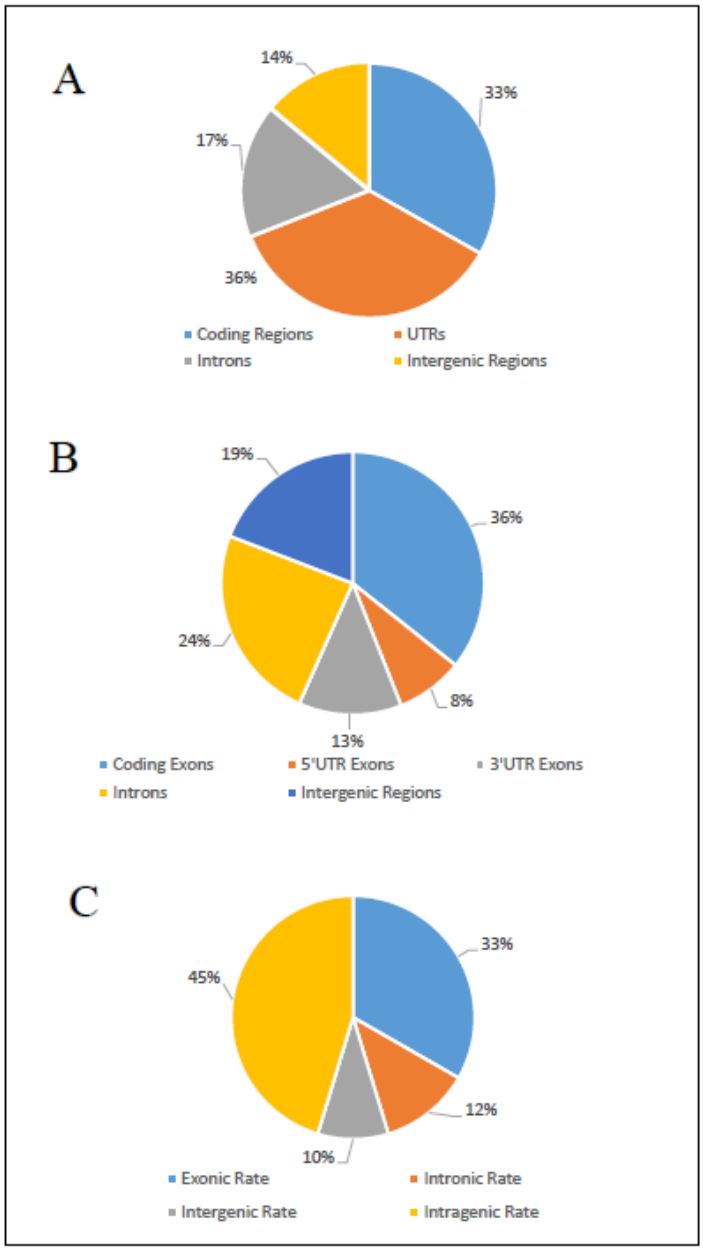


Figure 4.2 Average coding enrichment for Galas prepared samples from the cell culture trial. **A.** Prepared with PICARD **B**. Prepared with RseQC **C.** Prepared with BROAD Institute procedure.

3.6 Galas Lab 4N RNA library preparation OFC trial

After completion of the cell culture trial, one final trial was attempted with the Galas Lab 4N RNA Library Preparation method. Using human OFC tissue from the control (No MDD/suicide/ELA) (N=3), MDD/suicide/ELA (N=3), and MDD/suicide/no ELA groups (N=4).

Library construction was conducted as previously described for this protocol, with the exception of 16 cycles being used in the second round of PCR amplification. Two of the samples did not acceptably amplify for sequencing, so 1 identically prepared cell culture sample, and 1 identically prepared mouse brain sample were prepared to be sequenced with the two remaining OFC samples. Libraries were not visible following the second round of BluePippin size selection and were not sequenced.

3.7 Mouse PMI trial

A smaller sub-study was conducted concurrently to the SMARTer® smRNA-Seq Kit (Illumina®) human post-mortem brain tissue trial. Its purpose was to determine the effect of PMI on the ribosome profiling process using a PMI controlled mouse model. Mice were divided into the following experimental groups: control (immediately flash frozen), and then 4 experimental conditions which included spending 0,3,6, and 12 hours respectively at room temperature before being refrigerated for 24 hours and then flash frozen. Frontal lobe tissue was used in this trial, as it is most analogous to the human brain tissue regions that were also used throughout the larger study. An equal number of samples from male and female mice were used (N=32).

Lysis and RPF footprinting and purification were performed as described previously, and library preparation for both RPF and total RNA aliquots for each sample was performed according to previously described optimized SMARTer® smRNA-Seq Kit (Illumina®) protocol, including subsequent P5P7 additional amplification and high resolution agarose gel final size selection purification (this sub-study was conducted prior to BluePippin accessibility). Two samples did not adequately amplify for sequencing and both RPF and RNA samples were discarded (final N=60). Samples were sequenced via Miseq V3 150 cycle kit.

Measures of coding enrichment and triplicate periodicity did not reveal significant differences between any of the PMI conditions. Across RPF samples, we looked at the effect of group as a fixed factor, and within coding enrichment there was a significant trend toward significant effect of PMI (1-way ANOVA: $F(4,24)=2.135$, $p=0.108$). Average coding region rate for the RPF samples for all of the PMI conditions (when reads from when all fragment sizes were included) ranged from 4-7%, substantially below the expected ~66% for RPF samples. Reads attributed to UTRs, introns, and intergenic regions did not substantially differ across conditions either (Figure 5A). Contrary to the expectations of successful ribosome profiling, this low coding enrichment actually decreased when only >20bp fragments were taken into account. The total RNA samples' coding enrichments showed remarkable similarity to the RPF samples, and to each other, with an average of 5% of reads matching to coding regions. Just as with the RPF samples, the rate of reads aligning to coding regions decreased when only >20 bp fragments were analysed (Figure 5A).

Triplicate periodicity profiles generally revealed characteristic triplicate periodicity peak with a likely reading frame preference, whether in the RPF or total RNA samples. Like

other trials performed, and consistent with the >20nt fragment analysis, the peaks were observed outside the expected ~30nt range at ~12-16nt (Figure 5B).

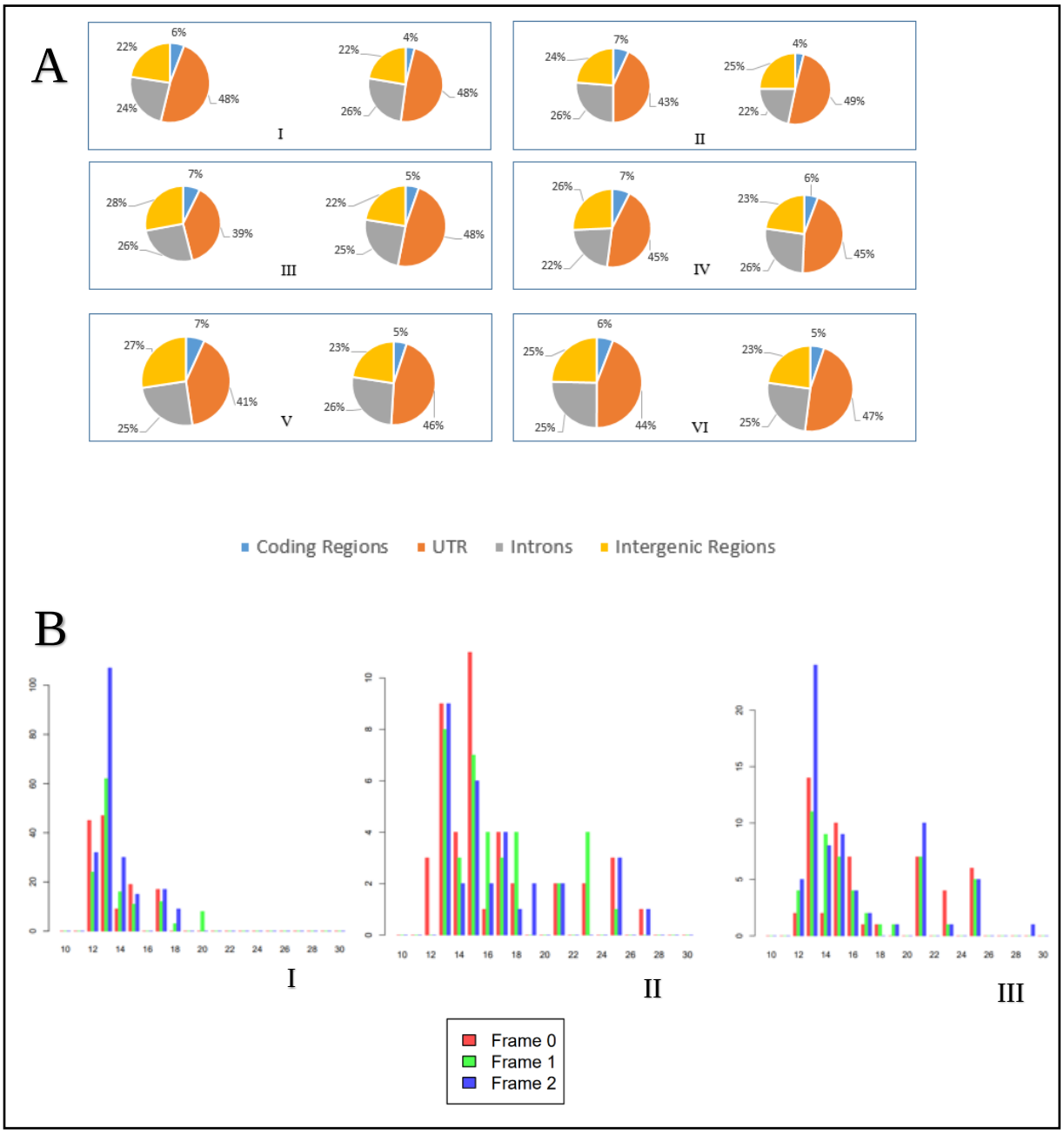


Figure 5. Mouse PMI trial metrics of RPF capture **A.** Average coding enrichment per cohort, RPF samples (left), and total RNA samples (right) **I.** Control cohort **II.** 0-hour samples **III.** 3-hour samples **IV**. 6-hour samples **V.** 12-hour samples **VI.** Average coding enrichment across all cohorts **B.** Triplicate periodicity **I.** 0-hour RPF sample **II.** 3-hour RPF sample **III.** 12 hour total RNA sample.

Chapter 4. Discussion

Ultimately, the aim to create a methodology for the ribosome profiling of human post-mortem brain tissue was not successful within the parameters of this study. There was also difficulty in performing successful ribosome profiling to different degrees across tissue types, PMIs, experimental protocol changes, and bioinformatics pipeline changes.

One possible explanation for a lack of significant changes in coding enrichment across PMIs in the mouse PMI trial may be that the tissue available did not adequately reflect the point at which PMI becomes a significant factor in monosome and RPF degradation. The maximum PMI obtained in this trial, 12 hours, is also significantly lower than the average PMI found in the human post-mortem brain tissue used in this study, which ranged from 18-64 hours. However, if the experimental and bioinformatic pipeline had been completed successfully, there should have still been a substantial coding enrichment (~66%) in the control (flash frozen) cohort regardless of effect of various PMIs. When data from Uttam et al. was subjected to our coding enrichment and triplicate periodicity analysis, it produced expected coding enrichment and triplicate periodicity profiles (Uttam et al., 2018) (Figure 4.1C). This data included a Truseq (Illumina®) adapter trimming sequence, which our protocol has not contained since the discontinuation of the Illumina® TruSeq Riboprofile (Mammalian) Kit. Given this, it is likely that any factors responsible for the results we received are from the experimental protocol or any step in the bioinformatics pipeline that precedes coding enrichment and triplicate periodicity analysis. Considering that our coding enrichment and triplicate periodicity never reached ideal profiles when using the Truseq kit (Illumina®), and following its exact pipeline protocol, it is doubtful that this step is the

source of failure of the protocol. In order to speculate on this, each major step of the methodology should be considered.

In performing cell lysis and ribosome footprinting, success at this stage was measured in the capture of predominately ~30nt fragments in the RPF samples, but not the total RNA samples. After Bioanalyzer (Agilent) analysis of fragment sizes was performed for the RNAse conditions following ribosome footprinting, this expected pattern was found (Figure 1.1). This likely indicates that there were RPF fragments successfully isolated at this point in the procedure, and that issues with the feasibility of the procedure occurred later in the protocol.

The TruSeq Riboprofile (Mammalian) Kit (Illumina®) methodology bears remarkable similarity to the original Ingolia protocol for the ribosome profiling of mammalian cells in both experimental procedure and bioinformatic pipeline (Ingolia et al., 2012). This is why, when the original two trials that used the TruSeq kit (Illumina®)- the initial BA 8/9 trial and the RNAse conditions trial- did not show a sufficiently large coding enrichment in the RPF samples, the expected triplicate periodicity profile in the expected ~30nt fragments, and the RNAse trial samples failed to amplify adequately for sequencing at all, it was determined that the experimental procedure defined by the Ingolia lab and the TruSeq Riboprofile (Mammalian) Kit (Illumina®) would not be sufficient of ribosome profiling of human post-mortem brain tissue (Figure 2.2).

Were the study re-attempted, a trial involving processing cell culture, mouse brain, and human brain tissue samples together using the Ingolia method (as the Kit itself was discontinued) would be of use. Since this methodology has been reliably used in mammalian cells and homogenized mouse brain tissue, failure of all samples to sequence

and produce expected coding enrichment and triplicate periodicity profiles could clearly indicate sources of incompatibility in the overall protocol workflow as responsible for not developing a feasible human post-mortem brain tissue ribosome profiling procedure (Ingolia et al., 2012; Uttam et al., 2018). As it stands, it is possible that the attempted combination of differently sourced lysis and ribosome footprinting, small RNA library construction, amplified fragment purification, and bioinformatic methods played a role in the cell culture and mouse brain samples prepared in later trials failing to show the hallmarks of successful ribosome profiling.

Of the subsequently tried small RNA library preparation methods, the SMARTer® smRNA-Seq Kit (Illumina®), and the Galas Lab 4N RNA Library Preparation protocol were the only ones that adequately amplified for sequencing. In each case, either high resolution agarose gel or BluePippin trimming revealed the presence of bands or peaks in the expected range for libraries (~183bp for the SMARTer® (Illumina®) libraries, and ~126-151bp for the Galas libraries), and not just the adapter range. This would suggest success with small RNA library construction up until this point, although that is not ensured by these results without performing read alignment, coding enrichment, and triplicate periodicity analysis post-sequencing (Galas, 2018).

The SMARTer® smRNA-Seq Kit (Illumina®) has been previously used, to success, on mammalian (mouse) brain tissue, and so should have produced ideal coding enrichment and triplicate periodicity profiles (as well as higher read counts) in at least the mouse PMI trial (Hornstein et al., 2016) (Figure 5). This does suggest a source of error in either the experimental protocol or the bioinformatic pipeline (perhaps in fragment trimming) that occurs before these analyses. The Hornstein study specifies using AMPure XP beads

(Illumina®) for final library purification, which we had previously found during our human post-mortem brain trials and the cell culture trial (which included a human brain RPF sample for comparison) results in an excess loss of material. It is possible that this method of purification, while successful in mouse brain studies, was not feasible for our end of goal of using human-post mortem brain tissue.

However, it is also possible that there was material lost or imprecision in excising the amplified fragments from the high resolution agarose purification. Since SMARTer® (Illumina®) libraries are present 153bp + the length of the library fragments (~30nt in the case of RPFs), the ideal length for excision would be ~183bp. However, to account for the slight degree of variance commonly seen in RPFs (~26-30nt), the possibility that the RPFs were over digested on each side on the monosome during footprinting, and to excise all of the top bands observed above the lower bands in the expected adapter range, we excised at a more liberal ~1650-190bp range, which was still ~10bp above the adapter dimer range. It is possible that this was not an adequate distance between the dimer range and expected library fragments, leading to contamination and low representation of coding fragments in our samples. The fact that the highest density of fragment lengths in the OFC trial were in the expected ~25-30bp range before poly(A) tail and adapter trimming- but reduced to the ~12-25bp range afterwards corroborates this reasoning. It is also possible that by starting excision only 12bp above 153bp, that fragments already in the ~12-25bp range were selected for sequencing and further analysis which could have represented over-digested RPF fragments (which would make coding region alignment more difficult). It is also possible that adapter annealing to non-RPF smaller RNAs that were not adequately filtered out of the samples during earlier purification occurred.

Something else of note is that while human imprecision in high resolution agarose excision may have contributed to errors in successful ribosome profiling across SMARTer® (Illumina®) trials (including the mouse PMI trial), these considerations do not apply to the BluePippin Purification method used in the Galas library preparation trials where fragment selection is done by the BluePippin machine without expanding the base pair range for purification. So, while it can be reasonably assumed that this step is not a source of error in the Galas cell culture trial (through the input for this step may have been contaminated or missing the RPF content from previous steps in the protocol), it is worth noting that after final BluePippin purification in the Galas OFC trial, there was insufficient material still present for sequencing. This contributes to the overall trend that the majority of trials performed suffered from an inadequate concentration, and likely quality, of usable RPF material.

This is evident in analyzing coding enrichment across ribosome profiling trials. In these trials, aligned reads corresponding to coding regions varied from 1-37% in RPF samples (and an overall exon rate 65% percent in the BROAD institute method applied to the Galas cell culture trial samples), which even at the highest is still substantially below the expected 65.7% percent for size-exclusion chromatography (s-400 spin columns specifically in these trials), as outlined in the Illumina® TruSeq Ribo Profile (Mammalian) Kit protocol (Figure1.2B, Figure 2.2A, Figure 3.1, Figure 4.1B, Figure 5.2). It cannot be assumed that the coding enrichment rate in the Galas cell culture trial is close to this rate, as UTRs would be included in the 65% exon rate, and UTRs made up approximately half of this percentage in the other coding enrichment analyses on these samples.

There are a number of potential explanations for these results including: either overzealous or imperfect adapter and poly(A) tail trimming early on in the bioinformatics pipeline before the coding enrichment analysis is performed, adapter ligation to non-RPF small RNA debris not properly filtered out of the samples during monosome purification, imprecise excision of Urea PAGE gels resulting in the loss of RPF material or contamination of non-RPF material, imprecise excision of high resolution agarose gels where this was used as the final library purification method, and potentially incompatibility of certain procedures when combined with each other for trials after the TruSeq kit (Illumina®) was discontinued.

It is also worth noting that while none of the trials reached the expected level of coding enrichment, there was still notable variance between library preparation methods. The coding region rate was 20.17% for the Truseq kit (Illumina®), 1-7% across trials using the SMARTer® kit (Illumina®), and 33-37% across trials using the Galas lab library construction method (and 65% exons using the BROAD Institute coding enrichment protocol) (Figure1.2B, Figure 2.2A, Figure 3.1, Figure 4.1B, Figure 4.2). This suggests that out of the library construction methodologies that amplified enough to be sequenced, the SMARTer® (Illumina®) protocol was the least successful in our trials at capturing RPFs, which is unexpected given that the SMARTer® (Illumina®) protocol has been successfully used to create RPF libraries of mammalian brain tissue previously, and the Galas method has not been published as doing so, and was created to construct sequencing libraries for extracellular RNAs (Galas, 2018; Hornstein et al., 2016).

In the expected coding enrichment profile, UTRs are supposed to represent the second highest fraction of aligned reads (at ~22%), followed by much lower rates of introns and intergenic regions (Figure 1.2A). Even with a reduced rate of coding enrichment, this

comparative relationship was not consistently upheld across trials with the highest rate of aligned reads across trials being: intergenic regions in the TruSeq (Illumina®) BA 8/9 trial, Introns in the SMARTer® (Illumina®) RNAse trial, UTRs in the SMARTer® (Illumina®) OFC trial (with still a high percent of introns and intergenic regions), and finally UTRs with a lower rate on introns and intergenic regions in the Galas cell culture trial across coding enrichment analysis methods. This final trial was the closest to the expected profile, which may be due in part to alterations made to the bioinformatic pipeline (including sequence trimming, perfect match or primary alignment selection, and coding enrichment program selection), but may also reflect greater success at capturing RPFs with the Galas library construction method (Figure 4.1B, Figure 4.2).

Overall, total RNA coding enrichment profiles also often differed from the expected profile. While in the original Truseq (Illumina®) BA 8/9 trial the coding enrichment profile did not differ substantially from the expected profile, this was not the case with the subsequent trials, where there was no clear difference observed between the coding enrichment profiles of RPF and Total RNA samples (Figure1.2B, Figure 2.2A, Figure 3.1, Figure 4.1B, Figure 4.2). The cause of this has not been definitively identified, but could contradict the possibility that the coding profile found in the Galas cell culture trial represents an improvement from previous trials in analysing coding enrichment, or that our fragments even represent RPFs and total RNA from the samples we were hoping to obtain.

This similarity of the RPFs and total RNA samples was also found in triplicate periodicity profiles across trials. In the original Truseq (Illumina®) BA 8/9 trial it was the RPF sample that resembled the expected total RNA profile, and in all subsequent trials both the RPF and total RNA samples typically showed the expected peak shape and reading frame reference

that defines successful RPF capture. In all of these cases, this peak was observed in the <20nt range, which while consistent with our results of fragment length density, is inconsistent with the hallmarks of successful ribosome profiling (Figure 2.2B, Figure 3.2, Figure 4.1C, Figure 5B). As mentioned when discussing fragment length density, this may be due to over-trimming of adapters, degradation of the monosomes and underlying RPFs during ribosome footprinting resulting in shorter than expected fragments (although this would be contradicted by the RNAse trial results from bioanalyzer visualization showing ~30nt fragments), imprecise excision in one of the purification steps resulting in loss of desired material, or adapter ligation to unfiltered smaller debris RNAs in the samples.

In light of the fragment distribution, coding enrichment, and triplicate periodicity, and total read counts, we likely did not achieve sufficient proof or quantity of successful RPF capture to continue with deeper DTG/DEG analysis as originally intended at the outset of this study.

While not sufficient for deeper DTG/DEG analysis, or to establish a replicable methodology, it could be argued that some degree of RPF capture may have been achieved in this study. The fragment length distribution observed via bioanalyzer after lysis and ribosome footprinting -but before sequencing- indicate that the lysis and RNAse conditions (when compared to total RNA controls) used in this study likely reflect a feasible procedure for capturing RPF fragments in human post-mortem brain tissue pre-cDNA library construction.

As well, while the expected coding enrichments and the triplicate periodicity peak ranges were lower than expected throughout the study, there was still a coding enrichment of 20-37% and triplicate periodicity peaks achieved. While issues in these analyses was still observed when applied to cell culture and mouse brain samples, suggesting a source of error

in either the experimental procedure and/or the bioinformatic pipeline preceding these analyses, it is possible that even if these factors were identified and corrected, coding enrichment and fragment length in human post-mortem brain tissue RPFs might still be lower than expected due to degradation and fragility of the tissue. Despite ultimately failing to produce a feasible and complete procedure for ribosome profiling in human post-mortem brain tissue in this study, there is evidence that with correction of the unidentified sources of error, creation of such a procedure is still possible.

In addition to the possible factors and sources of error already mentioned, it is plausible that the very nature of the attempted trial methodologies (following the initial TruSeq (Illumina®) BA 8/9 trial) being constructed of lysis, ribosome footprinting, monosome purification, cDNA library construction, library purification, and bioinformatic pipeline protocols from different sources or created protocols led to some incompatibility or error not caught in visualization check-points pre-sequencing. This could help explain why our cell culture and mouse PMI samples suffered similar inconsistencies with successful ribosome profiling as the human post-mortem brain samples across trials, but the data produced by Uttam et al. via the Truseq kit (Illumina®) and input into our coding enrichment and triplicate periodicity protocols did not (Uttam et al., 2018).

Chapter 5. Conclusions and future directions

Since the onset of this study, ribosome profiling has continued to be expanded in exploring translation in the human brain. One recent study, which performed ribosome profiling on striatal cell cultures affected by Huntington disease (HD), found evidence that the polyglutamine expansion of huntingtin (mHTT) impairs ribosomal translocation during translation elongation (Eshraghi et al., 2021). This is a mechanistic defect that could be explored in future clinical research into HD therapeutics.

While complete optimization of such a procedure was not attained in the scope of this work, achieving this aim in potential future research could have significance in a variety of ways. The successful creation of a human post-mortem brain tissue ribosome profiling procedure- if achieved through further experimentation- has the potential to be utilized as a novel methodology for the study of psychiatric pathologies such as MDD, bipolar disorder, PTSD, and schizophrenia. As well, it could be potentially used to examine alterations in translation in individuals who die by suicide: which has an approximately 60% comorbidity with MDD (Cavanagh, Carson, Sharpe, & Lawrie, 2003). Applied applications of such work could potentially include the identification of key DTGs which may have future potential to be utilized as target genes in clinical research into antidepressant therapies, therapies for other psychopathologies, or as risk biomarkers for suicide and any of the above mentioned psychopathologies. Outside of the direct scope of psychopathology, this methodology could also be used to explore DTGs across different brain regions.

Given advancements such as these, it is possible that mapping the translatome of the human brain could be a reality in the near future. The field of psychiatric epigenetics could be

poised to soon fill in a substantial step in their epigenetic alteration-to-psychopathy pathways, with a major methodological breakthrough on the horizon.

Supplemental Tables

Table 1. Deviations from the original Ingolia (2012) Ribosome profiling protocol paper, as demonstrated by a selection of ribosome profiling studies.

Phase of Procedure	**Method Used**	**Study**
Cell Lysis	Original	Calviello 2015 Ingolia 2011 Jang 2015 Uttam 2018
	Same buffer, altered centrifugation time	Gerashchenko 2017 Chew 2013
	Different buffer and centrifuge time	Amorim 2018 Bazzini 2012 Bazzini 2014 Hornstein 2016 Zhang 2017 Guo 2010 Hsu 2016 Jenson 2014
	Different buffer, same centrifuge time	Fisunov 2017 Dunn 2013 Liu 2019
Footprinting	RNase I for 45 min (Original)	Amorim 2018 Bazzini 2012 Bazzini 2014 Hornstein 2016 Zhang 2017 Uttam 2018 Jang 2015 Ingolia 2011 Calviello 2015

Table 1. Deviations from the original Ingolia (2012) Ribosome profiling protocol paper, as demonstrated by a selection of ribosome profiling studies, continued.

Phase of Procedure	Method Used	Study
	RNAse I for 1 hour	Hsu 2016
		Jenson 2014
	RNase I, S7, T1, A	Gerashchenko 2017
	RNase A + T1	Liu 2019
	RNase not Listed	Fisunov 2017
Footprinting Fragment Purification	15% TBE Urea Gel	Jang 2015
		Ingolia 2011
		Amorim 2018
		Bazzini 2012
		Bazzini 2014
		Hsu 2016
		Jenson 2014
		Dunn 2013
	10% TBE Urea Gel	Uttam 2018
		Chew 2013
		Guo 2010
	17% TBE Urea Gel	Calviello 2015
	Phenol Chloroform	Zhang 2017
	Not Specified	Gerashchenko 2017
		Hornstein 2016
		Fisunov 2017
		Liu 2019

Table 1. Deviations from the original Ingolia (2012) Ribosome profiling protocol paper, as demonstrated by a selection of ribosome profiling studies, continued.

Phase of Procedure	Method Used	Study
Library Preparation	Original	Jang 2015
		Ingolia 2011
		Uttam 2018
		Chew 2013
		Gerashchenko 2017
		Jenson 2014
		Liu 2019
	Different reverse transcription with no circularization	Bazzini 2012
	Alternative pre-existing RNA Library preparation method	Calviello 2015
		Guo 2010
	SMARTer	Hornstein 2016
	ARTseq Ribosome Profiling Kit, mammalian	Bazzini 2014
		Zhang 2017
	ARTseq/TruSeq Ribo Profile Kit (nonmammalian)	Hsu 2016
	SOLiD library preparation	Fisunov 2017
	Different RT but Same Circuligase as Ingolia	Dunn 2013
	Not specified	Amorim 2018
rRNA Depletion	Subtractive hybridization	Ingolia 2011
		Dunn 2013
	Bowtie	Calviello
		Uttam 2018
		Amorim 2018
		Fisunov 2017
	SILVA database	Chew 2013
	Ribo-Zero™	Bazzini 2014
		Zhang 2017
		Hsu 2016
		Liu 2019
	Subtraction oligo pool	Hornstein 2016
	Additional DNA amplification	Guo 2010
	Not specified	Jang 2015
		Bazzini 2012
		Jenson 2014

Table 2. Select ribosome profiling studies from chapter 1 by sample type.

Species	**Tissue type**	**Study**
M. musculus	Brain	Amorim 2018 Hornstein 2016 Liu 2019
	Embryonic stem cells	Chew 2013 Ingolia 2011
	Dorsal root ganglia and spinal cord	Uttam 2018
	Liver	Gerashchenko 2017
C. elegans	Whole worm lysate	Gerashchenko 2017
D. melanogaster	Embryos	Gerashchenko 2017 Dunn 2013
Arabidopsis	Plant roots and shoots from seedlings	Hsu 2016
D. Rerio	Zebrafish Embryo	Bazzini 2012 Bazzini 2014 Chew 2013
H. Sapiens	HEK293 cells HeLa cells U2OS cells	Calviello 2015 Guo 2010 Jang 2015
E. Coli	Bacterial cell culture	Gerashchenko 2017 Zhang 2017
M. gallisepticum	Bacterial cell culture	Fisunov 2017
S. cerevisiae	Yeast cell culture	Gerashchenko 2017
T. brucei	Whole parasite lysate	Jensen 2014

References

Abdolmaleky, H. M., Yaqubi, S., Papageorgis, P., Lambert, A. W., Ozturk, S., Sivaraman, V., & Thiagalingam, S. (2011). Epigenetic dysregulation of HTR2A in the brain of patients with schizophrenia and bipolar disorder. *Schizophr Res, 129*(2-3), 183-190. doi:10.1016/j.schres.2011.04.007

Amorim, I. S., Kedia, S., Kouloulia, S., Simbriger, K., Gantois, I., Jafarnejad, S. M., . . . Gkogkas, C. G. (2018). Loss of eIF4E Phosphorylation Engenders Depression-like Behaviors via Selective mRNA Translation. *J Neurosci, 38*(8), 2118-2133. doi:10.1523/JNEUROSCI.2673-17.2018

Asante, C. O., Wallace, V. C., & Dickenson, A. H. (2010). Mammalian target of rapamycin signaling in the spinal cord is required for neuronal plasticity and behavioral hypersensitivity associated with neuropathy in the rat. *J Pain, 11*(12), 1356-1367. doi:10.1016/j.jpain.2010.03.013

Bartholomaus, A., Del Campo, C., & Ignatova, Z. (2016). Mapping the non-standardized biases of ribosome profiling. *Biol Chem, 397*(1), 23-35. doi:10.1515/hsz-2015-0197

Bazzini, A. A., Lee, M. T., & Giraldez, A. J. (2012). Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. *Science, 336*(6078), 233-237. doi:10.1126/science.1215704

Borsani, G., Tonlorenzi, R., Simmler, M. C., Dandolo, L., Arnaud, D., Capra, V., . . . Ballabio, A. (1991). Characterization of a murine gene expressed from the inactive X chromosome. *Nature, 351*(6324), 325-329. doi:10.1038/351325a0

Bremner, J. D., Vythilingam, M., Vermetten, E., Nazeer, A., Adil, J., Khan, S., . . . Charney, D. S. (2002). Reduced volume of orbitofrontal cortex in major depression. *Biol Psychiatry, 51*(4), 273-279. doi:10.1016/s0006-3223(01)01336-1

Calviello, L., Mukherjee, N., Wyler, E., Zauber, H., Hirsekorn, A., Selbach, M., . . . Ohler, U. (2016). Detecting actively translated open reading frames in ribosome profiling data. *Nat Methods, 13*(2), 165-170. doi:10.1038/nmeth.3688

Cavanagh, J. T., Carson, A. J., Sharpe, M., & Lawrie, S. M. (2003). Psychological autopsy studies of suicide: a systematic review. *Psychol Med, 33*(3), 395-405. doi:10.1017/s0033291702006943

Chew, G. L., Pauli, A., Rinn, J. L., Regev, A., Schier, A. F., & Valen, E. (2013). Ribosome profiling reveals resemblance between long non-coding RNAs and 5' leaders of coding RNAs. *Development, 140*(13), 2828-2834. doi:10.1242/dev.098343

Cho, J., Yu, N. K., Choi, J. H., Sim, S. E., Kang, S. J., Kwak, C., . . . Kaang, B. K. (2015). Multiple repressive mechanisms in the hippocampus during memory formation. *Science, 350*(6256), 82-87. doi:10.1126/science.aac7368

Dunn, J. G., Foo, C. K., Belletier, N. G., Gavis, E. R., & Weissman, J. S. (2013). Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster. *Elife, 2*, e01179. doi:10.7554/eLife.01179

Eshraghi, M., Karunadharma, P. P., Blin, J., Shahani, N., Ricci, E. P., Michel, A., . . . Subramaniam, S. (2021). Mutant Huntingtin stalls ribosomes and represses protein synthesis in a cellular model of Huntington disease. *Nat Commun, 12*(1), 1461. doi:10.1038/s41467-021-21637-y

Fields, A. P., Rodriguez, E. H., Jovanovic, M., Stern-Ginossar, N., Haas, B. J., Mertins, P., . . . Weissman, J. S. (2015). A Regression-Based Analysis of Ribosome-Profiling Data Reveals a Conserved Complexity to Mammalian Translation. *Mol Cell, 60*(5), 816-827. doi:10.1016/j.molcel.2015.11.013

Fisunov, G. Y., Evsyutina, D. V., Garanina, I. A., Arzamasov, A. A., Butenko, I. O., Altukhov, I. A., . . . Govorun, V. M. (2017). Ribosome profiling reveals an adaptation strategy of reduced bacterium to acute stress. *Biochimie, 132*, 66-74. doi:10.1016/j.biochi.2016.10.015

Galas, D. (2016). Library Preparation for small RNA sequencing using 4N adapters. In.

Geranton, S. M., Jimenez-Diaz, L., Torsney, C., Tochiki, K. K., Stuart, S. A., Leith, J. L., . . . Hunt, S. P. (2009). A rapamycin-sensitive signaling pathway is essential for the full expression of persistent pain states. *J Neurosci, 29*(47), 15017-15027. doi:10.1523/JNEUROSCI.3451-09.2009

Gerashchenko, M. V., & Gladyshev, V. N. (2017). Ribonuclease selection for ribosome profiling. *Nucleic Acids Res, 45*(2), e6. doi:10.1093/nar/gkw822

Guo, H., Ingolia, N. T., Weissman, J. S., & Bartel, D. P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature, 466*(7308), 835-840. doi:10.1038/nature09267

Hanson, J. L., Chung, M. K., Avants, B. B., Shirtcliff, E. A., Gee, J. C., Davidson, R. J., & Pollak, S. D. (2010). Early stress is associated with alterations in the orbitofrontal cortex: a tensor-based morphometry investigation of brain structure and behavioral risk. *J Neurosci, 30*(22), 7466-7472. doi:10.1523/JNEUROSCI.0859-10.2010

Hornstein, N., Torres, D., Das Sharma, S., Tang, G., Canoll, P., & Sims, P. A. (2016). Ligation-free ribosome profiling of cell type-specific translation in the brain. *Genome Biol, 17*(1), 149. doi:10.1186/s13059-016-1005-1

Hsu, P. Y., Calviello, L., Wu, H. L., Li, F. W., Rothfels, C. J., Ohler, U., & Benfey, P. N. (2016). Super-resolution ribosome profiling reveals unannotated translation events in Arabidopsis. *Proc Natl Acad Sci U S A, 113*(45), E7126-E7135. doi:10.1073/pnas.1614788113

Huch, S., & Nissan, T. (2014). Interrelations between translation and general mRNA degradation in yeast. *Wiley Interdiscip Rev RNA, 5*(6), 747-763. doi:10.1002/wrna.1244

Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M., & Weissman, J. S. (2012). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc, 7*(8), 1534-1550. doi:10.1038/nprot.2012.086

Ingolia, N. T., Ghaemmaghami, S., Newman, J. R., & Weissman, J. S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science, 324*(5924), 218-223. doi:10.1126/science.1168978

Ingolia, N. T., Lareau, L. F., & Weissman, J. S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell, 147*(4), 789-802. doi:10.1016/j.cell.2011.10.002

Jang, C., Lahens, N. F., Hogenesch, J. B., & Sehgal, A. (2015). Ribosome profiling reveals an important role for translational control in circadian gene expression. *Genome Res, 25*(12), 1836-1847. doi:10.1101/gr.191296.115

Jensen, B. C., Ramasamy, G., Vasconcelos, E. J., Ingolia, N. T., Myler, P. J., & Parsons, M. (2014). Extensive stage-regulation of translation revealed by ribosome profiling of Trypanosoma brucei. *BMC Genomics, 15*, 911. doi:10.1186/1471-2164-15-911

Ji, Z., Song, R., Regev, A., & Struhl, K. (2015). Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. *Elife, 4*, e08890. doi:10.7554/eLife.08890

Liu, B., Molinaro, G., Shu, H., Stackpole, E. E., Huber, K. M., & Richter, J. D. (2019). Optimization of ribosome profiling using low-input brain tissue from fragile X syndrome model mice. *Nucleic Acids Res, 47*(5), e25. doi:10.1093/nar/gky1292

Lutz, P. E., Tanti, A., Gasecka, A., Barnett-Burns, S., Kim, J. J., Zhou, Y., . . . Turecki, G. (2017). Association of a History of Child Abuse With Impaired Myelination in the Anterior Cingulate Cortex: Convergent Epigenetic, Transcriptional, and Morphological Evidence. *Am J Psychiatry, 174*(12), 1185-1194. doi:10.1176/appi.ajp.2017.16111286

Michel, A. M., Mullan, J. P., Velayudhan, V., O'Connor, P. B., Donohue, C. A., & Baranov, P. V. (2016). RiboGalaxy: A browser based platform for the alignment, analysis and visualization of ribosome profiling data. *RNA Biol, 13*(3), 316-319. doi:10.1080/15476286.2016.1141862

Mohammad, F., Woolstenhulme, C. J., Green, R., & Buskirk, A. R. (2016). Clarifying the Translational Pausing Landscape in Bacteria by Ribosome Profiling. *Cell Rep, 14*(4), 686-694. doi:10.1016/j.celrep.2015.12.073

Obara, I., Tochiki, K. K., Geranton, S. M., Carr, F. B., Lumb, B. M., Liu, Q., & Hunt, S. P. (2011). Systemic inhibition of the mammalian target of rapamycin (mTOR) pathway reduces neuropathic pain in mice. *Pain, 152*(11), 2582-2595. doi:10.1016/j.pain.2011.07.025

Ozadam, H., Geng, M., & Cenik, C. (2020). RiboFlow, RiboR and RiboPy: an ecosystem for analyzing ribosome profiling data at read length resolution. *Bioinformatics, 36*(9), 2929-2931. doi:10.1093/bioinformatics/btaa028

Perkins, P., Mazzoni-Putman, S., Stepanova, A., Alonso, J., & Heber, S. (2019). RiboStreamR: a web application for quality control, analysis, and visualization of Ribo-seq data. *BMC Genomics, 20*(Suppl 5), 422. doi:10.1186/s12864-019-5700-7

Popa, A., Lebrigand, K., Paquet, A., Nottet, N., Robbe-Sermesant, K., Waldmann, R., & Barbry, P. (2016). RiboProfiling: a Bioconductor package for standard Ribo-seq pipeline processing. *F1000Res, 5*, 1309. doi:10.12688/f1000research.8964.1

Rao, J. S., Keleshian, V. L., Klein, S., & Rapoport, S. I. (2012). Epigenetic modifications in frontal cortex from Alzheimer's disease and bipolar disorder patients. *Transl Psychiatry, 2*, e132. doi:10.1038/tp.2012.55

Smircich, P., Eastman, G., Bispo, S., Duhagon, M. A., Guerra-Slompo, E. P., Garat, B., . . . Sotelo-Silveira, J. R. (2015). Ribosome profiling reveals translation control as a key mechanism generating differential gene expression in Trypanosoma cruzi. *BMC Genomics, 16*, 443. doi:10.1186/s12864-015-1563-8

Tonelli, L. H., Stiller, J., Rujescu, D., Giegling, I., Schneider, B., Maurer, K., . . . Postolache, T. T. (2008). Elevated cytokine expression in the orbitofrontal cortex of victims of suicide. *Acta Psychiatr Scand, 117*(3), 198-206. doi:10.1111/j.1600-0447.2007.01128.x

Uttam, S., Wong, C., Amorim, I. S., Jafarnejad, S. M., Tansley, S. N., Yang, J., . . . Khoutorsky, A. (2018). Translational profiling of dorsal root ganglia and spinal cord in a mouse model of neuropathic pain. *Neurobiol Pain, 4*, 35-44. doi:10.1016/j.ynpai.2018.04.001

Vasquez, J. J., Hon, C. C., Vanselow, J. T., Schlosser, A., & Siegel, T. N. (2014). Comparative ribosome profiling reveals extensive translational complexity in different Trypanosoma brucei life cycle stages. *Nucleic Acids Res, 42*(6), 3623-3637. doi:10.1093/nar/gkt1386

Weber, A. P. (2015). Discovering New Biology through Sequencing of RNA. *Plant Physiol, 169*(3), 1524-1531. doi:10.1104/pp.15.01081

Weinberg, D. E., Shah, P., Eichhorn, S. W., Hussmann, J. A., Plotkin, J. B., & Bartel, D. P. (2016). Improved Ribosome-Footprint and mRNA Measurements Provide Insights into Dynamics and Regulation of Yeast Translation. *Cell Rep, 14*(7), 1787-1799. doi:10.1016/j.celrep.2016.01.043

Woolstenhulme, C. J., Guydosh, N. R., Green, R., & Buskirk, A. R. (2015). High-precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. *Cell Rep, 11*(1), 13-21. doi:10.1016/j.celrep.2015.03.014

Zannas, A. S., Provencal, N., & Binder, E. B. (2015). Epigenetics of Posttraumatic Stress Disorder: Current Evidence, Challenges, and Future Directions. *Biol Psychiatry, 78*(5), 327-335. doi:10.1016/j.biopsych.2015.04.003

Zeng, C., Fukunaga, T., & Hamada, M. (2018). Identification and analysis of ribosome-associated lncRNAs using ribosome profiling data. *BMC Genomics, 19*(1), 414. doi:10.1186/s12864-018-4765-z

Zhang, Y., Xiao, Z., Zou, Q., Fang, J., Wang, Q., Yang, X., & Gao, N. (2017). Ribosome Profiling Reveals Genome-wide Cellular Translational Regulation upon Heat Stress in Escherichia coli. *Genomics Proteomics Bioinformatics, 15*(5), 324-330. doi:10.1016/j.gpb.2017.04.005