Dissecting the role of cytoplasmic lncRNAs in neuronal differentiation and disease

Andreas Somerville Kosteletos

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

Faculty of Biological Sciences

School of Molecular and Cellular Biology

June 2024

I confirm that the work submitted is my own, except where work which has formed part of jointly authored publications has been included. My contribution and the other authors to this work has been explicitly indicated below. I confirm that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapter 3; Figures 3.2, 3.3 and Table 3.1 were published in Douka et al., 2021. This was work completed by myself with input from Dr Julie Aspden and Dr Katerina Douka. Table 3.4 Translation efficiency was calculated by Dr Isabel Birds and was published in Douka et al., 2021.

Douka, K., Birds, I., Wang, D., Kosteletos, A., Clayton, S., Byford, A., Vasconcelos, E. J. R., O'Connell, M. J., Deuchars, J., Whitehouse, A., & Aspden, J. L. (2021). Cytoplasmic long noncoding RNAs are differentially regulated and translated during human neuronal differentiation. RNA, 27(9), 1082–1101. https://doi.org/10.1261/rna.078782.121

Acknowledgements

It's always nice when a gamble plays off. I can't say I felt too confident in my decision to move to a city I had never seen, to join a supervisor I had not met in person, in the middle of a pandemic, but it all worked out surprisingly well. I feel extremely lucky to have joined the group of my supervisor, Julie Aspden. Choosing to do my PhD with Julie ranks as one of my best decisions to date, in no small part due to her wonderful mentorship, support and enthusiasm. Also, a massive thank you to my co-supervisors Guillaume Hautbergue and Juan Fontana Jordan De Urries for their encouragement and help at every stage of my PhD and to James Poulter and Erica Harris for letting me share in their amazing organoid work.

To all the members of the Aspden lab, past and present, you made this PhD a joy. I am not sure I would have made it through my first year without the help of Mark Handley, who is a source of both seemingly endless knowledge and endless patience for my many, many mistakes. To Karl Norris, I will offer an apology for all the time I spent winding you up, but it was my favourite hobby during incubations. Thank you Eilidh Ward for bringing the balance back to the lncRNA side and to Tayah Hopes, Anna Blundell, Ioannis Tsagakis, Katerina Douka and Isabel Birds, thank you for making the last 3 and a half years so enjoyable.

Thank you to the many fantastic facilities at the University of Leeds. The bioimaging facility, particularly Ruth Hughes, for all the help with microscopy and FACs sorting and the Mass Spectrometry facility for enabling me to dip my toe into the world of proteomics.

Level 9 Garstang is filled with not just wonderful scientists but wonderful people who I am lucky to call friends. This is no truer for anyone than Katie Harper and Connor Hayward (the bog runners). I will never forget the countless moments of breathless laughter, both in labs and halfway up mountains and hope we will have many more adventures in the future. Plus, you were both occasionally helpful with science.

Finally, I would like to thank both mine and my partner's families. Especially my parents, Carol and Stefanos, and my brother Stathis for all their love, support, advice, and pretend interest in my PhD while I drone on at family dinners. Also, Pippa the dog for her keen scientific insights and for making sure I got plenty of fresh air while writing up. To my other half Chloe, thank you for everything over the last 4 years. You will never stop amazing me

and thank you for the perspective you give me when talking about your much more difficult job!

Abstract

Long non-coding RNAs (lncRNAs) account for ~30% of human genes and ~40% are specifically expressed in the brain. Their precise spatiotemporal expression patterns and dysregulation in neuronal disease suggests physiological and pathological relevance. Moreover, ribosome-profiling and proteomics have shown that many cytoplasmic lncRNAs contain unannotated or small ORFs (sORF) that can be translate into functional (micro)proteins. The Aspden group has previously performed Poly-Ribo-Seq on a human neuroblastoma cell line, SH-SY5Y, in immature and differentiated states identifying 45 translated ORFs within 35 annotated lncRNAs. These lncRNAs exhibited similar translational efficiencies to protein-coding genes.

To dissect the function of peptides translated from lncRNAs, a FLAG-tagged reporter assay was performed on 13 candidates selected based on their expression during neuronal development and dysregulation in disease. Of the translated lncRNAs investigated 6/13 produce proteins with subcellular localisations indicative of function, including nuclear and mitochondrial co-localisation. Additionally, siRNA knockdown of 4/6 translated lncRNAs resulted in dysregulation of neuronal differentiation, as characterised by the disruption of neuronal gene expression and attenuated neurite outgrowth. Of the 5 translated lncRNAs tested 4 were enriched in polysomes of cortical organoids, suggesting that they are translated in advanced models of human brain development.

One translated lncRNA, LIPT2-AS1, encodes a novel protein shown by FLAG-tagging assays in SH-SY5Y cells to localise to the nucleus. Knockdown of LIPT2-AS1-protein by CRISPR-Cas9 results in a significant reduction in SH-SY5Y neurite length. Furthermore, RNA-Seq identified a global downregulation of neuronal genes and an upregulation of genes associated with Proneural-Mesenchymal transition, demonstrating that LIPT2-AS1-protein is required for normal neuronal differentiation. Protein pulldown and Mass spectrometry revealed that LIPT2-AS1-protein interacts with nuclear proteins involved in transcription and RNA splicing.

Overall, this work identifies multiple translated lncRNAs required for neuronal differentiation and highlights the potential for lncRNA-derived proteins to expand the known proteome.

Table of Contents

List of figures

List of tables

Abbreviations

AD	Alzheimer's disease
ALS	Amyotrophic lateral sclerosis
BDNF	Brain derived neurotrophic factor
BMP	Bone morphogenetic protein
CDS	Coding sequence
CNS	Central Nervous System
Co-IP	Co-immunoprecipitation
CP	Cortical plate
CRABP2	Cellular retinoic acid binding protein 2
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethylsulfoxide
dNTPs	Deoxyribonucleotide triphosphate
dORF	Downstream open reading frame
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
eEF	Eukaryotic elongation factor
eIF	Eukaryotic initiation factor
ER	Endoplasmic reticulum
eRF	Eukaryotic release factor
eRNA	Enhancer RNA
FBS	Foetal bovine serum
FGF	Fibroblast growth factor
GABA	Gamma aminobutyric acid
GFP	Green fluorescent protein
GO	Gene ontology
hCSc	Human cortical organoids/spheroids
HD	Huntington's disease
hESCs	Human Embryonic stem cells
IF	immunofluorescence
iPSC	Induced pluripotent stem cells
IZ	Intermediate zone
kDa	Kilodaltons
LINE	Long interspersed nuclear elements
lncRNA	Long noncoding RNA
LTP	Long term potentiation
mECS	Mouse embryonic stem cells
miRNA	MicroRNA
mRNA	Messenger RNA
MIR	Mammalian-wide interspersed repeat
MS	Mass spectrometry

MZ	Marginal zone
NMD	Nonsense-mediated decay
NPC	Neural progenitor cells
nRNA	non-coding RNA
NSC	Neural stem cells
nt	Nucleotide
ORF	Open Reading Frame
P/S	Penicillin-Streptomycin
PBS	Phosphate buffer saline
PCA	Principal component analysis
PCR	Polymerase chain reaction
PD	Parkinson's disease
PNS	Peripheral nervous system
RA	Retinoic acid
RAR	Retinoic acid receptor
RARE	RA response elements
RBP	RNA binding proteins
RGC	Radial glial cells
RNP	Ribonucleoprotein
RPF	Ribosome Protected Fragment
rRNA	Ribosomal RNA
RT	Reverse transcription
RT-qPCR	Real-time quantitative PCR
RXR	Retinoid X receptor
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SEM	Standard error of the mean
SHH	Sonic hedgehog
shRNA	Short hairpin RNA
SINE	Short interspersed nuclear elements
siRNA	Small interference RNA
SNP	Single nucleotide polymorphism
sORF	Small Open reading frame
SVZ	Subventricular zone
TBE	Tris Borate EDTA
TH	tyrosine hydroxylase
TPM	Transcripts per million
tRNA	transfer RNA
TSS	Transcription start site
uORF	Upstream open reading frame
UTR	Untranslated region
VZ	Ventricular zone
WNT	Wingless/Integrated

WT	Wild type

Chapter 1: Introduction

1.1 Non-coding RNA

In the human genome ~20,000 protein coding genes are annotated (Derrien et al., 2012). Much of the remaining ~98% of the genome was described by some as “junk DNA” and was perceived as having little or no function. However, with advances in sequencing technologies and transcriptomics, evidence emerged for pervasive transcription of the genome. It has now become clear that ~85% of the genome is transcribed into RNAs of various sub-types (Hangauer et al., 2013).

One of these sub-types are long non-coding RNAs (lncRNAs). lncRNAs represent a large and diverse population of RNAs >200nt long, but which lack a traditional open reading frame (ORF) of >100 codons and were therefore initially deemed non-coding. As of 2024 GENCODE reports 20,424 annotated lncRNA genes in humans, accounting for ~32% of total genes in the genome, although estimates vary greatly between databases with LNCipedia listing 56,946 high-confidence lncRNA genes (Derrien et al., 2012; Volders et al., 2019). Approximately 40% of lncRNAs are specifically expressed in the brain, suggesting their roles are particularly important in the fine tuning of our most complex organ (Derrien et al., 2012). Unlike “housekeeping” ncRNAs such as tRNAs, which show high and wide-spread expression, lncRNAs tend to be expressed at lower levels in a cell type specific pattern, potentially indicating that lncRNAs play a regulatory role rather than a ubiquitous housekeeping one (Derrien et al., 2012; Djebali et al., 2012).

From the lncRNAs that have been characterised, it is clear many do have a function and they act through a range of mechanisms, highlighting that “lncRNA” is an umbrella term to describe a diverse population RNA species. While many lncRNAs share biochemical similarities with mRNAs, with 50% possessing a poly(A) –tail, 98% being spliced, and many being capped, they do possess distinguishing characteristics (Tsagakis et al., 2020). These include the tendency to form complex secondary structures, often including multiple hairpins and RNA structural domains that confer the means to interact with protein and nucleic acid targets. Also, compared to mRNAs, lncRNAs are generally less conserved, shorter and possess fewer exons (Derrien et al., 2012; Sparber et al., 2019). Further, transcriptomics has revealed that a single lncRNA gene typically gives rise to many different

transcripts. These individual transcripts from the same lncRNA gene may localise to different parts of the cell and even perform independent functions (Lewandowski et al., 2020).

Nuclear retained lncRNAs have been the focus for the majority of research to date. The most well-known of these is Xist, a nuclear lncRNA that is required for X-chromosome inactivation during development and that shows evolutionally-conserved function across placental mammals (Sparber et al., 2019; Yen et al., 2007). Yet, the mRNA-like properties of many lncRNA transcripts allows them to enter the cytoplasm, and they are found to localise to discrete cytoplasmic locations. Further, through techniques such as RNA-Seq, Ribosome profiling and proteomics, it has become clear that a significant population of cytoplasmic lncRNAs are ribosome-associated, and that many of these "non-coding" RNAs actually possess small or unannotated open reading frames (ORF) capable of producing biologically active proteins (Bazzini et al., 2014; Brunet et al., 2020; Couso & Patraquim, 2017; Orr et al., 2020). From characterised examples, and from the dynamic expression of cytoplasmic lncRNAs during development, it seems likely that they are involved in a range of physiological functions including organ development and synaptic plasticity (Grinman et al., 2021; Rani et al., 2016). Additionally, their dysregulation is observed in - and may potentially contribute to – cancers (Carlevaro-Fita et al., 2020), neurodegenerative diseases (Sosińska et al., 2015; Wu & Kuo, 2020) and developmental disorders (Shi et al., 2017).

Despite this some researchers still believe that lncRNAs are not required for normal function. The main evidence for lncRNAs lacking function is that the knockdowns of many lncRNAs during development does not lead to observable phenotypes (Goudarzi et al., 2019). This has led some to argue that lncRNAs are largely non-functional or mediate their functions solely by the act of being transcribed, with the transcripts produced a non-functional bi-product. However, there is an ever-growing body of evidence indicating that cytoplasmic lncRNAs, and the proteins some can produce, are biologically active and essential for normal cellular function(Anderson et al., 2015; Grinman et al., 2021; Rani et al., 2016). Perhaps the negative results seen by some are due to lncRNAs possessing subtle or context-dependant roles, lncRNAs generally being less sensitive to nucleotide mutations and being a heterogenous population likely containing both functional and non-functional genes (S. J. Liu et al., 2017). Therefore, more research is needed to uncover the precise functions

of individual lncRNAs and to elucidate what roles they play in neuronal development and disease.

1.1.1 Classification of long non-coding RNAs

The standard method to classify lncRNA genes is according to their genomic locations. Two subtypes make up the majority of lncRNA genes. **Intergenic** lncRNAs are located in so-called "gene deserts" with the nearest coding gene being >1Kb away in either direction. The second large group are **anti-sense** lncRNAs. These are transcribed from the anti-sense strand of protein coding genes and therefore contain sequence complementary to nascent and/or processed mRNAs. However, their roles are not restricted to acting on the gene they are anti-sense to, as has been made clear from the anti-sense lncRNAs so far characterised such as FGD5-AS1 which promotes tumour growth independently of its sense transcript, with FGS5-AS1 RNA acting as a sponge for miR-873-5p(Miao et al., 2020; N. Zhang et al., 2021; J. B. Zhao et al., 2020). Of the remaining three categories, two are also found within canonical genes, but running in the sense direction. These are termed **sense lncRNAs** or **intronic lncRNAs** depending on their localisation to either exons or introns, respectively. The final class are the **bi-directional lncRNAs,** which share a promoter region with a protein-coding gene but run in the opposite direction (Figure 1.1) (Tsagakis et al., 2020).

LncRNA function is defined as either *cis*- or *trans*-acting. *Cis*-acting refers to lncRNAs that functions near or at the site of transcription, for example anti-sense lncRNAs modulating the transcription of its sense gene. *Trans*-acting lncRNAs function away from their site of transcription, either within the nucleus or in the cytoplasm (Statello et al., 2021). While widely used, classification of lncRNAs based on their genes genomic context provides limited information regarding lncRNA physiological and molecular function. With further study, more informative sub-classifications may emerge once lncRNA functions are better understood.

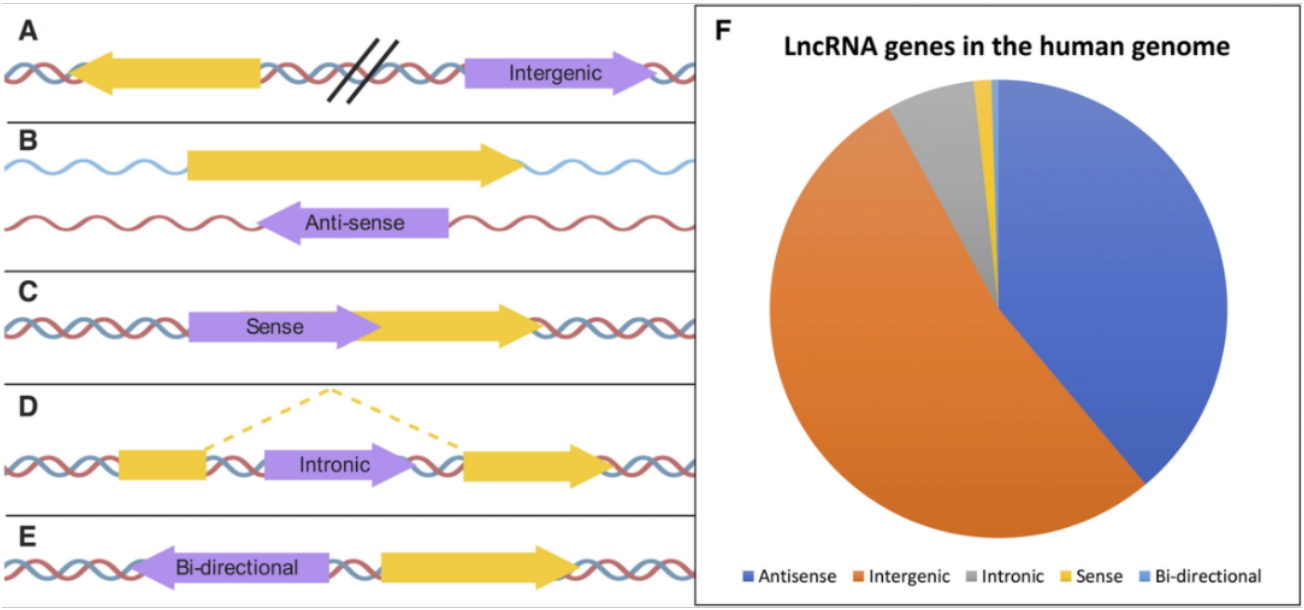


Figure 1.1: Schematic depicting the genomic structure and relative number of lncRNAs*. lncRNAs are purple and protein coding genes yellow A) Intergenic lncRNAs B) Anti-sense lncRNAs C) Sense lncRNAs D) Intronic lncRNAs E) Bi-directional promoter F) proportions of each class of lncRNA within the human genome. Figure from* (Tsagakis et al., 2020).

1.1.2 Cytoplasmic lncRNAs

Half of all lncRNAs are enriched in the cytoplasm (Carlevaro-Fita et al., 2016). The fact that nuclear export requires energy expenditure suggests this is functionally significant (Tsagakis et al., 2020). Indeed, once in the cytoplasm lncRNAs can localise to specific sub-cellular compartments, where some have been functionally characterised. To function, lncRNAs interact with a diverse range of partners (Figure 1.2). Examples of activities involving different partners include being a 'molecular sponge' for miRNAs to prevent mRNA degradation, providing a scaffold for protein complexes, and interacting with translational activators to mediate mRNA translation (X. Zhang, Wang, et al., 2019). Of note, individual lncRNAs can also have multiple activities. For example, in the human testis, TUG1 is essential for fertility and functions both in *cis* and *trans* as an RNA. Additionally, TUG1 transcripts can localise to the polysomes and appear to be translated into a micropeptide of unknown function (Lewandowski et al., 2020).

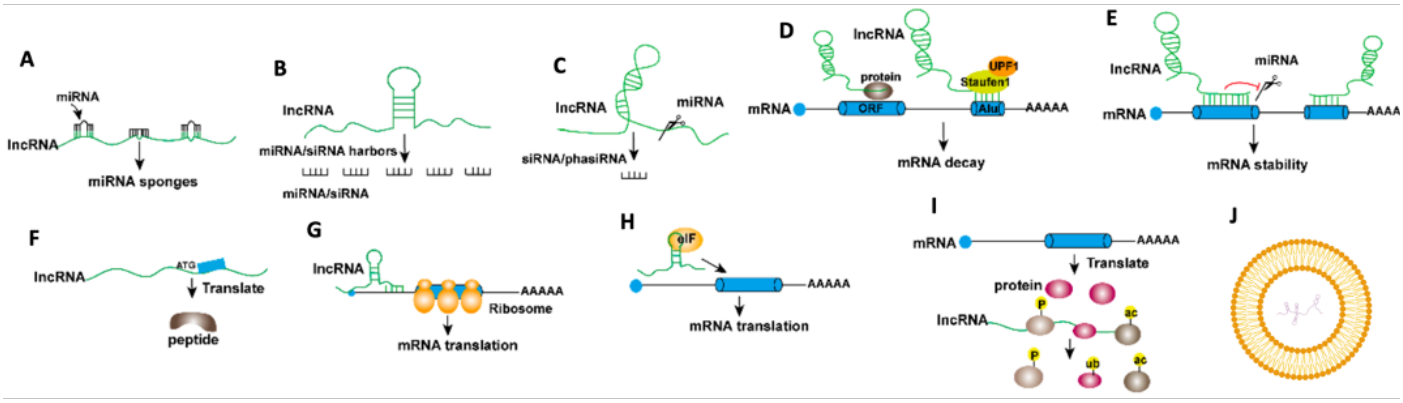


Figure 1.2: Mechanisms of action for cytoplasmic lncRNAs. *A) lncRNAs with complementary bases to miRNAs act as "molecular sponges" that sequester miRNAs* (Y. Wang et al., 2013; Q. Zhang et al., 2022). *B) lncRNAs can be degraded into smaller RNA subtypes including miRNAs and siRNAs (X. Cai & Cullen, 2007). C) miRNA targeted degradation of lncRNA can produce siRNAs and phasiRNAs in plants. D) lncRNAs can direct Staufen1-mediated decay of specific mRNAs (Gong & Maquat, 2011). E) lncRNAs can directly bind to mRNAs to increase their stability and protect mRNA from miRNA degradation* (Y. Zhao et al., 2018) *. F) lncRNAs containing sORFs can be translated to produce micropeptides* (Stein et al., 2018; C. Zhang et al., 2022). *G) lncRNAs can interact with mRNAs to regulate their translation* (Carrieri et al., 2012). *H) lncRNAs can interact with and direct localisation of proteins* (Campalans et al., 2004). *I) lncRNAs interact with proteins to regulate post-translational modifications. J) Some lncRNAs are enriched in exosomes and released from cells. Adapted from* (X. Zhang, Wang, et al., 2019).

1.1.3 Polysome associated lncRNAs

Polysomes are sites of active translation, where multiple ribosomal complexes translate a single mRNA transcript resulting in high levels of protein synthesis. Over 70% of human cytoplasmic lncRNAs are polysome-associated (Carlevaro-Fita et al., 2016). As a group, polysome-associated lncRNAs have distinct characteristics compared to other cytoplasmic lncRNAs, showing a greater cytoplasmic:nuclear localisation ratio than even mRNAs (Carlevaro-Fita *et al.*, 2016). Factors that correlate with enrichment at polysomes include transcript capping and presence of a long poly(A) tail (Carlevaro-Fita *et al.*, 2016). In contrast, the presence of the retroviral insertions LINEs and SINEs decreases the likelihood of a lncRNA transcript's polysomal association(Carlevaro-Fita et al., 2016; Zeng & Hamada, 2018). However, lncRNA enrichment at the polysome is not a static relationship; for example, dynamic polysomal-lncRNA-association throughout cardiomyogenesis of human embryonic stem cells has been reported (Pereira et al., 2020). The polysome-association of lncRNAs does not simply mirror transcript abundance levels, indicating that the regulation of polysomal-association is influenced by other unknown factors (Douka et al., 2021; Pereira et al., 2020). Also, different lncRNA transcripts from the same gene can show different levels of polysomal association, with the TUG1 locus producing both a nuclear retained transcript that modulates the transcription of genes in cis and a cytoplasmic TUG1 transcript that encodes a micropeptide (Lewandowski et al., 2020; Zeng et al., 2018). Of note, lncRNAs found enriched at the polysome have a higher median expression than both cytoplasmic free and nuclear lncRNAs and are expressed in a greater number of tissues (Carlevaro-Fita *et al.*, 2016).

Current research has only characterised a handful of polysomal lncRNAs and the molecular and physiological roles of the majority are unknown. One possibility is that polysomal lncRNAs undergo translation to produce (micro)proteins. While this is likely true for a small proportion (discussed later), the majority of polysomal lncRNAs are not identified as translated by Poly-Ribo-Seq (Couso & Patraquim, 2017). Another proposed function of polysome-associated lncRNAs is regulation of mRNA translation. This activity has been described for some anti-sense lncRNAs associated with cognate sense transcripts. For example, Uchl1-AS is an antisense lncRNA that regulates the mRNA encoding UCHL1. Both are enriched in mice ventral midbrains, specifically in dopaminergic neurons (Carrieri et al.,

2012). Transient Uchl1-AS expression has no effect on Uchl1 mRNA as shown by RT-qPCR levels but western blots showed a significant increase in UCHL1 protein levels after 24 hours, suggesting Uchl1-AS1 functions post-transcriptionally (Carrieri *et al.*, 2012). The apparent translational enhancement of UCHL1 expression can be attenuated by targeted deletion of Uchl1-AS1's SINB2 or Alu repeats, indicating these regions are essential for the lncRNA's function. Of note, UCHL1 protein, but not mRNA, levels are also increased by rapamycin treatment despite rapamycin being an inhibitor of cap-dependant translation. This effect is explained by the re-localisation of Uchl1-AS from the nucleus to the cytoplasm upon rapamycin treatment and its subsequent enhancement of the association between Uchl1 mRNA and polysomes (Figure 1.3) (Carrieri *et al.*, 2012). Additionally, MAPT-AS1 has been shown to regulate translation of its sense gene MAPT, which encodes tau protein. Polysome-profiling identified that overexpression of MAPT-AS1 caused MAPT mRNA to shift from heavy to light polysomes, indicating translational suppression (Simone et al., 2021). Furthermore, by mutating MAPT-AS1's mammalian-wide interspersed repeat (MIR) or natural anti-sense region to MAPT mRNA this effect was abolished indicating these regions are required for MAPT-AS1 to modulate the translation of MAPT (Simone *et al.*, 2021).

Of the thousands of lncRNAs found in polysomes only a proportion possess a translated-(s)ORFs or sequence complementary to mRNA. Thus, the roles and factors underlying localisation of the majority of polysomal lncRNAs remain unclear. It must be considered that ribosome-association of lncRNAs may simply control the levels of lncRNA transcripts that function elsewhere in the cell. This argument is supported by the observation that ribosome-associated lncRNAs are often highly sensitive to nonsense mediated decay (NMD), a translation-dependent surveillance pathway that degrades RNAs (Lloyd et al., 2020; Zeng et al., 2018). Yet, evidence supporting translation of some lncRNAs at the polysome includes the finding that unlike mRNA levels, lncRNAs levels do not show a linear relationship between transcripts found in RNA-Seq of polysomal fractions and those with translation signals detected by Poly-Ribo-Seq (Couso & Patraquim, 2017). This suggests that Poly-Ribo-Seq accurately discriminates the subpopulation of lncRNAs undergoing active translation (Aspden et al., 2014; Couso & Patraquim, 2017). Additionally, NMD targets are typically depleted in the polysomal fraction compared to total RNA and monosomes (Lloyd et al., 2020). Therefore, when considering this evidence alongside the few polysomal lncRNAs that

have been functionally characterised, it seems likely that the polysome contains multiple populations of lncRNAs including functional but non-coding lncRNAs, translated lncRNAs, and lncRNAs to be degraded.

However, in the context of the brain, specifically rat hippocampal neurons, efficient translation of mRNAs by monosomes has been observed. This indicates that in specific contexts, such as at the synapse where space is limited, monosomes may be the primary sites of protein production (Biever et al., 2020). Therefore, in neuronal contexts Poly-Ribo-Seq may exclude the translation of ORFs at the synapse. To date, no investigation of lncRNAs association and roles at synaptic monosomes have been conducted. However, multiple lncRNA transcripts have been shown to localise to synapses, often in response to stimuli including action potentials (Grinman et al., 2021).

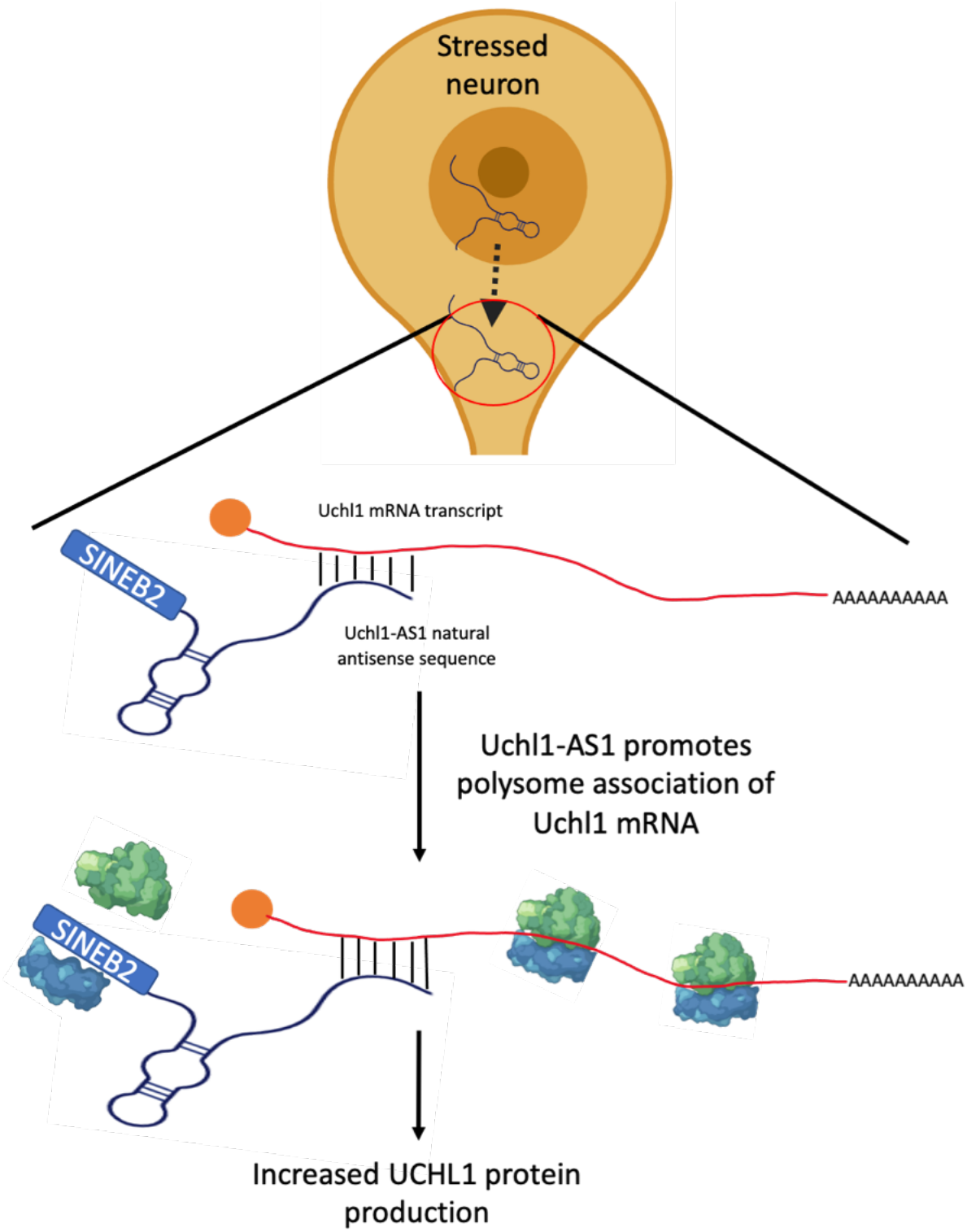


Figure 1.3: The lncRNA Uchl1-AS1 responds to stress to modulate translation of its sense gene. *Uchl1-AS1 is shuttled out of the nucleus in response to stress and interacts with its sense transcript, Uchl1, through a natural anti-sense region. Uchl1-AS1 promotes Uchl1 mRNA association with the polysome, and therefore translation, in a SINEB2 dependent manner* (Carrieri et al., 2012). *Figure generated in Biorender.*

1.2 Human neuronal differentiation

Neural development in humans begins during the third week of gestation and continues postnatally into adolescence, with the brain increasing in size four-fold in pre-school years reaching ~90% of adult volume by age 6 (Stiles & Jernigan, 2010). To reach the ~100 billion neurons at birth, 250,000 neurons per minute must be generated throughout pregnancy, and these neurons must be discreetly organised into regions to perform specific functions (Ackerman, 1992)

This process is orchestrated through series of essential stages starting with (1) expansion of undifferentiated brain cells, (2) migration of undifferentiated cells to a specific location within the immature brain, (3) initiation of differentiation as a suitable cell type for the location (4) accumulation of similar cells (5) formation of synaptic connections (6) competition between synapses leading to selective stabilisation and elimination, or pruning, of synapses (7) continued synaptic plasticity throughout life (Ackerman, 1992; Therianos et al., 1995).

Development of the human brain begins at week 3 in gestation. Neural induction is triggered by the inhibition of bone morphogenetic protein (BMP) signalling and activation of fibroblast growth factor (FGF) pathways by the release of factors from “organiser” cells underneath the ectoderm. This results in the stem cells of the ectoderm progressing towards a neuronal fate, forming the neural plate composed of neuronal progenitor cells (NPC) (Muñoz-Sanjuán & Brivanlou, 2002). Once the neural plate is formed, two ridges form at its edges with all NPCs in-between. The ridges close over to meet each other forming the hollow neural tube. Fusion happens initially at the centre of the forming neural tube and then spreads in both rostral and caudal directions. The neural tube will give rise to the brain and spinal cord. Patterning of the neural tube is required to specify regions with the anterior-posterior patterning primarily determined by FGFs, WNTs, and RA (Perrier et al., 2004; Yan et al., 2005). WNTs determine the anterior neural tubes development into precursors of the forebrain, mid-hindbrain, and anterior spinal cord. At the posterior end RA and FGFs regulate the segmentation of the spinal cord section. Patterning also occurs along the dorsal-ventral axes and is primarily controlled by BMP and WNT morphogens at the

dorsal side and high SHH levels at the ventral side. These morphogen gradients determine neural progenitor identity at specific regions of the neural tube (Tao & Zhang, 2016).

From the lateral margins of the neural plate that are not incorporated into the neural tube Neural crest cells differentiate and migrate away from the neural fold. Neural crest cells give rise to many cell types including neurons and glia of the peripheral and enteric nervous system as well as melanocytes, craniofacial connective tissue and bone, adrenal gland medulla cells and cardiac endothelium (Erb, 2006; Gilbert, 2000).

The neural tubes luminal surface is composed of germinal neuroepithelial cells (the germinal epithelium). Once neural tube folding is neuroepithelial cells begin expressing glial specific markers and transition into radial glial cells (RGC). RGC have the potential to differentiate into both neurons and glia and divide “vertically” with one daughter cell remaining as a RGC connected to the ventricular surface while the other daughter cell migrates away from the lumen and will not divide again (Martínez-Cerdeño & Noctor, 2018). Progressive division and migration builds layers up from the neural tube, eventually resulting on the generation of the intermediate zone and the germinal epithelium is now known as the ventricular zone. Cells migrating through the intermediate zone differentiate into neurons and glia. Neurons cell bodies then settle within the intermediate zone or (cortical plate in the embryonic cerebral cortex) and extend axons, forming connections between neighbouring cells and extending axons away from the lumen to create the marginal zone, composed primarily of axons, otherwise known as white matter (Figure 1.4) (Fernández et al., 2016; Gilbert, 2000).

Neuronal differentiation into specific neuronal cells is determined by the location they migrate into. Once differentiated, new neurons must form synapses with target cells to generate the complex wiring of the CNS. Axonal outgrowth is guided by a range of chemical signals that may either attract of repel growth. These include morphogens such as BMP, FGF, WNT, SHH as well as netrins and ephrins (Tamariz & Varela-EchavarrÃa, 2015). These are detected by guidance receptors, including Robo1, ephrin-B1 and neuropilin 1 (Nrp1) (Zang et al., 2021).

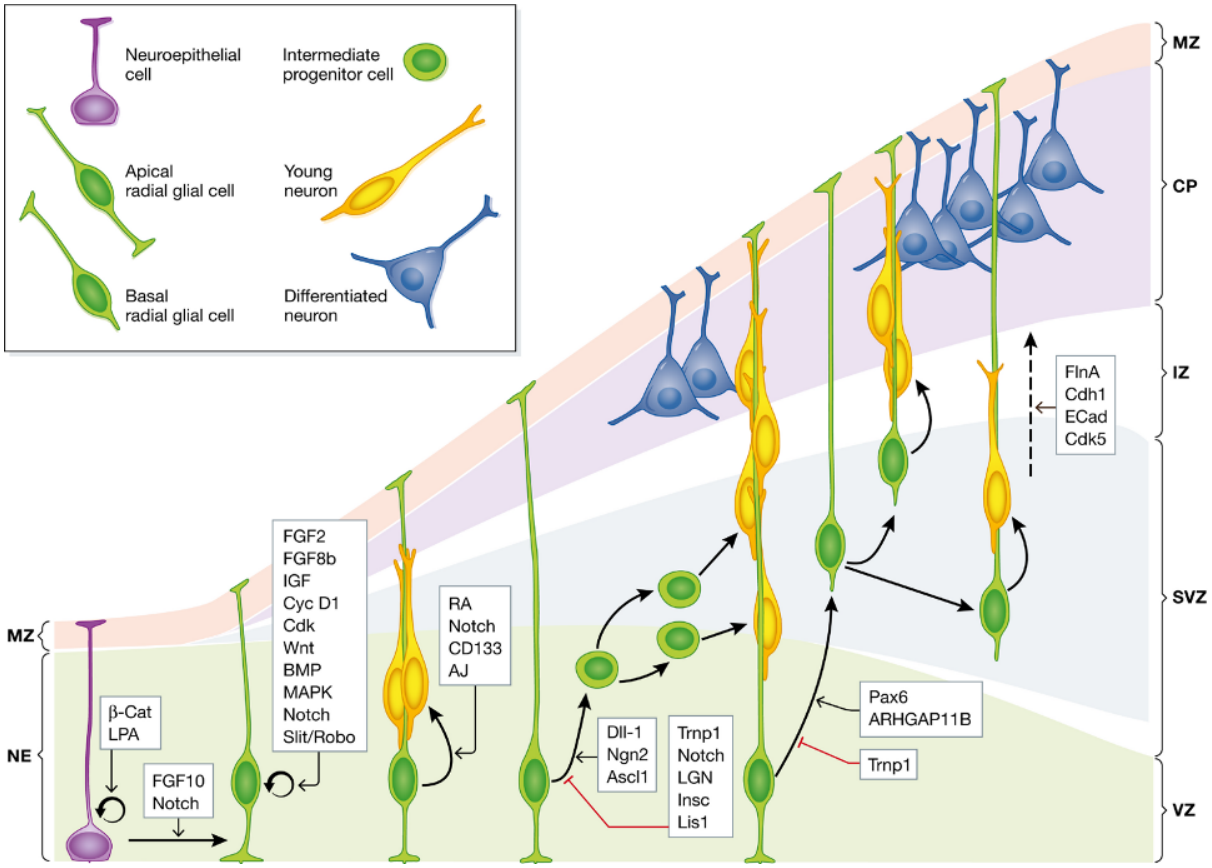


Figure 1.4: Schematic showing the development of the embryonic cortex. *With the fusion of the neural tube neuroepithelial cells start expressing glial specific markers and transition into radial glial cells (RGC). RGCs divide asymmetrically with one daughter remaining multipotent and the other daughter cell migrating away from the ventricular zone (VZ) through the subventricular zone (SVZ), intermediate zone (IZ) and undergo terminal differentiation once reaching the cortical plate (CP). Neurons of the CP extend axons into the marginal zone (MZ). At later staged of development RGCs migrate to the SVZ and continue developing into mature neuronal cells. Key molecules and signalling pathways regulating each stage are indicated. Figure from Cerebral cortex expansion and folding: what have we learned?* (Fernández et al., 2016)

1.3 lncRNAs in the brain

The human brain specifically expresses ~40% of all lncRNAs, compared to only ~14% of protein coding genes (Derrien et al., 2012; Sjöstedt et al., 2020). Many neuronal lncRNAs show precise spatial and temporal expression patterns throughout development and lncRNAs are often observed to be dysregulated in neuronal diseases (L. Li et al., 2019; Sarropoulos et al., 2019). Therefore, elucidating the function of lncRNAs will likely enhance our understanding of neuronal development and disease.

1.3.1 lncRNAs during neuronal differentiation

To achieve the formation of complex neuronal structures, precise temporal and spatial control of gene expression, protein production and modification is required. At a transcript abundance level, RNA-seq throughout early embryogenesis to adulthood identified that 18.6% of all lncRNAs detected showed developmentally dynamic expression (defined as differential expression at different time points within the same organ) (Sarropoulos et al., 2019). While this proportion is lower than that of protein coding genes (73%), it nevertheless represents ~6,000 dynamically regulated lncRNAs transcripts (Sarropoulos *et al.*, 2019). Of these ~6,000, ~20% are dynamically expressed during brain development. Multiple studies have reported large numbers of lncRNAs with differential expression throughout neurogenesis, with lncRNAs accounting for 1/3 of all transcriptionally altered genes during neuronal differentiation (B. Deng et al., 2017; M. Lin et al., 2011).

lncRNAs with an established role in neuronal development include the gene lncND. This lncRNA is enriched in neuronal progenitor and radial glial cells. lncND prevents neuronal differentiation by sequestering the miRNA miR-143-3p, which otherwise mediates degradation of the mRNAs of the receptors NOTCH-1/2, leading to neuronal differentiation (Rani et al., 2016). Another 'molecular sponge' lncRNA involved in development is Linc-RoR. siRNA knockdown of Linc-RoR in hESCs resulted in decreased levels of NANOG and OCT4 mRNA and protein. RNA- immunoprecipitation of linc-RoR, NANOG, OCT4 and SOX2 in HEK293 cells demonstrated that they are all targeted by a shared pool of miRNAs. Therefore, Linc-RoR prevents stem cells differentiation by blocking the degradation of key pluripotency genes (Y. Wang et al., 2013).

Beyond transcription, it has also demonstrated that translational control is essential for the development of specialised cell types and tissues. RNA-seq and Ribo-Seq of different tissues throughout mouse development has identified hundreds of mRNAs that were regulated primarily at the translational level during development (Fujii et al., 2017). Transcripts coding for factors in core signalling pathways that orchestrate cell fate and tissue formation such as Wnt, Hippo and Shh were particularly sensitive to translational control and can demonstrate tissue specific translation, but not transcription, during embryonic development (Fujii et al., 2017). The mechanism(s) underlying this tuneable translational control are not fully understood, but lncRNAs have been implicated in the translational control of key developmental factors (Statello et al., 2021; Yoon et al., 2012). Additionally, lncRNAs may regulate translation in mature neurons, a process thought to be important for synaptic plasticity and modulating the local environment of neuronal synapses, possibly in response to external signals. RNA-FISH and RT-qPCR demonstrated that lncRNA ADEPTR is directed to neuronal synapses in a cAMP-dependant manner in response to excitatory synaptic transmission but is repressed by GABA activity. Silencing of ADEPTR by shRNA and gapmers has been shown to inhibit mushroom spine formation and size indicating that ADEPTR is required for cAMP dependant synaptic structure modulation (Grinman et al., 2021). Additionally, transcript levels of BC200, a lncRNA of Alu-origin, are significantly upregulated at active synapses where they are known to interact with eIF4A to regulate synaptic translation (Mus et al., 2007; Shi et al., 2017; Sosińska et al., 2015).

Other important roles for lncRNAs include their involvement in maturation of non-neuronal cell types in the brain. OLMALINC - a brain-enriched lncRNA that is conserved in primates - shows significantly greater expression in white matter vs grey matter (Mills et al., 2015). OLMALINC is essential for oligodendrocyte maturation and its knockdown prevents the full maturation of radial glia due to altered expression of gene clusters involved in cellular adhesion and cytostructure. OLMALINC depletion also results in increased SOX4 mRNA levels, incompatible with differentiation (Mills et al., 2015). Additionally, lncRNAs appear to be involved in neuronal regeneration following damage. Nerve injury *in vivo* induces expression of many lncRNAs (Perry et al., 2018). Clusters of these lncRNAs can be separated based on temporal expression patterns, indicating lncRNAs may have time-point specific roles in neuroregeneration. Knockdown of two of these lncRNAs, *Silc1* and Norris1,

prevented normal regeneration of the injured nerve in cultured cells and mice, demonstrating that lncRNAs are not only essential for normal development of a range of CNS cell types but also regeneration of damaged neurons (Perry et al., 2018).

While not well understood, lncRNAs have been implicated in multiple processes where neuronal cells undergo dynamic changes, including maturation, synaptic plasticity, and regeneration (Perry et al., 2018; Shi et al., 2017). These potential roles must be validated but may indicate that lncRNAs are involved in the orchestration of neuronal cell changes - both during development, and in mature neurons reacting to signals or trauma.

1.3.2 lncRNAs in neurodegenerative disease

Neurodegenerative diseases are characterised by progressive atrophy of neurons and pose one of the greatest challenges for modern medicine with aging populations and a complete lack of disease modifying therapeutics. As of 2015 dementia and Alzheimer's disease (AD) became the leading cause of death in the UK (*Leading Causes of Death, UK - Office for National Statistics*, 2021). A similar clinical picture can be seen for Amyotrophic lateral sclerosis (ALS), Huntington's disease (HD) and Parkinson's disease (PD) where disease mechanisms are not fully understood, and current treatments are limited. lncRNAs have known roles regulating multiple pathways involved in apoptosis, mitochondrial function, neurotrophic factor availability, and neuronal differentiation (Qian et al., 2019; Rani et al., 2016; Shi et al., 2017). It is therefore not surprising that they are often seen dysregulated in neurodegenerative disease, including those listed above (Y.-Y. Wu & Kuo, 2020).

Dysregulation of BC200 is observed in AD. Northern hybridisation of BC200 in human brain autopsies revealed BC200 had abnormal distribution and upregulated in AD patient brains. Additionally, BC200 showed higher expression in brain regions most highly affected by AD pathology compared to regions typically spared from degeneration (Mus et al., 2007). Furthermore, the degree of BC200 expression correlates with clinical dementia score in AD patients (Mus et al., 2007). Other lncRNAs are directly linked to Aβ42 production, a hallmark of AD pathology produced by the enzyme BACE1 from amyloid precursor protein (APP). Antisense to the BACE1 transcript is the lncRNA BACE1-AS1, which stabilises BACE1 mRNA by forming an RNA duplex that protects BACE1 from miRNA-485-5p-mediated degradation.

Therefore, increased BACE1-AS1 expression leads to an increase in Aβ42 production (F. Li et al., 2019). lncRNAs may also offer candidates for novel biomarkers for AD. Both RP11-462G22.1 and PCA3 are differentially expressed in AD compared to healthy individuals, and as they are released in exosomes, their levels can easily be detected by CSF sampling (Gui et al., 2015). Similarly, BACE1-AS1 is significantly upregulated in blood plasma of AD patients (Fotuhi et al., 2019)(Figure 4).

ALS is a complex, multifactorial disease characterised by selective degeneration of motor neurons (Butti & Patten, 2019). SNAP25-AS, the antisense lncRNA to SNAP25, is downregulated in ALS-TDP43 patients. SNAP25 encodes an essential protein component of the *trans*-SNARE complex that mediates synaptic vesicle-membrane fusion and neurotransmitter release at the presynaptic terminal (Vangoor et al., 2021). However, the exact reason for, and significance of, SNAP25-AS downregulation in ALS, is not yet understood. Similarly, the functional significance of the deregulation of two lncRNAs, ZEB1-AS1 and ZBTB11-AS, in peripheral blood mononuclear cells from ALS patients spinal cords is also unknown, although both have sense genes involved in transcriptional regulation (Vangoor et al., 2021).

While it has been shown that some lncRNAs are dysregulated in neurodegenerative disease (*Table 1*) more research is required to determine which lncRNAs are involved, and whether they are primary drivers or secondary events in disease progression. Still, lncRNAs potentially offer much needed novel drug targets and biomarkers for the treatment and detection of neurodegenerative disease.

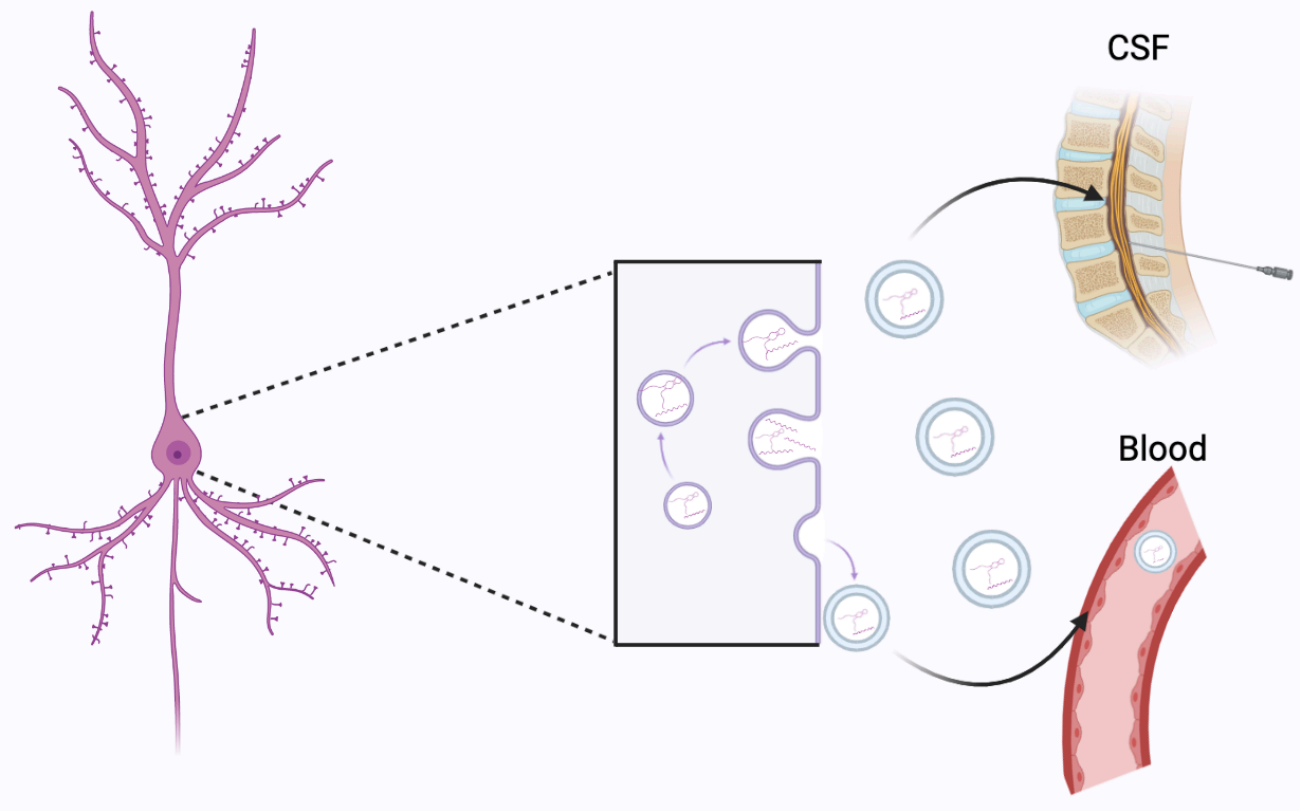


Figure 1.5: lncRNAs are potential biomarkers as can be released in exosomes from diseased cells. *AD sufferers show significantly increased CSF levels of RP11-462G22.1 and PCA3 and blood plasma levels of BACE1-AS1 compared to healthy controls* (Fotuhi et al., 2019; Gui et al., 2015). *Figure made in Biorender.*

1.4 lncRNAs in Cancer

Tumorigenesis is driven by a collection of genetic mutations that promote cancer phenotypes such as proliferation and migration. lncRNAs are currently being investigated for their roles in cancer biology as sequencing experiments consistently identify dysregulation of lncRNA expression in cancer. However, as the majority of lncRNAs have not been investigated for their functional roles in disease it is impossible to say how many cancer associated lncRNAs have a causative role in disease (Carlevaro-Fita et al., 2020).

Experimental evidence has confirmed that multiple lncRNAs contribute to cancer pathology and progression. For example, the cytoplasmic lncRNA LINC01116 is significantly upregulated in glioma compared to normal cells and is thought to contribute to glioma pathogenesis. Expression levels of LINC01116 correlated with higher tumour grade, increased recurrence and poor survival (Ye et al., 2020). Mechanistically, LINC01116 appears to enhance cell proliferation as knockdown causes glioma cells to accumulate in G0/G1-phase. siRNA knockdown of LINC01116 is reported to result in decreased levels of VEGFA, a key promoter of angiogenesis, increased levels of which are associated with enhanced tumorigenesis and metastasis. A dual luciferase experiment indicated that both LINC01116 and VEGFA mRNA are targeted by miR-31-5p. Therefore, LINC01116 may promote VEGFA by sequestering miR-31-5p, preventing silencing of VEGFA mRNA (Ye et al., 2020). Similarly, the antisense lncRNA FGD5-AS1's expression levels also correlate with glioma prognosis, and FGD5-AS1 is considered an oncogene in multiple tumour types. It is thought to regulate the Wnt/β-catenin pathway by sequestering miR-129-5p to prevent the degradation of HNRNPK, which in turn enhances the expression of proteins related with the Wnt/β-catenin pathway (J. B. Zhao et al., 2020). In glioma, upregulation of Wnt signalling is related to a worse prognosis and chemotherapy-resistance, and it is a noted characteristic of glioma stem cells (Zuccarini et al., 2018).

Furthermore, lncRNAs secretion from cells has been reported to be dysregulated in cancers. lncRNA abundance in exosomes does not mirror cellular levels indicating selective loading of specific lncRNAs, as is also seen for mRNA and miRNAs (Gezer et al., 2014). LincRNA-p21 is an anti-tumour lncRNA that is found significantly enriched in exosomes, up to 6500-fold in HeLa cells, and its exosomal enrichment is increased when cells are subjected to DNA

damage (Gezer et al., 2014). It is also secreted from human prostate cancers and can be detected at significantly higher levels in urine samples from cancer sufferers than in those from healthy individuals. Therefore, lincRNA-p21 could potentially serve as novel biomarker to improve the diagnosis of prostate cancer (Isin et al., 2015). Another lncRNA that shows promise as a biomarker is GIHCG, which is upregulated in a range of cancers including liver, renal and cervical, with higher expression being correlated with increased progression and migration. Serum GIHCG levels where significantly higher in cervical cancer patients compared to control individuals, again suggesting lncRNAs have potential as biomarkers (X. Zhang, Mao, et al., 2019).

Cytoplasmic lncRNAs with potential functions in neuronal disease					
Gene name	**Relevant disease**	**Expression in disease**	**Impact**	**Proposed mechanism of action**	**Reference**
NEAT1	AD	Upregulated	Pathogenic	Decoy for miRNA-124 leading to higher BACE1 levels and Aβ42 production.	(M.-Y. Zhao et al., 2019)
BACE1-AS1	AD	Upregulated	Pathogenic	Upregulated in AD and stabilises BACE-1 mRNA contributing to Aβ42 production.	(F. Li et al., 2019)
BC1/BC200	AD	Upregulated in areas most affected by AD	Pathogenic	Dysregulated in AD, is a known regulator of synaptic translation and plasticity. Has been shown to upregulated translation of APP causing Aβ aggregation.	(Mus et al., 2007; Sosińska et al., 2015)
LINC01116	Glioma	Upregulated	Oncogenic	Appears to be required for cell cycle progression, also sequesters miR-31-5p leading to enhanced VEGFA translation.	(Wang et al., 2020; Ye et al., 2020)
FGD5-AS1	Glioma	Upregulated	Oncogenic	Sequesters miR-129-5p to prevent the degradation of HNRNPK, leading to enhanced expression of proteins in the Wnt/β-catenin pathway.	(J. B. Zhao et al., 2020)
Linc-ROR	Glioma	Downregulated	Protective	Thought to supress tumour growth through inhibition of KLF4.	(Toraih et al., 2019)
TUG1	Glioma	Upregulated	Reports of both	Described both as oncogene and tumour suppressor.	(Katsushima et al., 2016; J. Li et al.,2016)
CRNDE	Glioma	Upregulated	Oncogenic	Expression promotes glioma growth, possibly through inhibiting miR-136-5p from degrading Bcl-2 and Wnt.	(D.-X. Li et al., 2017)
DLGAP1-AS2	Glioma	Upregulated	Oncogenic	Acts through unknown mechanism, possibly through modulation of the Hippo pathway, to target the transcriptional regulator YAP1.	(Miao et al., 2020)
HTT-ASv1	HD	Downregulated	Protective	Regulates mHTT expression in partially dicer dependant mechanism.	(Chung et al., 2011)
UCHL-AS1	PD	Downregulated	Protective	UCHL1-AS1 enhances the translation of UCHL1, loss of UCHL1 activity is reported in PD.	(Carrieri et al., 2012)
SNHG1	PD	Upregulated	Pathogenic	In neurons SNGH1 competes for miR-221/222 cluster members to indirectly regulate expression of p27/mTOR and in glia SNGH1 promotes neuroinflammation by sequestering miR-7.	(Cao et al., 2018; Qian et al., 2019)

Table 1.1: Summary of cytoplasmic lncRNAs involved in neuronal diseases and their proposed mechanism of action.

1.5 lncRNAs with novel translated open reading frames

When annotating the human genome, a set of arbitrary cut-offs were established to avoid spurious annotation of ORFs (Orr et al., 2020). However, this had the unintentional side effect of excluding thousands of potential ORFs from interest, including all sORFs as the original annotation rules excluded ORFs >300nts/100 codons in length (Figure 1.6) (Bazzini et al., 2014; Brunet et al., 2020; Yeasmin et al., 2018). Now, with the growing body of evidence supporting translation of sORFs and unannotated ORFs in lncRNAs from (Poly-)Ribo-Seq (Aspden et al., 2014; Douka et al., 2021), optimized MS (Fabre et al., 2021) and cloning experiments, translation from lncRNA (s)ORFs is gaining recognition as a true biological occurrence. Translated (s)ORFs from lncRNAs (as well as uORFs/dORFs) are collectively referred to as "non-canonical ORFs" (Prensner et al., 2023). Recent papers suggest ~22% of lncRNAs are translated (Chong et al., 2020; van Heesch et al., 2019), and so with GENECODE listing ~20,000 lncRNA genes (accessed March 2024), half of which localise to the cytoplasm, lncRNAs represent a large reservoir of potential novel (micro)proteins.

Figure 1.6: Schematic of novel ORFs found within lncRNAs. *(A) The majority of novel ORFs discovered in lncRNAs are small ORFs (sORFs) of less than 100 codons. (B) lncRNAs can also encode unannotated ORFs greater than 100 codons that have avoided annotation for alternative reasons. Figure made in biorender.*

1.6 Eukaryotic translation

Translation is a cyclical process comprised of 4 main stages: initiation, elongation, termination, and recycling. Each step of translation and the machinery required is highly conserved in eukaryotes (Schuller & Green, 2018).

Translation initiation refers to the assembly of elongation-competent 80S ribosomes along an RNA, culminating in base pairing of an anticodon loop of initiator tRNA with the initiation codon (AUG) of the translated RNA (Kapp & Lorsch, 2004). Initiation is composed of two steps: firstly, the formation of the 48S complex. This typically occurs through the scanning mechanism where the 43S pre-initiation complex (43S PIC), composed of the 40S subunit bound to initiation factors eIF1, eIF1A, eIF3 and eIF2-TC, binds the capped 5'-UTR region of an RNA to drive unwinding of any RNA secondary structure, aided by multiple eIFs including EIF4A/B/F. Secondly, once RNA secondary structure has been dismantled the 43S complex can begin scanning in the 5' to 3' direction until it reaches an initiation codon. Upon reaching an initiation codon the 43S complex is converted to the 48S complex through dissociation of eIFs allowing for the joining of a 60S subunit (Jackson et al., 2010; Kapp & Lorsch, 2004)

Elongation machinery, unlike initiation and termination, is conserved across eukaryotes, bacteria, and archaea (Kapp & Lorsch, 2004). Elongation begins once the elongation competent 80S ribosome is assembled and initiation eIFs displaced. The ribosome, loaded with an initiator fMet-tRNA^{fMet} in the P-site now accepts the first elongator tRNA and a peptide bond formed between the Met and the second amino acid. The first methionyl-tRNA, the Met which has now been attached to the second amino acid, is then released from the exit site (E-site). Elongation continues this cycle of cognate aa-tRNA entering the ribosome A-site with help of the EF-Tu, peptide bond formation between the latest aa-tRNA and the peptide being held by a tRNA in the P-site, release of the now empty tRNA from the E-site. The P-site is always occupied by the tRNA attached to the peptide chain and the nascent polypeptide chain extends by one amino acid residue with each elongation cycle. Elongation proceeds across the mRNA in a 5' to 3' in steps of 3 nucleotides until a stop codon is reached at the end of the ORF (Jackson et al., 2010; Schuller & Green, 2018).

Termination occurs through the recognition of a stop codon (UAA, UAG and UGA) by eukaryotic peptide chain release factor (eRF1). eRF1 then induces the release of the nascent peptide and now the 80S ribosome can be recycled into free 40S and 60S subunits allowing the cycle to occur again. The dismantling of the 80S ribosome is accomplished by ABCE1 which uses energy generated from ATP hydrolysis to separate the individual subunits (Schuller & Green, 2018).

For each of these steps there are non-canonical exceptions, for example translational initiation can also occur through other mechanisms, most notably through internal ribosome entry sites possessed by some mRNAs (Jackson et al., 2010). However, like canonical mRNAs, 98% of lncRNAs are spliced and ~50% possess a polyA tail (Derrien et al., 2012) and many have m^7G caps (Tsagakis et al., 2020). Therefore, it is likely that translational initiation of lncRNA ORFs, and the other stages of translation, occur along the canonical pathway described (Figure 1.7).

1.6.1 non-canonical translation initiation

Previously, it was believed that translation initiation only occurred at an AUG. However, ribosome profiling experiments have revealed that up to 10% of translation events occur at non-canonical start sites (Fedorova et al., 2022). These start sites are near-cognate triplets, most commonly CUG and GUG, and are typically recognised less efficiently than AUG, at ~1-10% of the frequency depending on the genomic context (Fedorova et al., 2022). However, translation of the protein POLGARF initiates at a CUG with an initiation efficiency of ~60-70% to that of AUG, indicating high efficiency translation initiations is possible from non-AUGs.

The use non-canonical start codons can be seen in mRNAs of critical genes including MYC and PTEN. PTEN possess multiple alternative proteoforms including three CUG and an AUU initiated proteoform. PTENs most common N-terminal extended proteoform comes from the AUU start codon, resulting in a 146 amino acid extension to canonical PTEN, which is conserved across eutherian mammals (Tzani et al., 2016). Further upstream, translation from a CUG results in a 173 amino acid N-terminal extension termed PTEN-long, which unlike regular PTEN is secreted from cells. PTEN-long can then enter neighbouring cells and

inhibit PI3K signalling (Hopkins et al., 2013). Similarly, upstream of the canonical AUG start codon for MYC is a CUG that encodes an alternative proteoform, overexpression of which inhibits the proliferation of Burkitt's lymphoma cell lines (Hann et al., 1984).

Non-AUG translation initiation in lncRNA ORFs has been reported with Ribo-Seq experiments revealing only ~50% of sORFs initiate with canonical AUG start codons (Ingolia et al., 2016, Cao et al., 2020, Lee et al., 2012). This highlights the wide spread use of non-AUG initiated translation, particularly in sORFs. While characterised examples of non-AUG are limited, the functions identified so far including regulation of translation for a downstream ORF (non-AUG uORFs) and production of novel proteins (Andreev et al., 2022). Therefore, based on the candidates so far characterised and their large number, non-AUG initiated ORFs are likely to have broad biological significance.

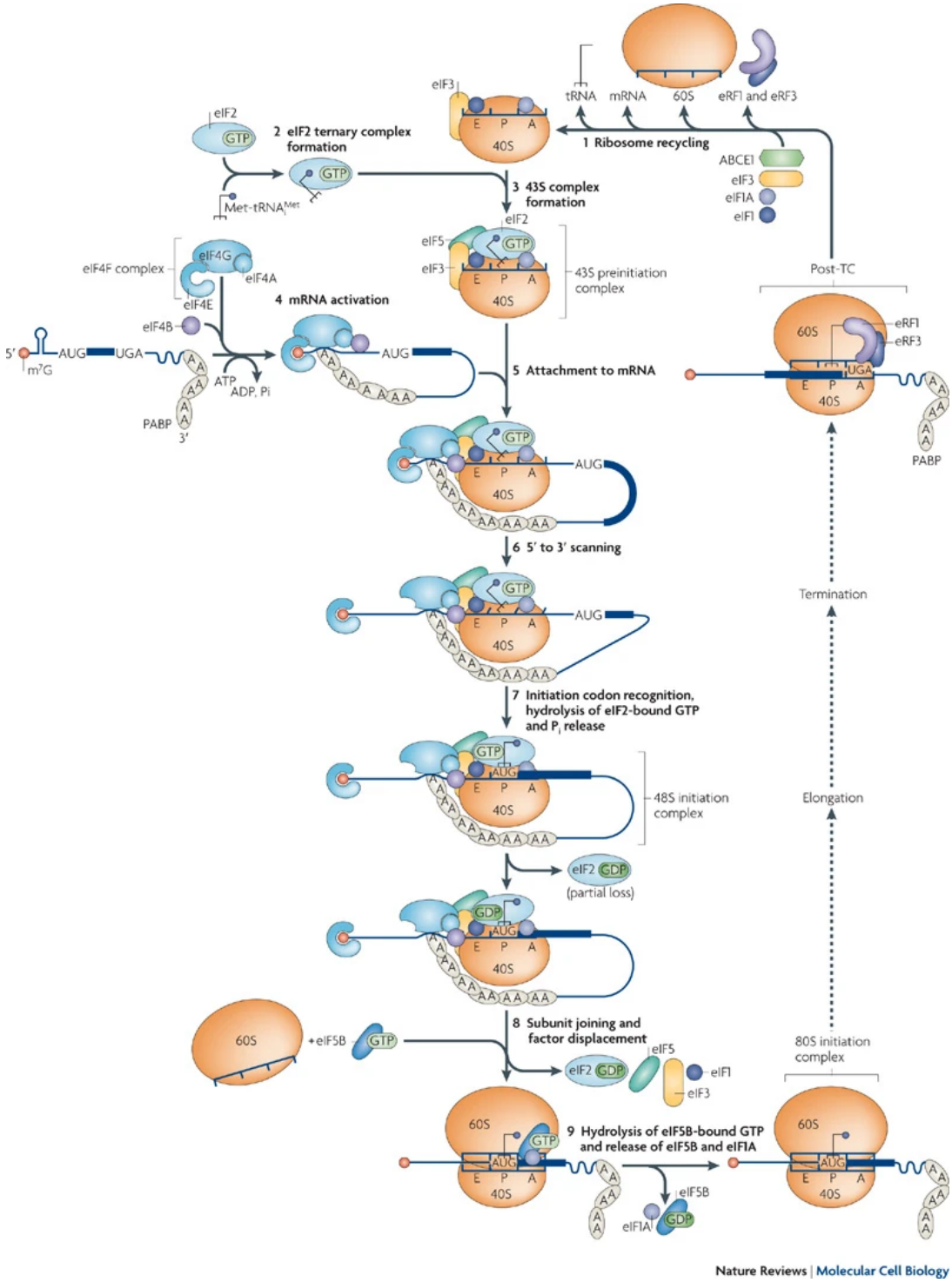


Figure 1.7: Schematic overview of eukaryotic translation. *Taken from "The mechanism of eukaryotic translation initiation and principles of its regulation"* (Jackson et al., 2010)*. Overview of canonical eukaryotic translation. (1-7) The cycle begins with 43S Pre-initiation complex formed of the 40S subunit and multiple eIFs. The 43S PIC binds and unwinds mRNA from the 5'UTR and begins scanning until an initiation codon is found. Once the initiation codon is recognised eIFs dissociate allowing for the formation of the 48S complex to form followed by requitement of the 60S subunit. Following initiation, elongation begins and can be summarised into 3 main steps, tRNA selection, peptide bond formation and translocation of tRNA/mRNA. This cycle continues until a stop codon is reached and recognised by eukaryotic peptide chain release factor (eRF1) and the nascent peptide is released and the 80S ribosome recycled into free 40S and 60S subunits ready for the cycle to begin again.*

1.7 Novel (s)ORF discovery

Open reading frame are simply an ATG with an in-frame downstream stop codon which are found distributed throughout the genomes. Short ORFs (>100 codons) are highly abundant with millions present in the human genome (Couso & Patraquim, 2017). However, only a small proportion of sORFs will be actively translated, with conservative estimates suggesting 7000 noncanonical ORFs are translated in humans, although other estimates are in the hundreds of thousands (Prensner et al., 2023). Therefore, extensive research has focused on the robust identification of translated (s)ORFs and detecting their peptide products.

The current gold-standard method for identification of noncanonical ORFs is Poly-Ribo-Seq (Aspden et al., 2014), whereby arresting translating ribosomes with cycloheximide and treating with RNaseI, short "footprints" of RNA protected within the ribosome at the point of arresting are generated. Poly-Ribo-Seq differs from Ribo-Seq as it first isolates polysomes – complexes of ribosomes undergoing active translation - from monosomes and single ribosomal subunits. Then, footprints can be sequenced to create a snapshot of actively translated RNAs (Figure 1.8). Poly-Ribo-Seq analysis also puts a strong emphasis on the "framing" of footprints, with characteristics including triplet periodicity over ORFs and a sharp drop of after stop codons giving a strong indication of transcripts undergoing active translation (Figure 1.9). Evidence from (Poly-)Ribo-Seq across multiple species suggests that hundreds of actively translated noncanonical ORFs are located within lncRNA transcripts (Anderson et al., 2015; Aspden et al., 2014; Bazzini et al., 2014; Chong et al., 2020; van Heesch et al., 2019). In the human heart, 22% of transcribed lncRNAs contain sORFs detected by Ribo-Seq, yielding microproteins with a median length of 49 amino acids (van Heesch et al., 2019). This finding was mirrored in the context of cancer where ~22% of transcribed lncRNAs in melanoma cells show evidence of being translated, with 55% of these verified by mass spectrometry (MS) (Chong et al., 2020).

Despite constant improvements to (Poly-)Ribo-Seq, critics are sceptical of its accuracy and question the low levels of microproteins detected by MS investigations. A clear explanation for MS poor ability to pick up microproteins is that their small size and low expression levels in comparison to canonical proteins make microproteins difficult to detect (van Heesch et al., 2019). This is due to limitations across the entire MS workflow including: (1) Protease

digestion; small proteins may not contain suitable cleavage sites and therefore will not be detected; (2) reference proteomes often do not contain predicted small or novel proteins so even if present peptides will have nothing to align to; (3) a “two-peptide rule” is often implemented in MS workflows where at least 2 unique peptides need to be identified for confident protein detection. Due to microproteins small size this threshold often is not reached; (4) MS is “amplification-free” and therefore is limited by protein abundance and detection is intrinsically bias towards the most abundant proteins (Ahrens et al., 2022). This likely explains some of the discrepancy between Ribo-Seq and MS ability to detect novel small proteins (Bazzini et al., 2014).

To address these issues the new field of “peptidomics” is developing new analytical workflows to analyse the presence of peptides from novel (s)ORFs. Alterations at every level of the MS pipeline are being implemented including optimising extraction of small proteins from cells, increasing the sensitivity if mass-spectrometers and custom bioinformatics to map small proteins. Each of these alterations will enrich for specific subsets of microprotein depending on their biochemical properties and subcellular localisation. Widespread use of these techniques will allow for a greater pool of data for the microproteins discovered by Rib-Seq to be validated with (Slavoff et al., 2013).

Another key consideration for lncRNA-translation is that active translation does not necessarily indicate that a functional microprotein is produced. Pervasive translation of noncanonical ORFs, regardless of function, has been proposed (Ruiz-Orera et al., 2018). Translation may be a mechanism for controlling lncRNA transcript levels through NMD with micropeptide production being a side product (Malabat et al., 2015). Furthermore, there is little guarantee that the (micro)proteins produced are stable, especially when considering that due to their small size, microproteins are particularly vulnerable to degradation in the cytosol (Makarewich, 2020). However, with multiple functional lncRNAs-encoded microproteins having been characterised this it seems unlikely that translation of lncRNAs is primarily a mechanism for controlling transcript levels.

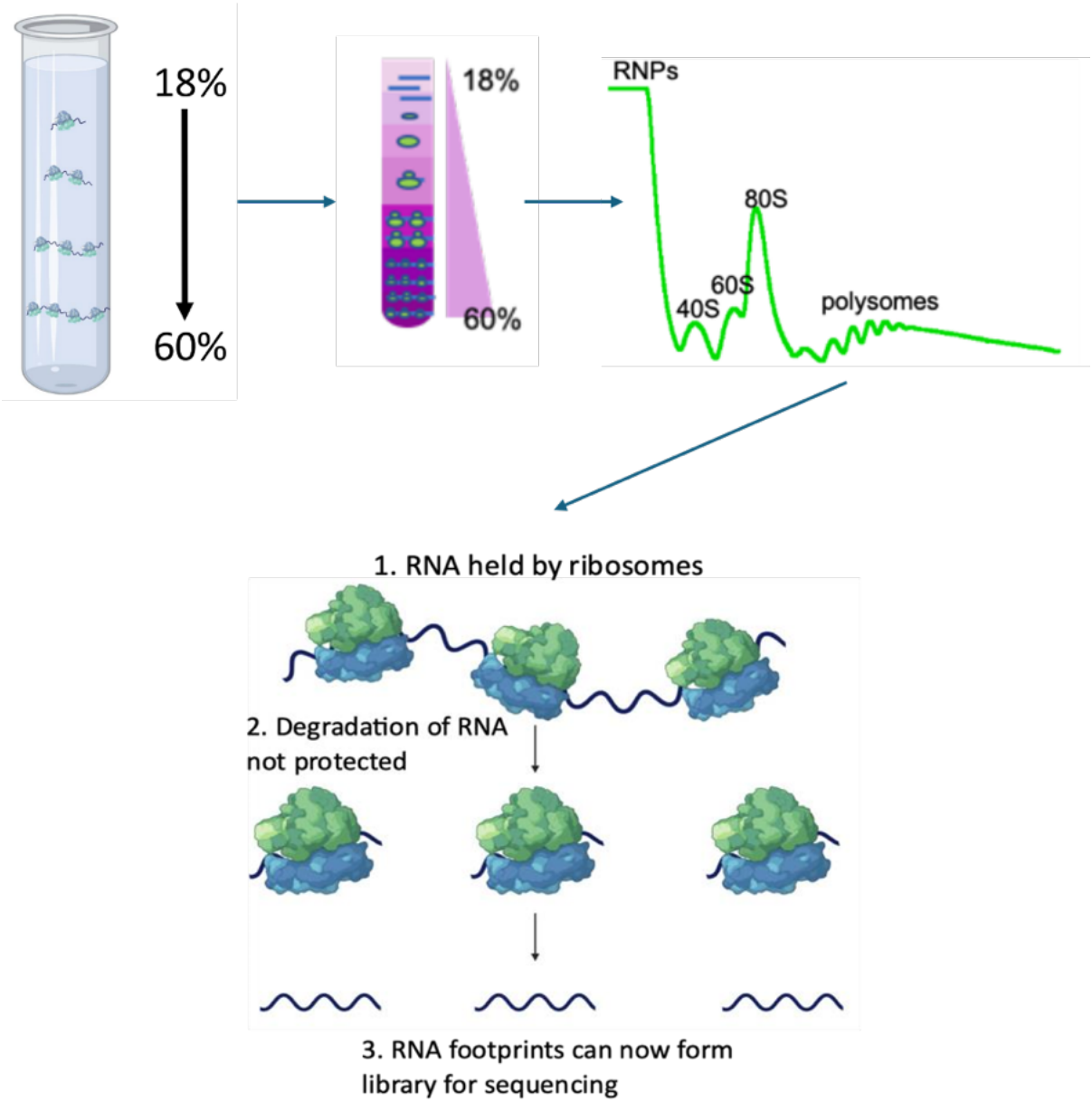


Figure 1.8: Schematic overview of Poly-Ribo-Seq. *Cell lysates are loaded onto a sucrose gradient (18% - 60%) before ultracentrifugation (121,355 x g_{avg} for 3.5 h, 4°C). More dense complexes will progress further through the gradient resulting in polysomes reaching the lower layers followed by disomes, monosomes, 60S and 40S subunits further up the gradient. Therefore, polysomes can then be isolated and Ribo-Seq performed on them. In brief Ribo-Seq involves degrading RNA not protected by ribosomes with ribonucleases (e.g. RNaseI). Ribosomal protected fragments of ~30 nucleotides can then be sequence to identify RNAs undergoing translation. Created in biorender.*

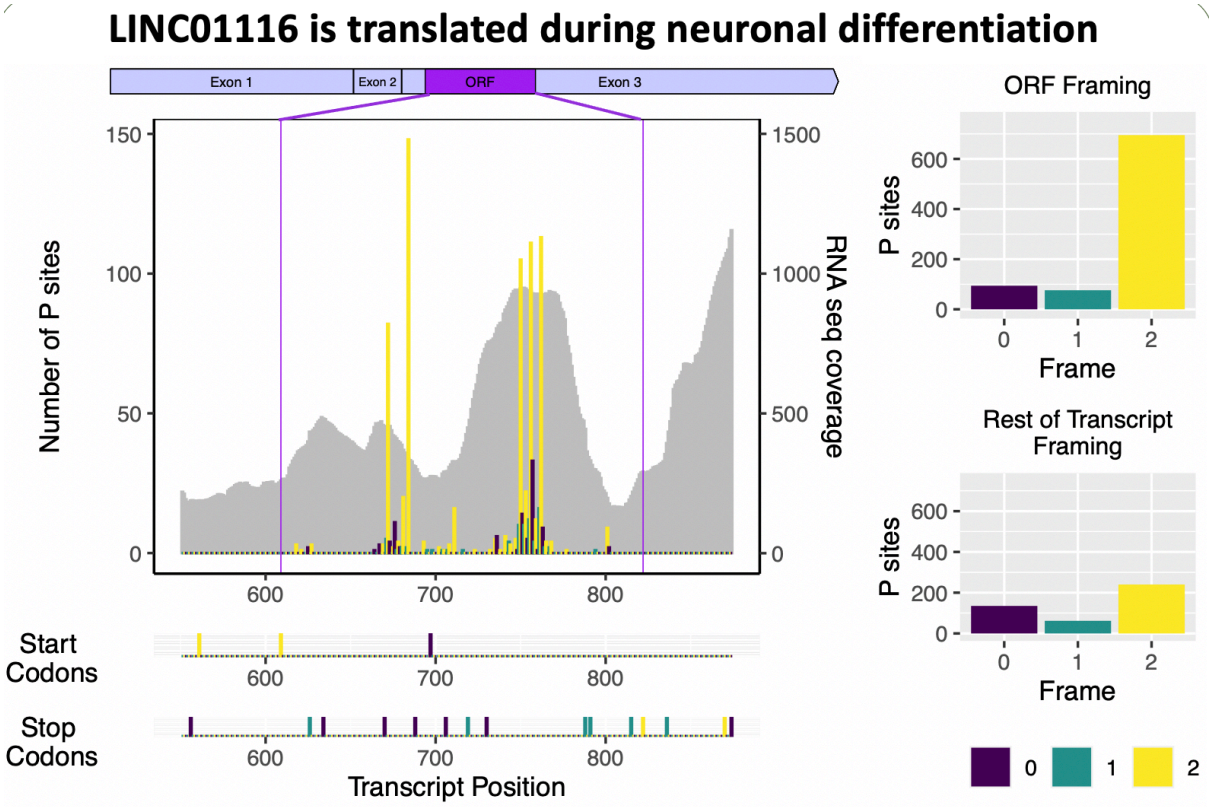


Figure 1.9: Poly-Ribo-Seq plot (produced by Dr Isabel Birds, Aspden group) showing Poly-Ribo-Seq data indicating that LINC01116 contains a translated sORF*. P-sites are shown as bars, color corresponds to frame. Start and stop codons and their frame are annotated below the main plot. Start and stop codons called by RiboTaper are highlighted by purple bars aligning to the transcript model above. RNA-seq reads are shown in gray. The plot shows that the majority of P-sites are within the sORF of the transcript and they show a strong bias towards frame 2 (yellow). This bias that is lost outside of the sORF* (Douka et al., 2021).

1.8 Evolution of long non-coding RNAs

lncRNAs are found in all organisms from bacteria to humans and the number of lncRNAs expressed in an organism positively correlates with its complexity, as defined by an organism's level of metabolic and development complexity e.g., the number of different cell types and how highly organized they are (Figure 1.10) (Taft et al., 2007). However, compared to protein-coding genes, lncRNAs show lower sequence conservation. This observation initially led to the dismissal of lncRNAs as transcriptional noise but importantly, lncRNAs are subject to greater selection pressure than neutrally evolving sequences such as intronic DNA. This indicates that despite evolving faster than protein-coding genes lncRNAs are under selective constraints (Necsulea et al., 2014). Additionally, due to the short average length of lncRNAs compared to mRNAs, conventional bioinformatic techniques often fail to detect what conservation is present (Herrera-Úbeda et al., 2019). This issue is exacerbated by the fact that lncRNA functions are often derived from small regions that mediate their interactions with proteins, RNA or DNA. In contrast to the coding regions of mRNAs, these critical domains can be as short as 10 nucleotides and may be sparsely dispersed within a lncRNA sequence, connected by "linker" sequences in which SNPs or small deletions are unlikely to impact molecular function (Herrera-Úbeda et al., 2019). Additionally, for truly non-coding lncRNAs, complex secondary structures are often more important for function than their underlying sequence, potentially making lncRNAs less sensitive to sequence changes. Therefore, bioinformatic techniques that analyse sequence variation only, while not considering overall RNA structure, will likely miss conserved lncRNAs. Xist is a prime example of this, as it performs X chromosome inactivation in all placental mammals, yet its orthologs show only low levels of sequence homology (Yen et al., 2007).

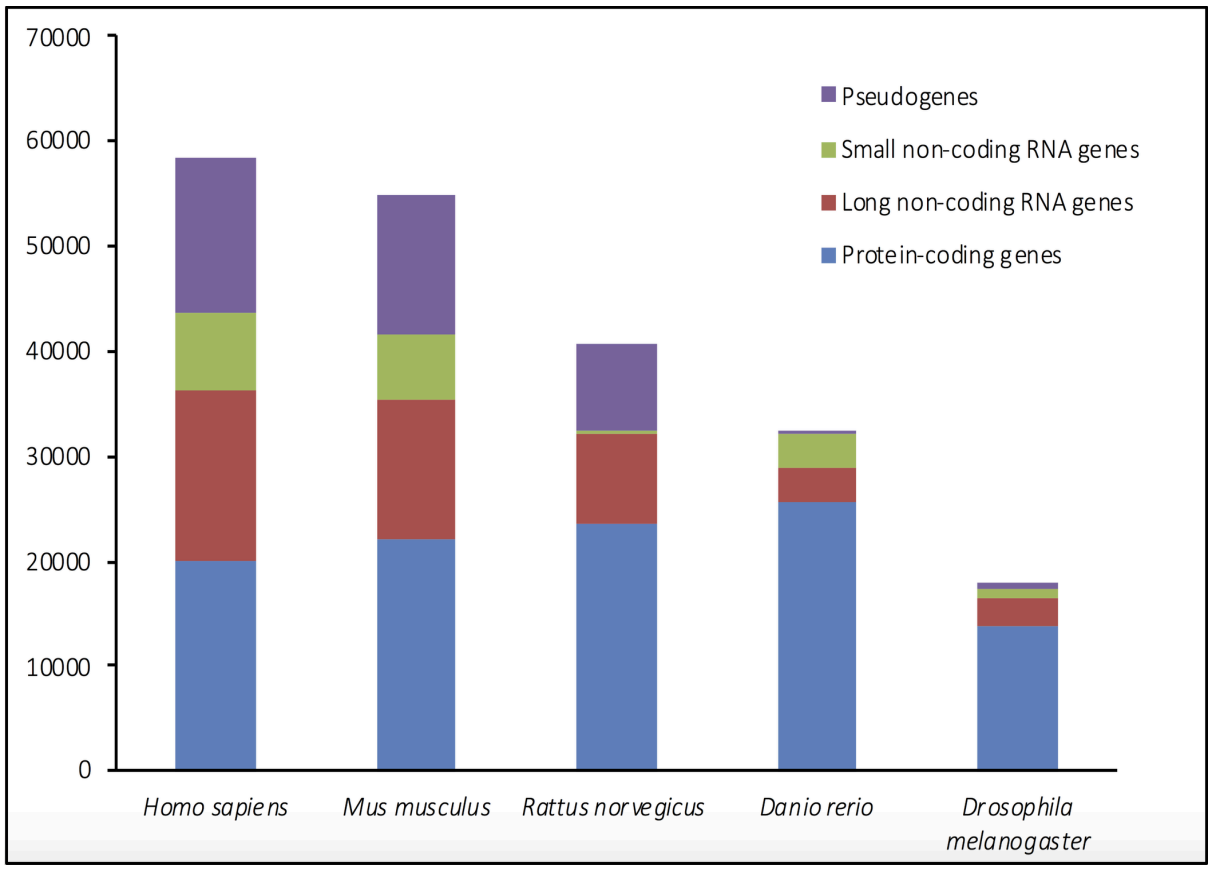


Figure 1.10: Proportion of genes in different species genomes*. Numbers of lncRNAs genes (in red) increases with organism complexity. Figure from* (Tsagakis et al., 2020).

1.8.1 Evolution of non-canonical peptides

The low sequence constraint observed for lncRNAs has been proposed to allow for the de-novo birth of protein-coding genes as many lncRNAs are not under any selection pressure so show neutral evolution. This combined with the pervasive translation of (s)ORFs allows for novel peptides to be produced and modified. If a lncRNA mutation then leads to a functional peptide it will be selected for and be conserved (Figure 1.11). Therefore, the moderate overall conservation of lncRNAs may, be in part due to an averaging of lncRNAs still neutrally evolving and lncRNAs that have gained function through translation of a (s)ORF or an RNA function (Ruiz-Orera et al., 2018). Additionally, as lncRNAs allow for the de-novo protein birth it is likely that many functional lncRNAs are evolutionarily recent and constrained just to humans and their closest relatives.

Analysis of novel sORF encoded proteins in humans revealed ~90% of sORF encoded microproteins lacked significant sequence homology to non-primate species, supporting the proposal that novel microproteins are evolutionarily young (Sandmann et al., 2023). This is a clear distinction from canonical proteins where ~70% show conservation beyond primates. The remaining 10% of novel sORF derived peptides show conservation across non-primate mammals, a significantly greater proportion than seen in the negative control composed of length matched UTR sequences where only 1% were conserved suggesting some “novel” (s)ORFs are conserved and present in non-primate mammals (Sandmann et al., 2023). Additionally, non-canonical proteins “young” status may partially explain the sub-optimal codon and non-canonical amino acid usage seen in lncRNA derived micropeptides (Patraquim et al., 2022). Specifically, microproteins on average contain more sulphur-containing amino acids, are enriched in positively charged and depleted in negatively charged amino acids. This appears consistent across microproteins identified in flies, mice, and humans (Couso & Patraquim, 2017). LncRNA ORFs on average contain more codons for rare tRNAs than canonical ORFs, perhaps indicating that translated lncRNAs have not been fine-tuned for coding efficiency as are evolutionarily young. The use of rare tRNAs is detrimental as can delay translation (Hershberg & Petrov, 2009; Patraquim et al., 2022).

Another characteristic of evolutionally young proteins is a deposition to be highly disordered. *De novo* gene birth from “non-coding” RNA has the potential to be deleterious as random peptide sequences are expected to be toxic. Novel peptides could act in a similar

manner to amyloids, resulting in uncontrolled protein aggregation and insoluble fibril formation with the potential to contribute to result in cell harmful phenotypes and eventually disease (Wilson et al., 2017). Therefore, the *de novo* birth of random peptides could be highly dangerous to an organism. A proposed explanation to avoid toxic peptides is through the preadaptation hypothesis where *de novo* genes originate with characteristics to avoid toxicity. Novel proteins from de novo genes have significantly higher degrees of intrinsic structural disorder, being inherently flexible and unstructured, and therefore are less likely to form toxic amyloids (Wilson et al., 2017).

Furthermore, disordered proteins evolve faster, perhaps reflecting their *de novo* origin and rapid evolution once established. The disordered nature of young proteins is primarily due to amino acid composition suggesting *de novo* gene birth is driven by amino acid composition with a short proteins being enriched in arginine and lysine but are depleted in aspartic acid, and glutamic acid (Couso & Patraquim, 2017). Additionally, on average young genes have significantly higher intrinsic structural disorder than sequences translated from randomly chosen junk DNA as well as when compared to canonical proteins again demonstrating young genes tendency to be intrinsically disordered (Wilson et al., 2017).

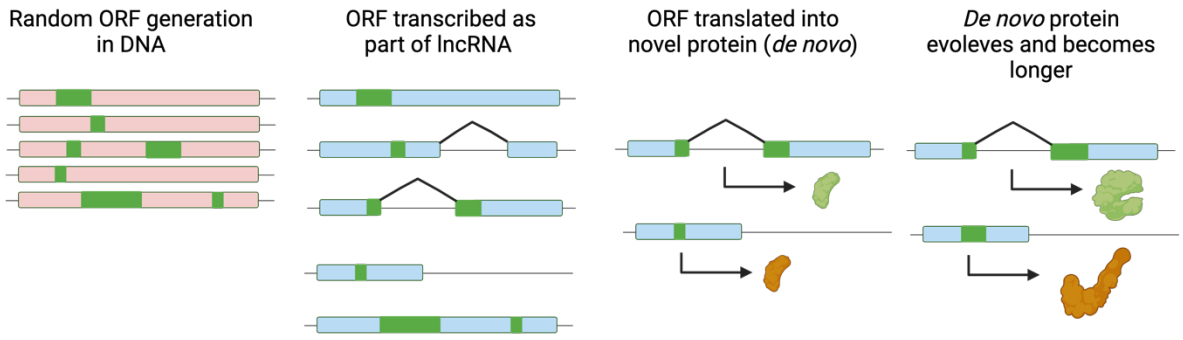


Figure 1.11: Schematic showing how lncRNAs can give rise to de novo proteins. Adapted from Classification and function of small open reading frames. *In frame start and stop codons are abundant throughout the genome. Of the millions of ORFs present outside of mRNA genes, many will be transcribed as part of a lncRNA. If within a cytoplasmic lncRNA with mRNA-like properties, the ORF may undergo translation to produce a novel protein. Novel proteins will then be subjected to natural selection. The low evolutionary constraint on many lncRNAs allows for mutation of novel protein, allowing them to refine their protein function (Couso & Patraquim, 2017).*

1.9 Functions of non-canonical peptides

Non-canonical peptides refer to any protein encoded from an RNA not annotated as an mRNA. The majority of these are small proteins >100 amino acids long, termed microproteins, and are encoded by small ORFs (sORFs). Microproteins differ from other small proteins with known functions such as Neuropeptide Y (Tatemoto et al., 1982) and Neuropeptide substance P (Nichols et al., 1999) as are translated directly as mature proteins instead of being derived from post-translational cleavage of a larger precursor protein (i.e., preproproteins) (Hassel et al., 2023).

sORFs, and therefore microproteins, were initially excluded form genome annotations due to calculated probabilities highlighting ORFs >300 nucleotides long are far more likely to generate stable proteins than sORFs (Basrai et al., 1997). However, recent investigation has revealed microproteins to have critical roles in both human physiology and pathology (Senís et al., 2021, Anderson et al., 2015, Miller et al., 2023).

1.9.1 Non-canonical peptides in neuronal physiology

As mentioned previously, ~40% of all human lncRNAs are specifically expressed in the brain (Derrien et al., 2012). If only a 10% fraction of these lncRNAs are translated into novel proteins lncRNAs could still account for ~800 novel neuronal proteins. To date only a few examples have been identified and characterised. FLJ33706 was first identified as potentially functional in a GWAS studies which identified a SNP associated with susceptibility to nicotine addiction in its 3'UTR (C.-Y. Li et al., 2010). Western blot in multiple human brain regions including midbrain and cerebellum detected the 194aa FLJ33706 protein. Although the exact function of FLJ33706 in the human brain is unknown, it is upregulated in Alzheimer's disease. Additionally, FLJ33706 has no identifiable homologues in any species suggesting that FJL33706 formed *de novo* in humans after the divergence from chimpanzee. Other novel proteins have also been discovered in GWAS studies with a heterozygous microdeletion (loss of one allele) at the genomic location of the lncRNA-derived microprotein Pants conveying a high risk for schizophrenia. When endogenously tagged in mice Pants-HA was identified in the CA2 and CA3 regions of the hippocampus and the molecular layer of the dentate gyrus where it is secreted from neurons in large dense-core

vesicles, a common pathway for neuropeptide release. Heterozygous (but not homozygous) deletion of Pants from mice resulted in greater extent of long-term potentiation (LPT) of hippocampal mossy fibers (Kragness et al., 2022).

While the majority of non-canonical ORFs are not conserved beyond great ape, functional micropeptides have been discovered with extensive conservation. One such highly conserved lncRNA-derived peptide is the 48 amino-acid microprotein (pTUNAR) originally discovered in zebrafish under the name *Megamind* (Ulitsky et al., 2011). TUNAR RNA is highly expressed in the human brain and IF and western blot can detect widespread pTUNAR expression in both adult mice brains and mouse embryonic stem cells (mESCs) (Senís et al., 2021). Knockdown of pTUNAR by shRNA prevents hESCs from undergoing neuronal differentiation(N. Lin et al., 2014). Functionally, over expression of pTUNAR in neurons derived from mESCs results in significantly increased calcium release in response to caffeine. Additionally, calcium re-uptake by the endoplasmic reticulum is significantly increase, likely explaining the greater response to caffeine as the ER will contain more calcium with pTUNAR overexpressed. This dysregulation of neuronal calcium dynamics inhibits the formation of neurites in different multiple models. Mechanistically, pTUNAR localises to the ER where Co-IP experiments have shown a pTUNAR to interact with the calcium pump SERCA2 (Figure 1.12) (Senís et al., 2021).

The 54 aa microprotein PIGBOS interacts with the ER protein CLCC1 to regulate the unfolded protein response (UPR). UPR dysfunction is known to contribute to multiple neurodegenerative disorders including Alzheimer's disease, Parkinson's disease, and Huntington's disease due to the accumulation of pathological, misfolded proteins (Hetz & Saxena, 2017). PIGBOS was initially identified through RNA-Seq and Ribo-Seq of HEK293T, HELA and K562 cell lines and shows high sequence conservation back to mouse. PIGBOS has been shown to be essential for inter-organelle communication, homeostasis, and cell survival. Exogenously expressed PIGBOS-FLAG localises to the outer mitochondrial membrane when it interacts with the ER protein CLCC1, as demonstrated by co-immunoprecipitation of PIGBOS-FLAG from HEK293 cell lysate. Knockdown of PIGBOS, by both siRNA KD or CRISPR-Cas9, in HEK293 cells leads to in an increase in the UPR and cell death (Figure 1.12) (Chu et al., 2019).

Micropeptides encoded from mitochondrial DNA have also been implicated in contributing to neuronal disorders including neurodegenerative disease and mental health conditions. SHMOOSE is encoded from a mitochondrial gene and localise specifically to neuronal mitochondria inner membrane in the human brain. SNPs within SHMOOSE that result in an amino acid substitution that is associated with Alzheimer's disease and pathological brain structure suggesting that the novel peptide SHMOOSE is essential for normal brain function and mutations may contribute to Alzheimer's pathology (Miller et al., 2023).

1.9.2 Micropeptides in cancer

RNA-Seq of disease states have identified an extensive number of lncRNAs with differential expression in virtually all pathologies, but to date, the field of cancer biology is leading the way for identification of pathological microproteins. In colorectal cancer models, the microproteins FORCP and RBRP, derived from putative lncRNAs, have both been shown to influence tumorigenicity. FORCP is expressed in healthy, well differentiated colorectal cells but is downregulated in cancer (X. L. Li et al., 2020). Overexpression of FORCP in HCT116 (colorectal carcinoma) cells has been shown to slow proliferation, but this effect is attenuated by mutagenesis of the sORFs start-codon, indicating that it is the peptide that is functional in this context (X. L. Li et al., 2020). Conversely, expression of RBRP correlates with an increased risk of death in colorectal cancer patients. *In vitro,* overexpression of RBRP increases cancer cell proliferation and invasion, even when the transcript it is derived from (LINC00266-1) is knocked-down (S. Zhu et al., 2020). Co-immunoprecipitation assays show that RBRP binds the mRNA binding protein IGF2BP1, among others. This interaction enhances the mRNA binding efficiency of IGF2BP1, allowing it to stabilise mRNAs of factors including c-Myc, thus explaining RBRPs role in tumorigenesis. Interestingly, RBRP shows no sequence conservation or homology to known protein domains, again highlighting the difficulties in identifying translated and functional sORFs (Figure 1.12) (S. Zhu et al., 2020).

The enrichment of novel translation events in cancer may, in part, be due to the dysregulation of translational control, which is a hallmark of cancer. Ribo-Seq performed on medulloblastoma biopsies and cell lines revealed the translation of 7,530 noncanonical ORFs with 6,740 of these being detected in at least 5 samples. Targeting 2,019 of these novel

ORFs with a CRISPR screen revealed 387 ORFs were required for cancer cell viability with 121/387 being required for viability in ≥2 cell lines (Hofman et al., 2024). This indicates these noncanonical ORFs are contribute to cancer cell viability, potentially due to their translation into functional proteins.

Similarly, in a pan-cancer noncanonical ORF CRISPR screen, 553 candidates were randomly selected from noncanonical ORF datasets. CRISPR was used to target ORFs in 8 different human cancer cell lines. Knockout of ~10% for noncanonical ORFs produced a viability phenotype in ≥2 cell types (Prensner et al., 2021). This is compared to ~17% of canonical genes that produce a viability phenotype in ≥2 cell types. Following the initial screen a secondary sgRNA screen using dense tiling of sgRNAs across the genomic locus of each ORF. Of the genes tested, 44% only produced a phenotype with sgRNAs targeting the predicted protein coding region, including OLMALINC and LINC001116 which were detected as translated in Poly-Ribo-Seq of SH-SY5Y cells (Douka et al., 2021; Prensner et al., 2021).

1.9.3 Membrane microproteins

A sub-group of micropeptides receiving particular attention is that of membrane microproteins, as association with biological membranes is suggestive of functionality and would offer micropeptides with protection from degradation. Evidence for micropeptides as membrane components includes the observation that many are predicted to contain a transmembrane α-helix motif (Aspden et al., 2014). Additionally, micropeptides are thought to engage with larger protein complexes to be functional, hence cell membranes would provide microproteins with an array of multiprotein complexes to engage with (Makarewich, 2020). For example, Mitoregulin is a lncRNA-encoded micropeptide that despite lacking an obvious mitochondrial localisation signal is found on the inner mitochondrial membrane. Mitoregulin knockout mice show reduced supercomplex formation, which is required for optimal electron transfer, and fatty acid oxidation (Stein et al., 2018). Furthermore, the sarco/endoplasmic reticulum calcium ATPase (SERCA) is target for a family of inhibitory micropeptides including phospholamban (PLN), sarcolipin (SLN) and myoregulin (MLN) (Anderson et al., 2015; MacLennan & Kranias, 2003; Pant et al., 2016). MLN, is derived from a sORF within lncRNA LINC00948. Overexpression of both the full-length transcript and the sORF alone in HEK293 cells inhibited the reuptake of Ca^{2+} by

SERCA. This effect was abolished by introduction of a frameshift mutation that disrupted the protein coding ability of MLN. Additionally, knockout of MLN in mice significantly enhanced muscle performance (Anderson et al., 2015).

Overall, while the proportion of (s)ORFs that are actively translated to produce biologically active micropeptides is still debated, if this is significant, they represent an area of importance for future research with the potential to greatly expand the proteome. Additionally, the identification of novel proteins in disease states could offer new therapeutic targets and biomarkers in a range of diseases (C. Zhang et al., 2022).

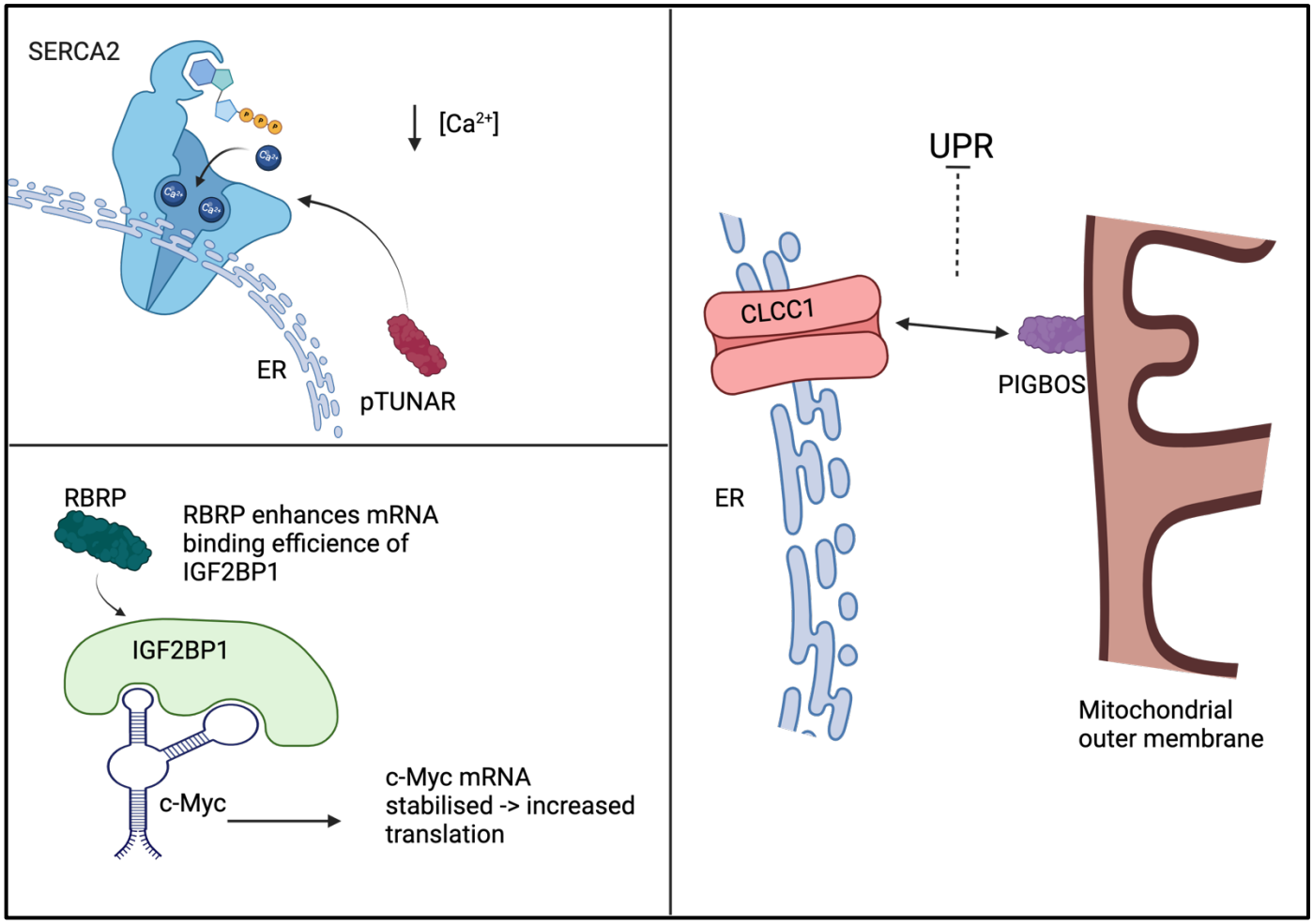


Figure 1.12: Schematic showing function of 3 micropeptides: pTUNAR, RBRP and PIGBOS. *pTUNAR interacts with SERCA2 at the ER where it increases calcium reuptake* (Senís et al., 2021)*. RBRP binds the RNA binding protein IGF2BP1 to increase its RNA binding efficiency. IGF2BP1 then binds multiple mRNAs including c-Myc increasing c-Myc translation* (S. Zhu et al., 2020)*. PIGBOS localises to the outer mitochondrial membrane where it interacts with ER protein CLCC1 at ER–mitochondria contact sites to reduce the unfolded protein response* (Chu et al., 2019).

1.10 SH-SY5Y cells as a model for neuronal differentiation

SH-SY5Ys are a widely used model of neuronal differentiation. These cells are derived from three successive subclones of the parental cell line SK-N-SH which originated from a metastatic neuroblastoma tumour in the bone marrow (Kovalevich & Langford, 2013). Neuroblastoma is a cancer of the neuroectodermal origin, therefore SH-SY5Y cells under standard culture conditions phenotypically resemble neuronal progenitor cells. Despite multiple subcloning's SH-SY5Y populations are still composed cells with two distinct phenotypes: neuroblast-like (type "N") and a smaller proportion of epithelial-like cells (type "S") (Kovalevich & Langford, 2013). These populations show distinct unique features with only the type N cells expressing catecholaminergic neuronal markers including tyrosine hydroxylase (TH) and dopamine-β-hydroxylase. However, type N and S can undergo bidirectional switching from one phenotype to the other (Cohen et al., 2003).

The most useful properties of SH-SY5Ys are their ability to "differentiate" into a neuronal-like phenotype characterised by the upregulation of mature neuronal marker genes, the extension of neurites (precursors to axons and dendrites) and the slowing of proliferation and eventual arrest of differentiated cells in G0 and G1 (Şahin et al., 2021). Therefore SH-SY5Y can be used to modelling key physiological steps in the maturation of NPC to neuron or to generate mature neuronal like cells to investigate disease states (Dravid et al., 2021; Şahin et al., 2021). In particular SH-SY5Ys are used to model for Parkinson's disease due to their catecholaminergic neuronal attributes once differentiated (Krishna et al., 2014; Pandey et al., 2024). Additionally, SH-SY5Ys have been used extensively to investigate tauopathies (M. Bell & Zempel, 2022). While multiple methods for inducing SH-SY5Y differentiation are described the most widely used treatment with all-trans retinoic acid (RA) (Forster et al., 2016; Kovalevich & Langford, 2013). RA is a diffusible molecule generated from vitamin A metabolism and is involved in multiple biological processes including patterning of the forebrain.

RA together with sonic hedgehog (SHH) and bone morphogenetic proteins (BMP), regulates the patterning of the dorsoventral axis of the neural tube (Maden, 2007; Tao & Zhang, 2016). Additionally, RA induces the differentiation of multiple subtypes of neurons and glia. For RA to function it must first entre the nucleus through the action of CRABP2 (Cellular

Retinoic Acid Binding Protein2). In the nucleus RA binds and activates a transcription complex comprised a heterodimer of the RA receptor (RAR) and retinoic X receptor (RXR), of which there are 3 of each. The heterodimer pair then bind DNA at retinoic acid-response element (RARE) modulating the expression of > 500 genes, although some of these differentially expressed gene will be due to non-RARE activities (Maden, 2007). RA signalling upregulates the expression of transcription factors, cell signalling molecules, structural proteins, enzymes, and cell-surface receptors, including key genes such as components of the WNT signalling pathway, SOX1/6, and MAP2 (Shibata et al., 2021)

Overall, SH-SY5Ys are a powerful and widely utilised model to investigate neuronal differentiation, despite their tumorigenic origin. SH-SY5Y key strengths as a model include that they are of human origin, are easy to handle and can be expanded for high throughput methods or techniques that require lots of material, such as Poly-Ribo-Seq (Dravid et al., 2021).

1.11 Project objectives

lncRNAs have been identified as key players in neuronal development and disease. However, until recently the role of lncRNAs exported to the cytoplasm and subsequently translated has not been investigated in the context of human neuronal development.

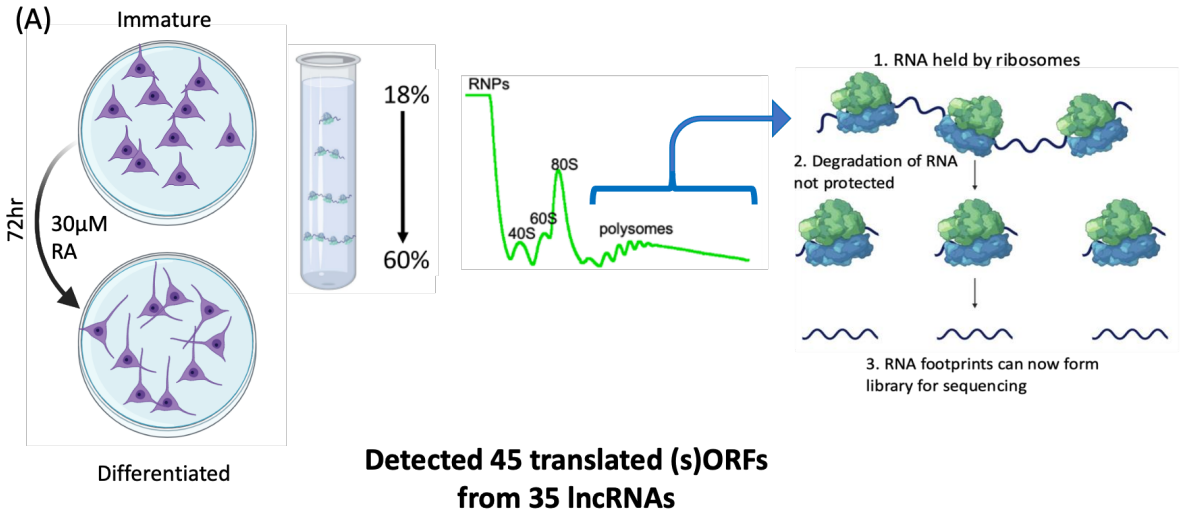
Previous work within the Aspden group by Dr Katerina Douka and Dr Isabel Birds identified cytoplasmic lncRNAs translated in neuronal differentiation (Douka et al., 2021). Using Poly-Ribo-Seq of immature and differentiated (RA treated) SH-SY5Ys, 45 sORFs from 35 annotated lncRNAs were identified as translated (Figure 1.13 A) (Douka et al., 2021).

Of the 45 translated sORFs, 42 were novel and only one, CRNDE, had been characterised at the micropeptide level (Szafron et al., 2015). Therefore, this provides a population of potentially functional micropeptides to investigate in the context of neuronal development and disease. The translated lncRNA ORFs showed translation efficiency similar to that of translated ORFs, were identified evenly across immature and differentiated SH-SY5Y (Figure 1.13 B) and were primarily sORFs with only 4/45 ORFs being longer than 300 nucleotides (Figure 1.13.C).

Initial characterisation focused on 2 candidates, LINC01116 and LINC00478. FLAG-tagged overexpression assays validated their translation and protein products were identified to localised to the cell periphery, specifically clustering at neurites (LINC01116) and in puncta in both the nucleus and cytoplasm (LINC00478). siRNA knockdown of LINC01116, but not LINC00478, significantly inhibited SH-SY5Y neurite outgrowth following RA treatment (Douka et al., 2021)

This thesis aims to determine which of the translation events described in Douka et al produce stable, functional novel peptides and elucidate their role in early neuronal differentiation. This will be achieved through these objectives:

1) Identification of which translated lncRNAs are most likely to have physiological and pathological roles through analysis of publicly available data.
2) Validation of translation from the identified lncRNA ORFs, and to characterise their protein products through FLAG tagged reporter assays, immunofluorescence, and western blot. Additionally, validate the expression and translation of our lncRNAs in an advanced model of human neuronal development will be investigated.
3) Determine the contribution of these translated lncRNAs to neuronal differentiation. Data collected from publicly available data and the reporter assays will be used to select candidates for an siRNA knockdown screen in an SH-SY5Y cells. Differentiation will be assayed through neurite extension assays and RT-qPCR of mature neuronal and pluripotency markers.
4) In-depth characterisation of a single candidate, LIPT2-AS1-peptide, to elucidate its role in neuronal differentiation and begin understanding its functional mechanism through CRISPR editing, RNA-Seq and MS to identify interaction partners.



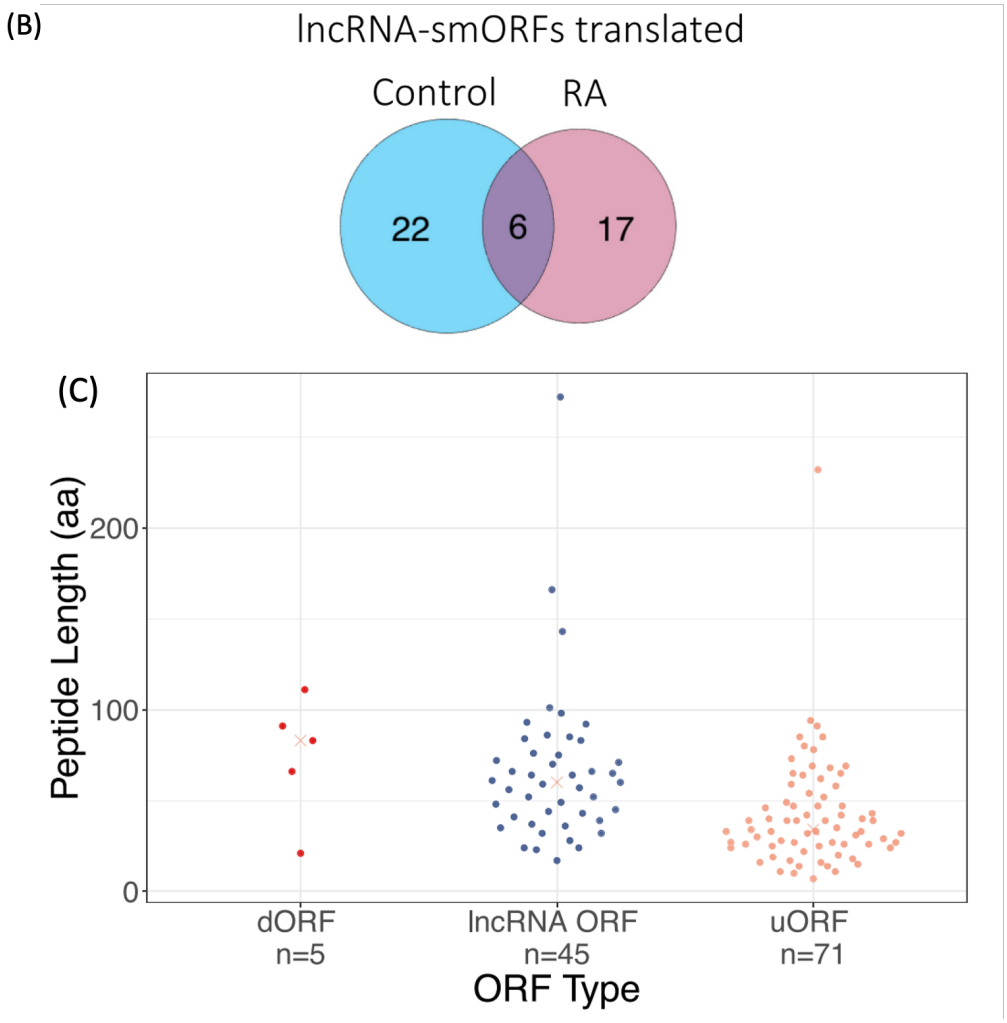


Figure 1.13: Summary of most prevalent data from Douka et al for this thesis. *(A) Schematic of ORF discovery in immature and differentiated SH-SY5Y cells. Cell were lysed and polysomes isolated and Ribo-Seq performed on them. Poly-Ribo-Seq identified 45 translated ORFs from 35 annotated lncRNA genes. (B) what condition translated ORFs were identified in. (C) Size distribution of dORFs, uORFs and lncRNA ORF identified.*

Chapter 2: Materials and Methods

2.1 Publicly available data

2.1.1 lncRNA gene disease relation

The disease databases lncRNADisease, Differential Expression Atlas and Cancer RNA-Seq Nexus contain RNA-Seq data comparisons between healthy and disease states (J.-R. Li *et al.*, 2016; Bao *et al.*, 2019; Papatheodorou *et al.*, 2020). These datasets were searched for the 35 translated lncRNAs genes (Douka et al., 2021). Genes were defined as having "disease relation" if they showed significant differential expression >1 $\log_2$ fold in diseased samples compared to control, or if a lncRNAs role in disease has been experimentally validated (Table 2.1).

Group	Diseases included
CNS cancer	Medulloblastoma, Astrocytoma, Glioblastoma, Oligodendroglioma, Glioma, Ependymoma
Neurodegenerative disease	Alzheimer's disease, Huntington's disease, Amyotrophic Lateral Sclerosis
Neurodevelopmental	Autism, Downs syndrome
Mood disorder	Borderline personality disorder

Table 2.1: Conditions included in each "disease group".

2.1.2 lncRNA expression during human development

lncExpDB contains RNA-Seq data from multiple human organs at different stages of development (Cardoso-Moreira et al., 2019; Sarropoulos et al., 2019). This data had been analyzed using the R package maSigPro, an R package that analyses transcriptomics time-courses to identify developmentally dynamically lncRNAs using the threshold of $R^2 > 0.3$.

2.1.3 LncAtlas – nuclear and cytoplasmic distribution of lncRNAs

LncAtlas contains RNA-Seq data of nuclear and cytoplasmic fractions from multiple cell lines (Mas-Ponte et al., 2017). Data from the SH-SY5Y precursor cell line, SK-N-SH, was collected for the 28 translated lncRNAs found in lncAtlas and 88 polysome associated but not translated lncRNAs as a control.

2.1.4 RNA expression across tissue

RNA-Seq data in transcript per million (TPM) for translated lncRNAs was accessed through the Illumina Body Map which contains RNA-Seq from 3 studies covering 16 tissues (Asmann et al., 2012; Barbosa-Morais et al., 2012; Derrien et al., 2012).

2.2 Novel peptide structural predictions

2.2.1 ColabFold protein structure prediction

Peptide sequences were inputted to ColabFold (Mirdita et al., 2022) in FASTA format and baseline parameters were used (Table 2.2). PDB files opened and viewed in Mol* 3D Viewer (Sehnal et al., 2021).

Parameters	Setting
num_relax	0
template_mode	none
msa_mode	mmseqs2_uniref_env
pair_mode	unpaired_paired
model_type	auto
num_recycles	3
recycle_early_stop_tolerance	auto
relax_max_iterations	200
pairing_strategy	greedy
max_msa	auto
num_seeds	1

Table 2.2: Parameters for ColabFold.

2.2.2 Phyre2

Phyre2 was also used with all amino acid sequences added in a batch and investigated under normal stringency (Modelling Mode: Normal) (Kelley et al., 2015).

2.2.3 EMBL-EBI protein sequence similarity search (SSS)

All translated lncRNA protein sequences were assessed for sequence homology to know protein domains using the online server SSS with default settings (Protein database: UniProtKB/Swiss-Prot, Program: FASTA) (Madeira et al., 2022).

2.2.4 Tis-Transformer

Tis-Transformer was used in browser with default settings (rank (k): 0-1, TIS matches on transcript: >0) to predict translational start sites from lncRNA transcript sequences (Clauwaert et al., 2023).

Database/tool	Use	Thesis section	Reference
Illumina body map	lncRNA expression levels in healthy tissues	3.2.1	Asmann et al., 2012; Barbosa-Morais et al., 2012; Derrien et al., 2012.
LncExpDB	Developmental dynamics of lncRNAs throughout development	3.2.2	Cardoso-Moreira et al., 2019; Sarropoulos et al., 2019
lncRNADisease	Disease association	3.2.3	Bao et al., 2019
Differential Expression Atlas	Disease association	3.2.3	Papatheodorou et al., 2020
Cancer RNA-Seq Nexus	Disease association	3.2.3	J.-R. Li et al., 2016
LncAtlas	Cytoplasmic vs Nuclear localisation of lncRNA	3.2.5	Mas-Ponte et al., 2017
Tis-Transformer	Translational start site prediction	3.4.1	Clauwaert et al., 2023
ColabFold	Protein modelling	3.14	Mirdita et al., 2022
Phyre2	Protein modelling	3.14	Kelley et al., 2015

Table 2.3: Databases used to investigate translated lncRNAs.

2.3 Cell culture

2.3.1 Cell maintenance and passage

Human cell lines SH-SY5Y (kindly provided by Dr Eric Hewitt) and HEK293 (kindly provided by Dr Iosifina Sampson, Bayliss group) were maintained in DMEM (Thermo) media supplemented with 10% Foetal Bovine Serum (Thermo) and 1% Penicillin/Streptomycin (P/S, Gibco) and incubated at 37°C, 5% CO_2. Cells were grown in T-75 flasks and passaged once confluency had reached ~70-80%. To passage growth media was removed, cells washed once with 1 XPBS and then covered in 1x trypsin (Merck) for 5 minutes at 37°C. An equal volume of full media was then added to neutralise the trypsin. Cells were harvested and pelleted at 400 xg for 5 minutes at RT. The supernatant was then discarded, and cells resuspended in DMEM and plated into a fresh flask. SH-SH5Y cells were split no more harshly than 1:5 and only passaged up to P15 before fresh stock was used. SH-SY5Y were induced to differentiate by supplementing DMEM with 30 μM of RA (dissolved in DMSO) over 72 hours. Control cells were treated with an equal volume of DMSO.

2.3.2 Cryogenic storage and recovery of mammalian cell-lines

Cells were grown to ~70%-80% confluency in DMEM media supplemented with 10% FBS and 1% P/S. Cells were then trypsinized and pelleted at 400 xg for 5 minutes at RT. Cell pellets were resuspended in freezing media (Table 2.4) to give an approximate concentration of $1x10^6$ cells/ml. Cryovials were filled with 1 ml of resuspended cells and placed at -80°C for >24 hours before moved to liquid nitrogen for long term storage.

Component	Final concentration	
Cell line	SH-SY5Y	HEK293
DMEM	-	70%
FBS	90%	20%
DMSO	10%	10%

Table 2.4: Freezing media composition.

2.4 Cloning of FLAG tagged expression vectors

2.4.1 Trizol RNA extraction

Cell pellets were re-suspended in Trizol (Invitrogen) and incubated for 5 minutes at RT before 100% chloroform was added (to a ratio of 5:1), shaken vigorously and then centrifuged at 12000 xg for 15 mins at 4°C. The aqueous layer was added to 1.5X volume of 100% ethanol and mixed. This solution as then added to an RNA spin column (Zymo) and washed following manufacturers protocol. RNA was eluted in 30 µl ddH20. RNA concentration was determined by Qubit RNA BR using 2 µl of purified RNA (Thermo Fisher, Q10210).

2.4.2 cDNA synthesis for cloning

cDNA was produced using the First Strand cDNA Synthesis kit (NEB #M0368). RNA (1 µg) was combined with dNTPs and RNase free water added to bring the reaction volume to 8µl. Then Random primer Mix was added (Table 2.5)

The reaction mix was heated to 65°C for 5 minutes before being placed on ice. Finally, 10µl of master mix containing all remaining components (Table 2.6) was added and tubes incubated at 25°C for 5 minutes, 42°C for 60 minutes and finally 65°C for 20 minutes.

Initial components	Final concentration
RNA	1µg
dNTP	1mM
Random primer mix	12mM
ddH2O	Up to 10µl

Table 2.5: initial components for first strand cDNA synthesis.

Secondary components	Volume per reaction
5X protoscript buffer	1x
DTT	10mM
ddH2O	Up to 10µl
RNase Inhibitor	0.4 U/µl
ProtoScript II RT	10 U/µl

Table 2.6: cDNA synthesis components.

2.4.3 Cloning primer design - pcDNA 3.1 Hygrow expression vector

Primers were designed to anneal from the start (AUG) of sORFs and to contain a 5' artificial Kozak (GCCACC) and NheI restriction site. Reverse primers targeted the 3' end of a sORF excluding the stop codon to allow for continual translation of the sORF and FLAG-tag within the same ORF. Additionally, 6 translated lncRNAs were also cloned to include their endogenous 5'-UTRs, (s)ORF with a FLAG-tag. pcDNA3.1 Hygrow expression vector (Addgene) was used for transient transfection of all FLAG-tagged sORFs for protein synthesis and localisation studies. pcDNA3.1 contains an Ampicillin resistance gene for selection. Primers were designed for 16 translated lncRNAs for (s)ORF only cloning (Table 2.7) and 6 translated lncRNAs were cloned with their endogenous 5'-UTRs (Table 2.8).

2.4.4 PCR to amplify translated sORFs

To isolate and amplify sORF sequences for cloning into expression vectors primers specific to (s)ORFs and containing the appropriate restriction sequences were combined with SH-SY5Y cDNA and sORF amplification was performed using Q5 High-Fidelity DNA Polymerase (NEB, M0491S, 2,000 units/ml) following manufactures instructions, with minor adjustments (Table 2.9, 2.10).

Gene	Direction	Sequence	Tm	Amplicon
AC008124.1	Forward	GCTAGCGCCACCATGCATCCGAGTTTGGAAACC	59.5	96
	Reverse	GGTACCAAAACGCGTAGAGGTGGTAGAA	60.1	
AC020928.1	Forward	GCTAGCGCCACCATGCGAGCCAAAGGATTCTTG	59.5	258
	Reverse	GGTACCCACGAAATAGTACCTATCGTTATTC	60.9	
AC244102.3	Forward	GCTAGCGCCACCATGGGGATCGGATGGGGG	60.8	276
	Reverse	GGTACCGAACCAGGTACTATGATCATTCAC	62	
AP001372.2	Forward	GCTAGCGCCACCATGAAAAATAGAAAAACACTTCATAAAGC	60.5	816
	Reverse	GGTACCTTTTATTTTTAGGTGAGAAGAAGCATC	60.8	
APTR	Forward	GCTAGCGCCACCATGGCCGAGGTAGCGATCG	61.6	177
	Reverse	GGTACCACAAAAAGAAGGACTTGATTCCCG	62	
EBLN3P	Forward	GCTAGCGCCACCATGGATACCTCTGAGCCTCTAT	60.1	129
	Reverse	GGTACCTGTCCTCCACTTCCCCAGT	59.5	
ENTPD1-AS1	Forward	GCTAGCGCCACCATGGTTCAGTTGACTATTGCAAAG	60.3	294
	Reverse	GGTACCGTTCCCAACAAGGTCTTCATCT	60.1	
GIHCG	Forward	GCTAGCGCCACCATGAAGTGGACACAGCATGTGAA	60.9	192
	Reverse	GGTACCAACAATGTCATCAGGACTTGGTC	60.9	
LINC00221	Forward	GCTAGCGCCACCATGGAGACTCTGAGAACATCCA	60.1	84
	Reverse	GGTACCTATGCAAACAGCCTCTCGTGG	61.2	
LINC00478	Forward	GCTAGCGCCACCATGGTTTGTGAAAGAATATCTGTGTG	61.7	111
	Reverse	GGTACCGTAGTAAAATGCTCTCTGATGAATTT	60.1	
LINC00839	Forward	GCTAGCGCCACCATGCTGGAGGCTGCTCCG	60.8	498
	Reverse	GGTACCACCCGGGTTTGAGCCCCT	60.8	
MAP4K3-DT	Forward	GCTAGCGCCACCATGCTAACGAGCTGTGAAATGAC	60.9	135
	Reverse	GGTACCGTATAGTCGGCTCTCCATCTG	61.2	
OLMALINC	Forward	GCTAGCGCCACCATGCATGTGACATTTGGTGCTG	60.1	225
	Reverse	GGTACCCCCGATTTGAAATTGGCAAGATG	60.9	
PSMA3-AS1	Forward	GCTAGCGCCACCATGTTTCTGGCGAGAAGGGAA	59.5	198
	Reverse	GGTACCGGTGATCCACCCGCCTC	59.8	
SNHG8	Forward	GCTAGCGCCACCATGGATGATGGAAACATAAGACTATC	61.7	108
	Reverse	GGTACCTTCGGAACACCCGTTTCCC	59.5	
EMSLR	Forward	GCTAGCGCCACCATGTCATATAGAGAATTGTGGAAACT	60.1	255
	Reverse	GGTACCGATTTTAAGTTTTGAGAGTTTTTTTTCTG	60.5	

Table 2.7: PCR primers for isolation of sORFs from SH-SY5Y cDNA. *Restrictions sites highlighted in red (NheI) and orange (KpnI). The artificial Kozak added to the 5' primer is highlighted in green.*

5'UTR-sORF primers			Tm	Expected amplicon
MAP4K3-DT	Forward	GCTAGCGTTCCGGCTCCGCGCG	60.9	351
	Reverse	GGTACCGTATAGTCGGCTCTCCATCTG	61.2	
LIPT2-AS1	Forward	GCTAGCGTACCGGGTGGCAAAGCAG	61.6	2010
	Reverse	GGTACCTTTTATTTTTAGGTGAGAAGAAGCATC	60.8	
GIHCG	Forward	GCTAGCATTCAGAGGTCTTTTAGGATGCG	60.09	715
	Reverse	GGTACCAACAATGTCATCAGGACTTGGTC	60.09	
LINC00839	Forward	GCTAGCTCACCCTGCTCCAGGCAG	60.8	592
	Reverse	GGTACCACCCGGGTTTGAGCCCCT	60.8	
APTR	Forward	GCTAGCTTCGGAGGTCAGAGCTCGG	61.6	319
	Reverse	GGTACCACAAAAAGAAGGACTTGATTCCCG	62	
OLMALINC	Forward	GCTAGCCTCAGTCTGCCCTACCCTG	61.6	623
	Reverse	GGTACCCCCGATTTGAAATTGGCAAGATG	60.9	

Table 2.8: Primers for isolating 5'UTR and (s)ORF. Restrictions sites highlighted in red (NheI) and orange (KpnI).

Component	Final concentration
Forward primer	0.5 µM
Reverse primer	0.5 µM
5x Q5 Reaction Buffer	1X
Betaine	1M
dNTPs	200µM
CDNA	< 1000ng
Q5 polymerase (2 U/µl)	0.02 U/µl
Total volume made up to 25µl with H20	

Table *2.9: Q5 polymerase PCR components*

Step	Temperature	Time	Cycles
Initial denature	98	30 seconds	
Denature	98	5 seconds	35
Anneal	59	20 seconds	
Extension	72	30 seconds	
Final extension	72	2 minutes	
Hold	4		

Table 2.10: Q5 running conditions

2.4.5 Agarose gels

Agarose gels (1%) were made up with 1X TAE and SYBR Safe (0.01% v/v; Thermo Fisher). Samples were mixed with 6X purple loading dye (NBE, B7024S). Up to 30μl of sample was loaded onto the gel in 1XTAE. Samples were run at 120V for 1 hour alongside 1kb plus DNA ladder (ThermoFisher, 10787018). Gel were either then imaged under UV using a Chemidoc XRS+ (BioRad) or processed to isolate bands from the gel.

2.4.6 Gel extractions

PCR fragments ran on agarose gel were excised from the gel. PCR fragments were then isolated using a QIAquick Gel Extraction Kit (Qiagen) following the manufacturer's instructions.

2.4.7 Blunt ended DNA ligation

Isolated PCR fragments of the correct size were incorporated into pCR blunt II TOPO (Invitrogen) by blunt-end ligation following manufacturers protocol. Plasmids were the transformed into chemically competent *E. coli* (DH5α).

2.4.8 Restriction enzyme DNA digest

DNA digests were performed with either KpnI-HF and NheI-HF for sORFs to be incorporated into pcDNA 3.1 (all enzymes from NEB). For diagnostic digest 500 ng of plasmid was digested with 10U of enzyme at 37^{o}C for 1 hour. For cloning 4 μg of (s)ORF containing plasmid was digested with 10U of enzyme for >3 hours at 37^{o}C. All digests were performed in 1XCutSmart® buffer (NEB). (s)ORF fragments and linearized plasmid were then separated on a 1% agarose gel in 1XTAE and purified using QIAquick Gel Extraction Kit (Qiagen) following the manufacturers protocol.

2.4.9 Sticky ended ligation

(s)ORFs with 5' end overhangs were ligated into plasmid vectors with complementary overhangs using T4 DNA ligase. Next, 1μl of (s)ORF DNA fragment and 1μl of linearised plasmid were added to 2.5μl of 2X rapid ligation buffer (Promega, C6711) before 0.5μl of T4 DNA ligase was added and contents gently mixed. The reaction was left at RT for 30 minutes before being transformed into competent bacteria.

2.4.10 Bacterial transformation

Chemically competent DH5α (50μl) were added to either plasmid DNA or 5μl of ligation reaction mix. DH5α was incubated for 30 minutes on ice before being heat shocking at 42^{O}C for 30 seconds in a pre-warmed water bath. Cells were immediately place back on ice for 2 minutes before 150 μl of either S.O.Cs or LB medium was added and cells were incubated at 37^{O}C in a shaking incubator (220 RPM) for 1 hour. Cells were then plated on warmed LB/agar plates containing either 25 μg/ml Kanamycin or 100 μg/ml Ampicillin, depending on plasmid. Plates were incubated overnight at 37^{O}C.

2.4.11 Small scale (mini) plasmid prep

LB medium containing 25 μg/ml Kanamycin or 100 μg/ml Ampicillin, depending on plasmid, were inoculated with single bacterial colonies and grown overnight (16 hours) at 37^{o}C in a shaking incubator (220 RPM). The resulting culture was centrifuged at 13,000 xg for 5 minutes. Plasmid isolation from DH5α pellets was then performed using QAIprep Spin mini-prep kit (Qiagen) according to manufactures instructions.

2.4.12 Medium scale (midi) plasmid prep

LB medium (5ml) containing the appropriate antibiotic was inoculated and incubated at 37^{O}C in a shaking incubator (220 RPM) for ~8 hours. The starter culture was used to inoculate 150ml LB medium which was then grown overnight at 37^{O}C in a shaking incubator. Plasmids were then isolated using the HiSpeed Plasmid Midi (Qiagen) kit following manufactures instructions.

2.5 Transient transfection for FLAG-tagged micropeptide expression

SH-SY5Y cells were plated in either 24-well plates containing 13 mm coverslips (VWR) or glass bottom 96 well plates (Cornering), coated with PLL (Sigma). Cells were seeded at 1x10^{5} in 24-well plates or 2x10^{4} in 96 well plates, respectively and incubated for 24 hours. Approximately, 30 minutes prior to transfection culture medium was replaced with DMEM +10% FBS but with no P/S. Plasmid DNA (200 ng/well for 24 well plate or 50 ng/well for 96 well plates) was diluted in Opti-MEM (ThermoFisher) with the addition of reagent P3000 (2 μl/1 μg of DNA). Lipofectamine 3000 was diluted in an equal volume of Opti-MEM as the DNA was. The DNA:Lipofectamine (ng:μl) ratio for SH-SH5Y transfection was 1:3.75 and 1:2 for HEK293 transfection. Equal volumes of diluted DNA and lipofectamine were then

combined and incubated RT for 25 minutes. DNA:lipofectamine was added in drops to cells. Cells were then incubated for 48 hours before fixation and immunostaining.

2.6 Imaging and image analysis

2.6.1 Immunofluorescence

Culture media was removed, and cells were washed once with 1XPBS containing 0.9 mM $CaCl_2$ and 0.5 mM $MgCl_2$. Cells were then fixed with 4% ice-cold paraformaldehyde for 20 minutes at room temperature. Wells were then washed twice with 1XPBS before blocking in 1XPBS containing 0.1% Triton-X and 5% normal goat serum (NGS) for 30 minutes at RT. Block was removed and replaced with primary antibody in 1XPBS containing 0.01% Triton-X and 0.5% NGS and incubated overnight at 4°C. Cells were then washed once in 1XPBS and then incubated with secondary antibodies for 2 hours at RT (Table 2.11). For 13 mm slides, secondary antibodies were removed and cells washed twice with 1X PBS before slides were removed, allowed to air dry and mounted onto slides using Vectashield antifade mounting medium with DAPI (Vector Laboratories). For 96 well plates, after secondary anti-bodies were removed cells were washed once with 1XPBS. Nuclear staining was achieved with 1X PBS containing 0.1 μg/ml Hoechst 33342 for 5 minutes at RT. Cells were finally washed 3X with 1XPBS, with the last wash being left on to prevent cells drying out.

Type	Name	Species raised	Usage (IF/WB)	Company	Dilution IF	Dilution WB
Primary	Anti-FLAG M2	Mouse	Both	Sigma	1:1000/	1/10,000
	Anti-FLAG	Rabbit	Both	Sigma	1:1000/	1/10,000
	Anti-Beta-tubulin	Rabbit	Both	Proteintech	1:400	
Secondary	Alexa 488 anti-mouse	Goat	IF	Thermofisher	1:500	
	Alexa 594 Anti-mouse	Goat	IF	Thermofisher	1:500	
	Alexa 633 anti-rabbit	Goat	IF	Thermofisher	1:500	
	LI-COR 680nm anti-mouse	Goat	WB	LI-COR		1/10,000
	LI-COR 800nm anti-rabbit	Goat	WB	LI-COR		1/10,000

Table 2.11: Antibodies and dilutions used.

2.6.2 Widefield and confocal microscopy

FLAG-tagged (s)ORF transfections were first screened using an EVOS fluorescent microscope (M7000, Invitrogen) before detailed imaging with a confocal microscope (LSM880, ZEISS). Representative images were taken with a 40X oil objective (Pln Apo 40X/1.3 Oil DICIII) and processed using Fiji software (Schindelin et al., 2012).

2.6.3 Co-localisation analysis

SH-SY5Y cells expressing GIHCG-FLAG were stained with Mitotracker CMXRos (ThermoFisher, M7512). Culture media was removed from cells and replaced with DMEM (no FBS or P/S) supplemented with 200nM Mitotracker for 30 minutes at 37 ^{o}C. SH-SY5Y cells were then fixed and stained for FLAG to identify transfected cells (Methods 2.5). Cells were imaged using a confocal microscope (LSM880, ZEISS). Representative images were taken with a 40X oil objective and analysed in FIJI using the Just Another Colocalization Plugin (BOLTE & CORDELIÈRES, 2006).

2.6.4 Mitochondrial structure analysis

SH-SY5Y overexpressing GIHCG-FLAG or FLAG alone for 72 hours (to allow for differentiation to be induced) were stained with Mitotracker before being fixed and immunofluorescence used to identify cells overexpressing GIHGC-FLAG. Cells were imaged using a confocal

microscope (LSM880, ZEISS). Representative images were taken with a 40X oil objective and analysed in FIJI using the plug in MiNA - Mitochondrial Network Analysis using a previously described pipeline (Valente et al., 2017). In brief images were converted to binary and skeletonised before analysis with the MiNA plugin.

2.7 Quantification of RNA levels

2.7.1 RNA extractions for RT-qPCR and RNA-Seq

RNA was purified using ZYMO Quick-RNA™ Miniprep according to the supplier's instructions. RNA was eluted in 30 μl of RNase free ddH2O. TURBO-DNase treatment was performed for 30 mins at 37°C following manufactures instructions (6U of TURBO DNase I per 10 μg of RNA). RNA was then re-purified with RNA Clean & Concentrator-25 (Zymo, R1017) following manufacturer's instructions. RNA concentration was determined by Qubit RNA BR using 2 μl of purified RNA.

2.7.2 cDNA synthesis for RT-qPCR

cDNA generation from 1 μg of RNA was performed with the Q-script cDNA synthesis kit. Purified RNA was combined with 5X qScript Reaction Mix and 1 μl of qScript RT to a total volume of 20 μl. Samples were then placed in a thermocycler and run at 22°C for 5 minutes, 42°C for 30 minutes and finally 85°C for 5 minutes.

2.7.3 Quantitative Real Time PCR (RT-qPCR)

RT-qPCR was performed using CFX Connect thermal cycler and SYBR Green fluorescent dye (PowerUp SYBRTM Green Master Mix, Thermo) according to the manufacturer's instructions. Primers were designed to anneal at 60°C. Negative controls of no template (NTC) and no RT (NRT) were included for each primer pair used. cDNA was diluted 1/10 in RNase free ddH2O with 5 μl (25 ng) cDNA used per well. qPCR was performed in 96-well plates (Appleton Woods-BP049) and plated in triplicate. Cycle details and well components summarised below (Table 2.12/ 2.13). Relative quantification using the ΔCq method was performed with HAUS8 and NDUFV2 used as the reference. Primers used are listed in table 2.14.

Component	Final concentration
PowerUp SYBR Green master mix	1X
Forward Primer	300 nM
Reverse Primer	300 nM
cDNA	5 ng/µL

Table 2.12: RT-qPCR components.

Step	Temperature (^{O}C)	Duration	Cycles
UDG activation	50	2 minutes	1
Activation (Dual-Lock™ DNA polymerase)	95	2 minutes	1
Denature	95	15 seconds	40
Anneal/extend	60	1 minutes	
Melt curve	65-95; increment of 0.5^{o}C	5 seconds	

Table 2.13: RT-qPCR cycle parameters.

Gene	Direction	Sequence	Exon spanning
APTR	Forward	GGAAGTGTTCAAAATCATTCGATGTGG	Yes
	Reverse	GGCTGAACTTCTGTGAATCCATGGG	
OLMALINC	Forward	CAGCAAAACACACTAAAGATTGGGCAG	Yes
	Reverse	CCGTGTGAAGAAACCACCAAACAG	
LINC00839	Forward	TGCCCATACGGACCTACTGC	No
	Reverse	TCCTGGAACAATCTGGGCCTG	
GIHCG	Forward	GTGAAGACAATTCTTTCAAGAAGTTTGGC	Yes
	Reverse	CAGACAAACCTCCATGTGGTTGC	
LIPT2-AS1	Forward	GCGCTTGATACAGCGGTGC	Yes
	Reverse	GCATGAACTGGGTGTTGTGCA	
MAP4K3-DT	Forward	CCTCGGCTTTTCTGCGGAC	Yes
	Reverse	CCCTTCTGCTGCCAGATCTTTG	
NTN4	Forward	CGAGTGCAGAACCTGCAAGTGT	Yes
	Reverse	CATCTGGAGCTGAGAAGGGTC	
RET	Forward	CCAGCATCTCTACGGCACGTA	Yes
	Reverse	CCGCGGTTGCGGACACTGA	
SOX2	Forward	ACATGAACGGCTGGAGCAA	No
	Reverse	GTAGGACATGCTGTAGGTGGG	
LIPT2 (Sense)	Forward	CAAGATCTGCGCGATCGGAGT	Yes
	Reverse	CTCAAACCACGTGAGGTCGG	

Table 2.14: qPCR primers.

2.7.4 RNA-Seq sample preparation and sequencing

RNA was isolated from LIPT2-AS1$^{-/-}$ (CRISPR knockout cells) and WT SH-SY5Y in both immature and differentiated conditions using Zymo Quick-RNA™ MiniPrep kit. RNA was quantified by Qubit and 200 ng of RNA was run on a non-denaturing agarose gel to assess quality through the comparison of 28S and 18S rRNA bands. Purified RNA was sent to Novogene for Poly-A selection and paired-end sequencing of 150 nucleotides.

2.7.5 RNA-Seq analysis (performed by Eilidh Ward)

Raw data was kindly processed and analysed by Eilidh Ward (Aspden group, University of Leeds). In brief data was first put through FastQC (v0.12.1) and MultiQC to generate a quality report and identify adapters followed by Cutadapt (v4.6) in paired end mode to remove adapter sequences and poly-A reads and FastP (v) was used to remove poly-G reads. Poor quality reads and r/tRNA contaminants were removed using fastx and Bowtie2, respectively. Once data had been cleaned reads were aligned to the human genome (GENCODE V45) using STAR and gene counts quantified with featureCounts in paired end mode. Finally, DESeq2 was used to identify differentially expressed genes.

2.7.6 GO term analysis

Using the differentially expressed genes identified by Eilidh Ward I used g:Profiler to identify gene ontology terms (Reimand et al., 2007).

2.8 CRISPR CAS-9

2.8.1 sgRNA design

Custom sgRNAs were designed and inserted into a plasmid encoding the CRISPR-Cas9 machinery and EGFP (pSpCas9(BB)-2A-GFP (PX458, addgene #48138). This plasmid encodes wildtype Cas9 from S. pyogenes, which recognizes the PAM site NGG. Multiple sgRNA composed of the 20 bp 5' upstream from PAM site were selected and filtered for efficiency using the CRISPR Efficiency Predictor (Housden et al., 2015), retaining sgRNAs with an efficiency score of > 7.5 as suggested by creators of the CRISPR Efficiency Predictor. Additionally, sgRNA specificity was checked with CRISPR-Cas9 guide RNA design checker (IDT, https://eu.idtdna.com/site/order/designtool/index/CRISPR_SEQUENCE). Where possible two sgRNAs per ORF were then cloned into pSpCas9(BB)-2A-GFP, one targeting the ORF ATG to mutate the start codon and prevent initiation and the other downstream of the ATG to create a frameshift or truncation mutation. sgRNAs were used individually (not pooled) to maximize chance of small deletion and precise targeting. sgRNAs were ordered with containing 5' overhangs complementary to BbsI to allow ligate into pSpCas9(BB)-2A-GFP plasmid (Table 2.15).

2.8.2 sgRNA incorporation into pSpCas9(BB)-2A-GFP plasmid

sgRNAs were ordered as ssDNA oligos. sgRNA oligos were first annealed by mixing 25 μl of each oligo (200 nM) with 50μl of annealing buffer (100mM NaCl, 2 mM EDTA, 20 mM Tris-HCl pH8). Mixture was then heated to 95^{o}C for 10 minutes before being slowly cooled to RT for > 2 hours. Annealed oligos were then ligated into pSpCas9(BB)-2A-GFP (PX458) cut with BbsI with T4 DNA ligase (see Methods 2.4.8/9).

Gene name	Number	Strand	sgRNA sequence (+ sticky ends)
LINC00839	1	Sense	CACCGTGGTCTGTATGGCTCCGCCC
		Antisense	AAACGGGCGGAGCCATACAGACCAC
	2	Sense	CACCGTCCAGGCAGCGTGGTCTGTA
		Antisense	AAACTACAGACCACGCTGCCTGGAC
GIHCG	1	Sense	CACCGATTGAAAAGACAGGATGAAG
		Antisense	AAACCTTCATCCTGTCTTTTCAATC
	2	Sense	CACCGGACAATTCTTTCAAGAAGTT
		Antisense	AAACAACTTCTTGAAAGAATTGTCC
LIPT2-AS1	1	Sense	CACCGTTGATCGTGTATTGAAAGAG
		Antisense	AAACCTCTTTCAATACACGATCAAC
	2	Sense	CACCGTGTGTTCACTGCGACGCTGA
		Antisense	AAACTCAGCGTCGCAGTGAACACAC
OLMALINC	1	Sense	CACCGACACGGATGCATGTGACATT
		Antisense	AAACAATGTCACATGCATCCGTGTC
	2	Sense	CACCGGTTGTCGCCCTCACTCCGTG
		Antisense	AAACCACGGAGTGAGGGCGACAACC

Table 2.15: sgRNA sequence including complementary sticky ends to BbsI cut sites.

2.8.3 CRISPR/Cas9-mediated knockouts

SH-SY5Y cells were transiently transfected using lipofectamine 3000 (ThermoFisher, L3000015) with plasmid encoding sgRNAs, CRISPR-Cas9 and EGFP (addgene #48138). Cells were left for 24 hours post transfection cells then harvested and GFP positive cells were individually sorted into wells 96 well plates containing 180 µl of conditioned media using a FACS melody (BD Biosecience). FACS was performed by Dr Ruth Hughes of the University of Leeds Bioimaging facility. Cells were incubated until clonal colonies had formed then expanded into 2 wells of a 6 well plate. One well was then frozen for recovery (see Methods 2.3.2) and the other for gDNA extraction and genotyping.

2.8.4 Conditioned media for CRISPR cell recovery

Recovery media (Table 2.16) was added to SH-SY5Y cells at ~70% confluency and left overnight. The next day the media was collected off SH-SY5Ys and mixed with equal parts of fresh recovery media before being sterile filtered and is now referred to as conditioned media. Conditioned media was then used immediately (Table 2.16).

Component	Concentration
DMEM	69%
FBS	30%
P/S	1%
HEPES (7.5pH)	25 mM

Table 2.16: Recovery media composition.

2.8.5 gDNA extraction from CRISPR clones

Genomic DNA was extracted from CRISPR clone pellets using DirectPCR Lysis Reagent Cell (Viagen). In brief, 100 µl of DirectPCR Lysis reagent containing 5 µg of Proteinase K was used to resuspend each pellet. The lysate was incubated overnight at 55°C followed by a 3-hour incubation at 85°C to inactivate Proteinase K. DNA was diluted 1:10 in ddH2O, ready to be used in PCR.

2.8.6 PCR amplification of gDNA targeted by sgRNA

The region of gDNA targeted by the sgRNA was amplified using flanking primers (Table 2.17) and a non-proofreading Taq polymerase (ThermoFisher, EP0402) (Table 2.18-2.19). PCR products were then purified by gel extraction and incorporated into TOPO plasmid through TA cloning.

Gene	Direction	Sequence	Amplicon size (bp)	Tm
LINC00839	Forward	TACACACTGAATATGCTGCTCCCAG	905	62
	Reverse	AGAGTCAGACGATCTGCCAGG		61
GIHCG	Forward	GGGATGTGTTCTAATGCGCT	611	59
	Reverse	CCCAATTTGTGGCACCCCTC		62
OLMALINC	Forward	TCATTGCTATGGTTTCAAATTCCCA	894	59
	Reverse	CCGCAATTGACTTGTCACCA		60
LIPT2-AS1	Forward	GCACAACACCCAGTTCATGC	1427	60
	Reverse	GCCAGCACAGTAGTGTCACC		61

Table 2.17: Genotyping primers.

Component	Final concentration
NH4 buffer (x10)	1X
dNTP	200 µM
Forward primer	0.5 µM
Reverse primer	0.5 µM
MgCl2	6 µM
Taq pol EP0402 (5U/µl)	0.625 U
Template	0.4 ng/µl
ddH_20	Up to 25 µl total volume

Table 2.18: Taq PCR components

Step	Temperature (°C)	Time	Cycles
Initial denature	95	3 minutes	
Denature	95	30 seconds	35
Anneal	55	30 seconds	
Extension	72	1 minute	
Final extension	72	30 minutes	
Hold	4		

Table 2.19: Taq PCR cycling condition

2.8.7 TA cloning

TA cloning kit (ThermoFisher, K200001) was used following manufacturer's instructions. In brief, PCR fragments from gDNA template generated by a non-proof reading Taq (leaves overhang of single Adenosine) were purified and ligated into TOPO plasmid with complementary T overhangs. Components and volumes used outlined in Table 2.20. Ligation mixes were then transformed into DH5α for screening.

Component	Volume (µl) per 10 µl reaction
ddH_2O	4
5X reaction buffer	2
Linearised vector	2
Insert	1
DNA ligase	1

Table 2.20: Reaction components for TA ligation

2.8.8 Colony PCR

From each plate 8 single colonies were picked and mixed into PCRs containing DreamTaq (Thermo scientific, K1081) and primers complement to specific region of gDNA sgRNAs target. DreamTaq was used following manufacturer's instructions. Half of PCR reaction mix was run on a 1% agarose gel (see Methods 2.4.5) to ensure PCR was successful and the remainder cleaned using ExoSap-IT (ThermoFisher, 75001.1. EA) and sent for sequencing to determine if CRISPR has been successful. Sequencing used the same primers as were used for gDNA amplification.

2.9 siRNA knockdowns

siRNAs targeting lncRNAs of interest were transfected into SH-SY5Y using RNAiMAX following manufacturer's instructions. In brief, siRNA and RNAiMAX were diluted in equal volumes of Opti-MeM (Thermo) then combined. The siRNA and RNAiMAX mix were incubated for 25 minutes at RT. After incubation mixture was added to an empty well of a 6 well plate. SH-SY5Y cells were then added directly to the well at 400,000 cells/well in DMEM containing 10% FBS but no P/S.

2.10 Neurite extension assay

SH-SY5Y cells were plated in 24 wells plates on 13 mm coverslips coated with Poly-L-lysine at $4x10^4$ cells/well and incubated for 24 hours before siRNA knockdown. Knockdown media was changed for full DMEM with RA or DMSO 24 hours after siRNA treatment. SH-SY5Ys were harvested and fixed after 72 hours before being stained with mouse-anti TUJ1, (Proteintech) and visualised with Alexa 488 (Goat anti-mouse, ThermoFisher) (see Methods 2.6.1). Coverslips were mounted onto glass slides with Vectashield containing DAPI (Vector Laboratories). Images were taken at 10X magnification (EVOS-M7000, Invitrogen). >100 neurites were measured from 5 representative images for each well, with each condition plated in duplicate. For each condition, 3 biological replicates were performed. Images and neurite measuring occurred under blinded conditions where the condition and time point of each slide was hidden.

2.11 Cell growth curves

Cells were plated in 24 well plates at 25000 cells/well in DMEM + 10% FBS + 1% P/S. Cells were allowed to recover before being treated with either RA or DMSO. Cells were harvested and mixed with an equal volume of Trypan Blue solution (0.4% v/v) before counting on a haemocytometer. Counts were performed one day after plating and then at 48-hour intervals over 4 time points.

2.12 ddPCR of cortical organoid cDNA

Cortical organoids were kindly provided by the Poulter group (University of Leeds, School of Medicine) and were grown by Erica Harris (University of Leeds, Poulter group). Starting with iPSCs self-organizing neural/cortical organoids can be developed through multiple stages of different culture conditions. Cortical organoid growth accurately models *in vivo* corticogenesis and therefore provide a powerful model to investigate normal and aberrant brain development (Eigenhuis et al., 2023; Lancaster et al., 2013)

2.12.1 RNA extraction from cortical organoid

RNA extraction from Cortical organoids was performed using RNeasy Micro Kit (Qiagen) following manufacturer's instructions and RNA was quantified by Qubit BR RNA. To measure lncRNA levels from neural induction to neuronal differentiation 6 organoids were pooled per biological repeat and time points from day 3 to day 26 at ~3-day intervals. cDNA was generated using Q-Script (see Methods 2.7.2)

2.12.2 Polysome profiling of cortical organoids

Polysome profiling was done with the help of Dr Karl Norris. Non-linear Sucrose gradients were generated containing a range of sucrose concentrations from 18% to 60% (w/v). Cycloheximide, Dithiothreitol (DTT) and protease inhibitor (Roche) were added to pre-made sucrose solutions just before pouring (Table 2.21). Gradients were made in 12.5 mL POLYCLEAR centrifuge tubes (Seton Scientific). Each layer of sucrose was snap-frozen in liquid nitrogen before progressing to the next level

Per biological replicate, ~80 individual day-25 organoids were collected and combined. Cycloheximide (CHX) was added to media (100 μg/ml) and organoids are incubated at 37 ºC for 5 minutes. Media was removed and organoids washed with ice cold 1XPBS + CHX (100 μg/ml) then pelleted at 400 xg for 5 minutes. Pelleted organoids are resuspended in lysis buffer (Table 2.22) and incubated on ice for >30 minutes. Lysed cells are then pelleted at 17,000 xg for 5 mins and supernatant loaded onto sucrose gradient. Gradients were then spun at 121,300xg (r_{AV}) for 3.5 hours at 4^{o}C in a SW40 rotor.

Sucrose gradients were fractionated using a GRADIENT STATION instrument (Biocomp) and the RNA trace generated by measuring absorbance at 254 nm using a Bio-Rad UV-vis spectrophotometer. Samples containing 40S, 60S, 80S and polysomes were collected for RNA extractions.

2.12.3 RNA extraction from sucrose gradient fractions

Fractions were collected, and the absorbance at 254 nm used to group fractions into 40S, 60S, 80S and polysome. To each fraction, 1.5x volume of isopropanol was added, followed by NaCl to a final concentration of 380 mM and 1µl Glycoblue. Mixtures were then incubated at -80 °C for >24 hours.

Samples were spun at 17,000 xg, 4 °C for 20 minutes and supernatant discarded. Two sequential washes with ice cold 70% ethanol were performed. RNA was resuspended in ddH2O, and DNase treatment was performed (Turbo DNase-see Methods 2.7.1) and cleaned and purified (Zymo RNA clean & concentrate) following manufacturers instruction and purified RNA eluted in 30µl of ddH2O.

Component	Concentration
Sucrose	18%-60% (w/v)
Tris-HCl pH8	50 mM
NaCl	150 mM
$MgCl_2$	10 mM
DTT	1 mM
Cycloheximide	100 µg/ml
Protease inhibitor	0.33%

Table 2.21: Sucrose gradient composition

Component	Concentration
Tris-HCl pH8	50 mM
NaCl	150 mM
$MgCl_2$	10 mM
IGEPAL	1%
DTT	1 mM
Cycloheximide	100 µg/mL
Turbo DNase	24 U/mL
RNasin plus	90 U
Protease inhibitor, Roche EDTA free	1X
ddH2O	

Table 2.22: Polysome profiling lysis buffer composition.

2.12.4 ddPCR of lncRNA from cortical organoids

For ddPCR of total RNA from organoids 20 ng of cDNA per well was used. cDNA was generated from purified RNA using the Q-script cDNA synthesis kit (Methods 2.7.2). Due to low RNA yield from the sucrose gradients 2.5 ng of cDNA was used per well when looking at RNA levels across 40S, 60S, 80S, and polysomal fractions.

cDNA was mixed with QX200™ ddPCR™ EvaGreen Supermix (Bio-Rad #1864034), forward and reverse primer and ddH20 up to 22 µl per well (Table 2.23). Once plated, automated droplet generation was performed using QX200 AutoDG Droplet Digital PCR System (Bio-Rad) and PCR performed (cycle details – Table 2.24). Droplets were read using QX200 Droplet Reader (Bio-Rad #1864003).

Component	Final concentration
EvaGreen Supermix	1X
Forward primer	100 nM
Reverse Primer	100 nM
cDNA	Variable (2.5ng-20ng)
ddH2O	Up to 22 µl

Table 2.23: Component per well for ddPCR.

Step	Temperature (°C)	Duration	Cycles
Enzyme activation	95	2 minutes	1
Denature	95	2 minutes	40
Anneal/extend	60	1 minute	40
Signal stabilisation	4	5 minutes	1
	90	5 minutes	1
Hold	4	Infinite	1

Table 2.24: Cycle conditions for ddPCR

2.13 FLAG-tagged protein pulldowns

SH-SY5Y cells (6 x10^6) were plated in 10 cm dishes and transfected (see Methods 2.5) with 14 μg of plasmid DNA encoding FLAG-tagged protein of interest. Transfected cells were incubated for 48 hours at 37°C before harvesting with a cell scraper. Cells were then pelleted for 5 mins at 400 xg at 4 °C. Cells were washed twice in ice cold 1X PBS. Cleaned cell pellets were resuspended in 400 μl lysis buffer (Table 2.23) and incubated on ice for 20 minutes. Lysate was cleared by centrifugation at 16,000 xg for 10 minutes (4°C) and supernatant added to low-bind Eppendorf's containing 30 μl of magnetic beads coated with M2 anti-FLAG antibody (Millipore, M8823) that had undergone 3 washes with lysis buffer (Table 2.25).

Lysate and beads were incubated overnight at 4°C with gentle rotation (30 RPM). Beads were isolated using a magnet and washed 3 times in ice cold wash buffer (Table 2.26). To elute proteins, beads were boiled for 20 minuets at 50°C in Lamelli buffer (Biorad). Success of pulldown was then assessed through western blot to detected FLAG tagged protein of interest.

Component	Concentration
KCL	150 mM
Tris-HCl pH 7.5	25 mM
EDTA pH 8.0	5 mM
NP-40	0.5 %
DTT	0.5 mM

Table 2.25: Components of pull-down lysis buffer.

Component	Concentration
Tris-HCl pH 7.5	10 mM
NaCl	150 mM
EDTA	0.5 mM

Table 2.26: Components of pull-down wash buffer.

2.13 SDS-PAGE

Transfected cells were harvested and pelleted. Cells were resuspended in 4x Laemmli buffer and boiled at 95°C for 5 minutes. Samples were run on a 4%-20% polyacrylamide gradient gel (Bio-Rad) for 1 hour at 120V in running buffer (Table 2.27) with an ice block.

2.13 Western blot of FLAG-tagged peptides

After separation on an SDS-PAGE, proteins were transferred onto 0.2 μm nitrocellulose membrane (GE Healthcare) in cold transfer buffer (Table 2.28) using BIO-RAD transfer apparatus. Transfer was performed for 70 minutes at 200 mA. Success of transfer was checked using Ponceau Red (VWR). Membranes were washed in dH2O until excess ponceau was removed and then blocked (5 w/v% fat-free milk powder, 1X TBS) for 1h at RT on a roller. Following blocking membranes were incubated overnight with primary antibodies in TBST + 5 w/v% fat-free milk at 4°C. Membranes were then washed 3 times in TBST (5-minute washes) before an incubation with secondary antibodies conjugated to LI-COR probes (Table 2.11) in TBST + 5 w/v% fat-free milk for 2 hours at RT. Following incubation membranes were washed 3 times in TBST (5-minute washes) before being visualised (LI-COR).

Reagent	1x Final conc.
Tris-base	25 mM
Glycine	192 mM
SDS	0.1%

Table 2.27: SDS-PAGE running buffer.

Reagent	1x Final conc.
Tris-base	25 mM
Glycine	192 mM
Methanol	10-20%

Table 2.28: Western blot transfer buffer.

2.12 Silver stain

To visualise total protein from pulldown experiments 10% of elute was loaded onto an SDS-Page 4%-20% gradient gel (Bio-Rad) and run for 1 hour at 120V. Gels were then fixed and stained using a silver stain kit (ProteoSilver™ Silver Stain Kit, Sigma Aldrich) following manufacturer's instructions.

2.13 Label free Mass Spectrometry

Label free quantitative Mass Spectrometry (MS) and subsequent analysis was performed by Dr Sri Ranjani Ganji (University of Leeds Mass Spectrometry team). MS was performed for pulldowns of LIPT2-AS1-FLAG and FLAG alone as a negative control. Samples were provided as elute from magnetic beads coated with M2 anti-FLAG antibody in 2X Lamelli buffer. MS data was processed using MaxQuant (Cox & Mann, 2008)

I then calculated Log2Fold enrichment and proteins enriched ≥1.5 log2fold in LIPT2-AS1 pulldown compared to FLAG-alone pulldown were analysed by string analysis and GO term enrichment analysis. String analysis was performed in browser using the String website using baseline parameters (Szklarczyk et al., 2023).

2.14 Statistical tests

Statistical significance between independent samples with equal variances was determined with a two-tailed paired student's T test unless specified. A $p<0.05$ was accepted as significant.

Chapter 3: Assessing the potential function of translated lncRNAs

3.1 Introduction

Neuronal differentiation is a highly complex and tightly regulated process involving dramatic changes in gene expression of both mRNAs and non-coding RNAs (Albert & Huttner, 2018; Burke et al., 2020). Many lncRNAs have been identified as key factors in neuronal progenitor cells decision to self-renew or differentiate into mature neurons, as well as contributing to the progression of neurodevelopmental disorders and glioma (Y. Zhao et al., 2020)

Non-canonical proteins derived from lncRNAs have been identified as crucial components in human physiology and disease with microproteins as small as 11 amino acids functioning to modulate insect morphogenesis (Kondo et al., 2007). In humans, the smallest annotated ORF to date is MOTS-c, which is translated to produce a 16 amino acid peptide. MOTS-c is encoded from the mitochondrial genome and acts as a signalling peptide, regulating skeletal muscles response to insulin (C. Lee et al., 2015). However, sORFs encoding peptides of between 3 and 15 amino acids have been described, but not annotated, and show clear evidence of translation in Ribo-Seq data (Sandmann et al., 2023), suggesting current cut-offs of 16 amino acids for micropeptide identification used by the micropeptide consortium is too stringent (Mudge et al., 2022).

Previous work by the Aspden group identified the translation of 45 (s)ORFs from 35 annotated lncRNA using Poly-Ribo-Seq of SH-SY5Y cells. Poly-Ribo-Seq only identifies ORFs that are being actively translated by multiple ribosomes and therefore avoids false positives generated by Ribo-Seq reads from sporadic binding of single ribosomes or ribosomal subunits (Aspden et al., 2014). Additionally, despite the limitations of detecting micropeptides by mass spectrometry (MS) (Tharakan & Sawa, 2021)the presence of several candidate proteins was validated by MS of SH-SY5Y cells (Douka et al., 2021).

However, the presence of a novel protein in a cell, even if detected by MS, does not necessarily ensure that protein is functional. A common measure to infer function is to assess the level of evolutionary conservation, working on the assumption that a genes function has arisen throughout evolution (Prensner et al., 2023). However, the majority (~ 90%) of lncRNA derived proteins display minimal conservation and are termed ‘evolutionarily young’ either being human specific or conservation limited to great apes (Sandmann et al., 2023). Importantly, their recent de novo birth does not preclude a novel

peptide from having a function (Papadopoulos & Albà, 2023). Current literature contains multiple examples of functional translated lncRNAs that are evolutionarily young including the hominoid specific gene ENSG00000205704, which was previously annotated as a lncRNA named LINC00634. ENSG00000205704 produces a 107 aa protein which is involved in neuronal development and its expression required for cortical organoid development from hESCs. Knockout of ENSG00000205704 resulted in significantly smaller organoids than wild-type controls due to accelerated neuronal maturation (An et al., 2023).

This chapter aims to identify which of the 45 translated lncRNA ORFs identified in SH-SY5Ys are most likely to produce functional novel peptides. Firstly, to identify candidates lncRNAs potentially involved in neuronal development and disease so we can characterise candidate most likely to be functional in humans. The most promising candidates were then cloned to allow for confirmation of their ability to produce stable peptides and initial characterisation of these peptides. Further characterisation was performed using structural prediction tools to elucidate potential functions to be investigated.

3.2 Analysis of publicly available data on translated lncRNAs

3.2.1 Translated lncRNAs are expressed in multiple human tissues

To determine the expression pattern of translated lncRNAs across human tissues, publicly available RNA-Seq data was accessed through the Illumina Body Map (Asmann et al., 2012; Barbosa-Morais et al., 2012; Derrien et al., 2012). Only 23 out of our 35 translated lncRNA genes were present in the Body Map dataset and their expression in the brain varied from 0 (not detected) to 340 transcripts per million reads (TPM) (Figure 3.1). lncRNAs such as SHNG8 and FGD5-AS1 are highly expressed in all tissues sequenced, whilst LINC01515 is only detected in 6 of the 15 tissues highlighting the diversity in tissues expression for translated lncRNAs. Out of the 23 translated lncRNA gene in the database, only OLMALINC, GIHCG and AC020928.1 showed their highest RNA abundance in the brain (Figure 3.1).

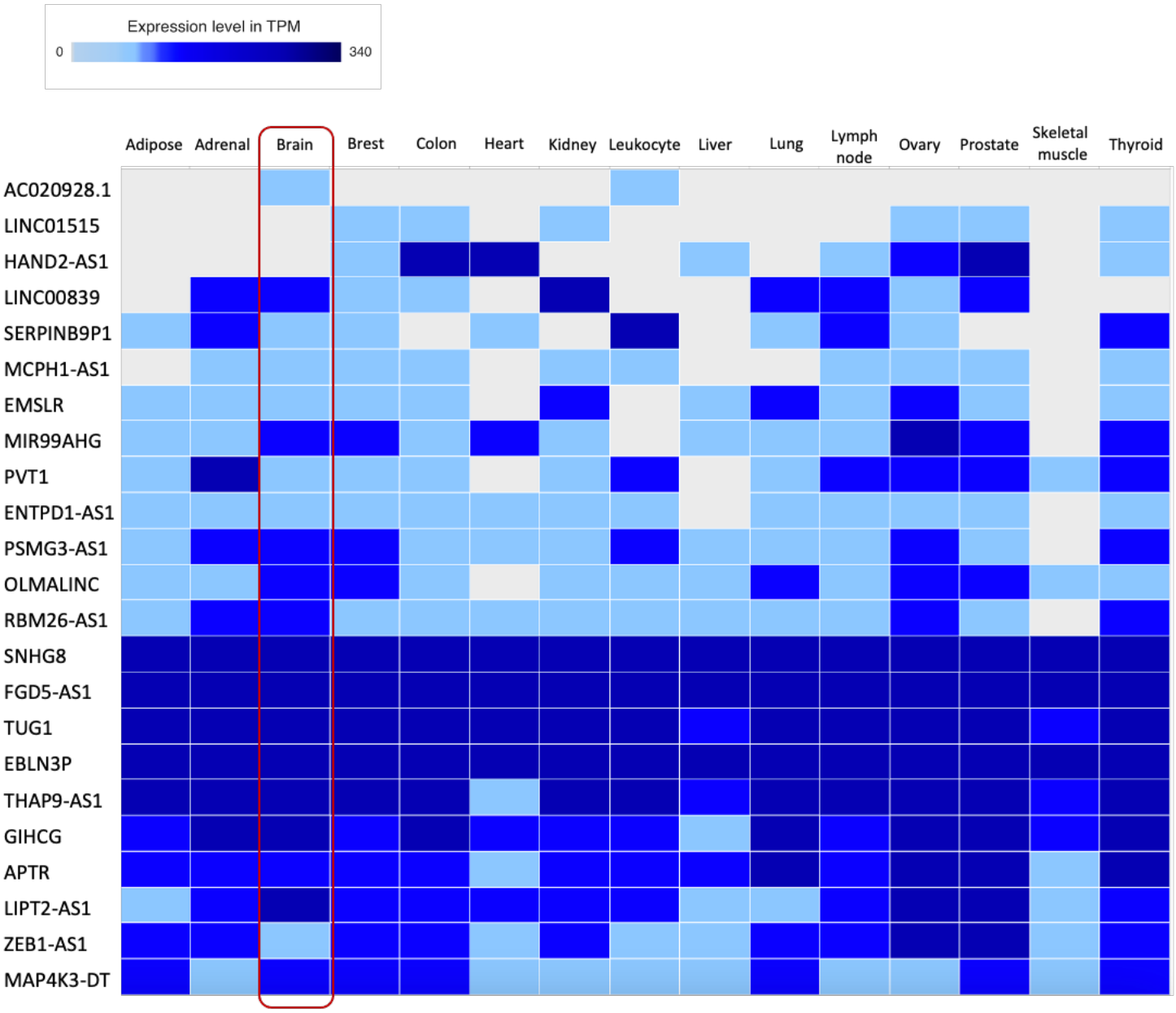


Figure 3.1: RNA abundance level of translated lncRNA in human tissues. *TPM of translated lncRNAs across 15 human tissues. Data collected from Illumina Body Map* (Asmann et al., 2012; Barbosa-Morais et al., 2012; Derrien et al., 2012)*. Brain expression is highlighted in red box.*

3.2.2 Translated lncRNAs genes show developmentally dynamic expression

LncExpDB contains human lncRNA expression data from various tissues throughout human development (Cardoso-Moreira et al., 2019; Sarropoulos et al., 2019). Searching for the 35 genes detected by Poly-Ribo-Seq in this dataset revealed that translated lncRNAs are more likely to show developmentally dynamic expression – defined as genes that show large changes in expression level within an organ during human development - than lncRNAs overall. Out of the 35 translated lncRNA genes 14 (39%) showed developmentally dynamic expression in humans, compared to just 19% of all lncRNAs (Figure 3.2A). Additionally, over half of the translated lncRNAs that showed developmentally dynamic expression did so in the brain (57%). Whereas for all developmentally dynamic lncRNAs in LncExpDB, only 17% were dynamic in the brain, with the majority of dynamic expression for all lncRNAs being found in the testis (Figure 3.2B).

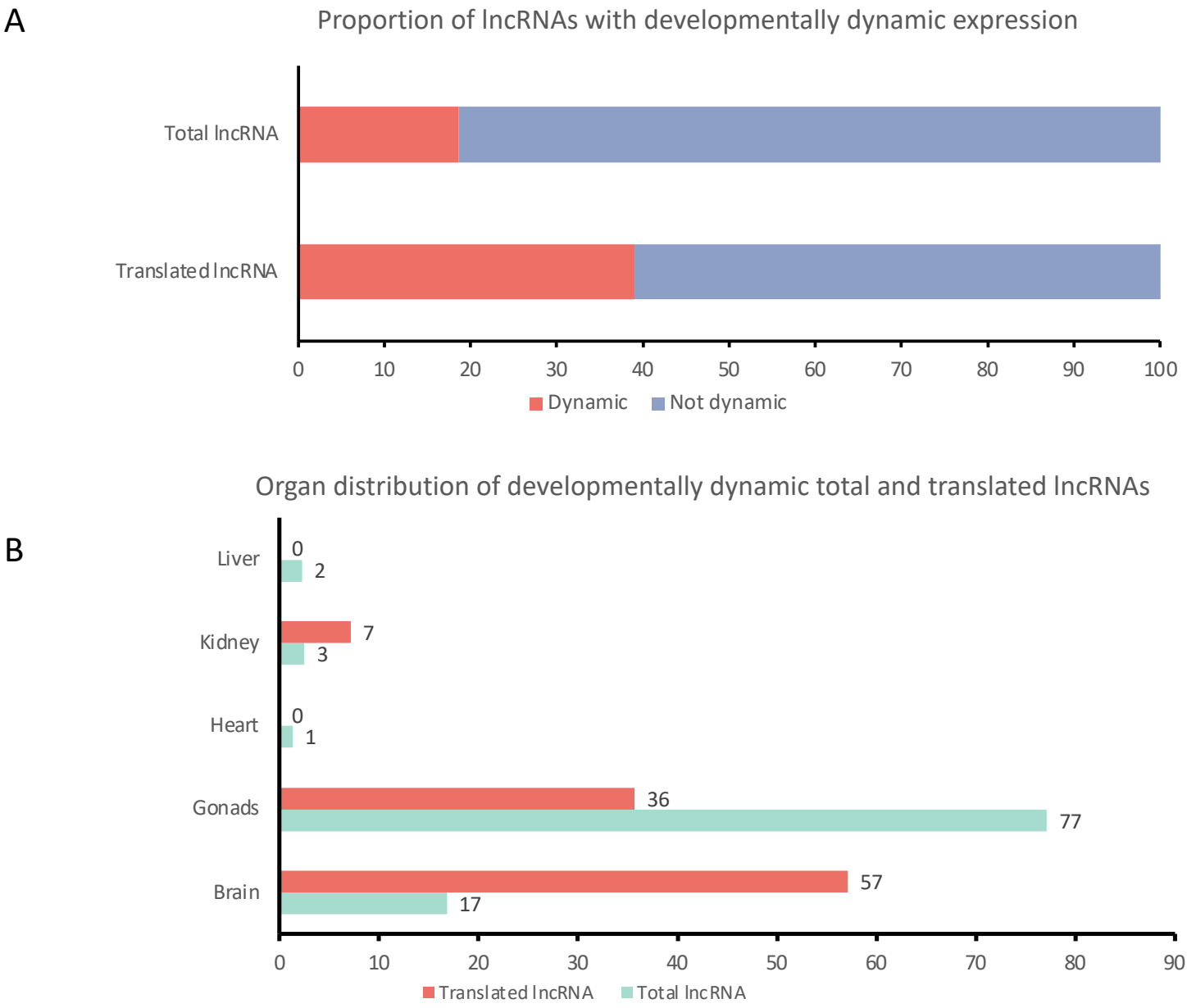


Figure 3.2: Translated lncRNA genes are enriched for developmentally dynamic expression in the brain. *(A) Percentage of all lncRNAs and our translated lncRNAs that show developmentally dynamic expression in humans. (B) Organ distribution of developmentally dynamic lncRNAs for total lncRNAs (green) and our translated lncRNA population (red). Data collected from lncExpDB* (Cardoso-Moreira et al., 2019; Sarropoulos et al., 2019).

3.2.3 Over half of translated lncRNA genes are associated with neuronal disease

To identify lncRNA candidates most likely to be clinically relevant, the disease association of the 35 translated lncRNAs genes was assessed. Utilising multiple disease databases (Bao et al., 2019; J.-R. Li et al., 2016; Papatheodorou et al., 2019), translated lncRNAs were found to be associated with neuronal diseases, as defined by a lncRNA gene showing significant differential expression >$\log_2$ fold-change of >1 in diseased samples compared to control, or if a lncRNAs role in disease has been experimentally validated. Perhaps unsurprisingly, as Poly-Ribo-Seq was performed in neuroblastomas cells (SH-SY5Y), 68 % of translated lncRNAs were associated with neuronal cancers (Figure 3.3).

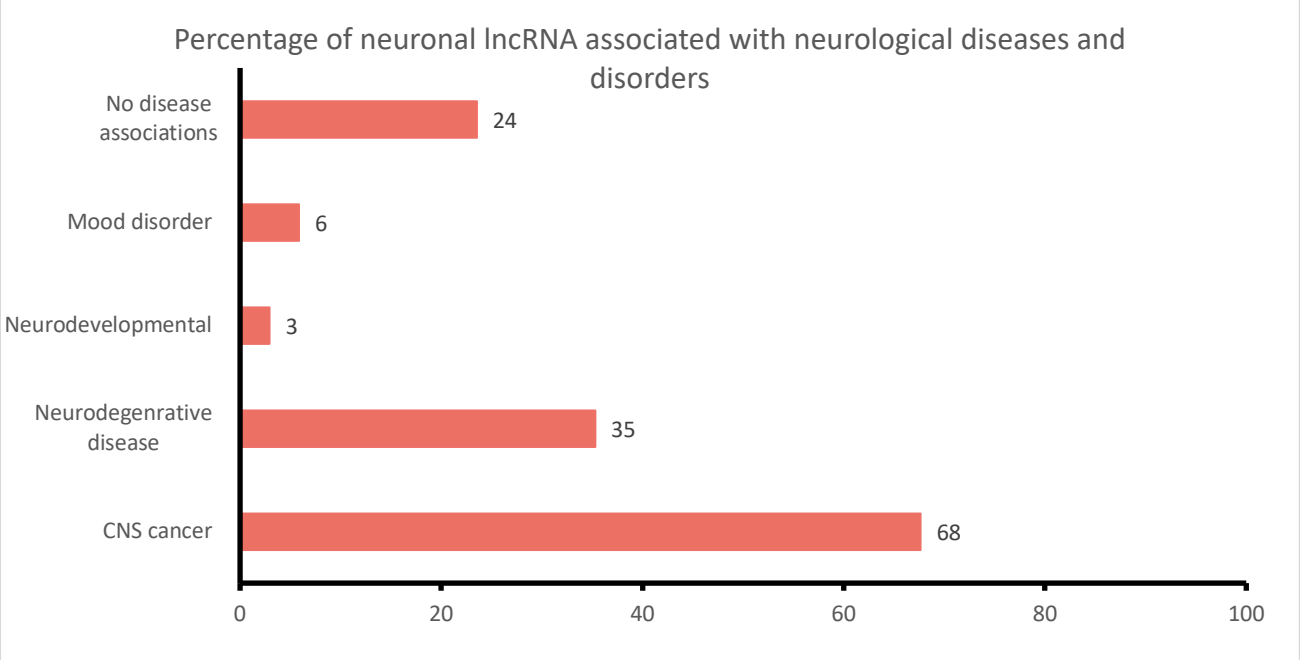


Figure 3.3: Over half of translated lncRNA genes are deregulated in neuronal cancers. *Percentage of translated lncRNA genes (red) found to be associated with neuronal diseases and disorders according to lncRNADisease, Differential Expression Atlas, Cancer RNA-Seq Nexus. X axis is percentage.* (Bao et al., 2019; J.-R. Li et al., 2016; Papatheodorou et al., 2019).

3.2.4 Analysis of FANTOM6 lncRNA knockdown data reveals that silencing of three translated lncRNAs downregulates transcript within neuronal gene sets

The FANTOM (Functional ANnoTation Of the Mammalian genome) project is an international consortium focused on functional annotation of the human genome. Previous FANTOM projects have focused on the generating accurate transcriptomes and transcriptional start sites with CAGE (Cap Analysis Gene Expression) in mouse and human cells (Kawaji et al., 2017). Iteration 6 of the FANTOM project investigated lncRNA cellular functions by performing systematic antisense oligonucleotides silencing of 285 individual lncRNA genes, randomly selected from FANTOM lncRNA annotations. This including 3 identified as translated by Poly-Ribo-Seq in SH-SY5Y cells; FGD5-AS1, LINC01116 and TUG1 (Ramilowski et al., 2020). Knockdowns were performed in nontransformed human primary dermal fibroblasts and silencing of these 3 genes resulted in dysregulation of neuronal pathways as identified by RNA-Seq and gene set enrichment analysis (Table 3.1). This finding is particularly striking as the knockdowns were performed in non-neuronal cells.

lncRNA	Downregulated pathways and Normalised enrichment score
LINC01116	GO_MYELIN_SHEATH
	GO_NEURON_MIGRATION
	GO_LARGE_RIBOSOMAL_SUBUNIT
	GO_SMALL_RIBOSOMAL_SUBUNIT
	GO_POSITIVE_REGULATION_OF_CANONICAL_WNT_SIGNALING_PATHWAY
FGD5-AS1	GO_NEURON_PROJECTION_GUIDANCE
	GO_NEURON_PROJECTION_MEMBRANE
	GO_NEURON_RECOGNITION
	GO_MAIN_AXON
	GO_SYNAPTIC_SIGNALING
	GO_NEURON_PROJECTION_MORPHOGENESIS
TUG1	GO_SYNAPTIC_MEMBRANE
	GO_SMAD_BINDING
	GO_NEGATIVE_REGULATION_OF_WNT_SIGNALING_PATHWAY
	GO_NEURON_PROJECTION_MORPHOGENESIS
	GO_CELL_MORPHOGENESIS_INVOLVED _IN_NEURON_DIFFERENTIATION
	GO_CANONICAL_WNT_SIGNALING_PATHWAY

Table 3.1: GO terms pathways significantly altered following Antisense oligonucleotides silencing of 3 translated lncRNAs (LINC01116, FGD5-AS1, TUG1) in human dermal fibroblasts. Data from FANTOM6 (Ramilowski et al., 2020).

3.2.5 Translated lncRNAs are enriched in the cytoplasm

For an RNA to be translated it must be exported from the nucleus into the cytoplasm. To assess the proportion of translated lncRNAs that are found predominantly in the cytoplasm publicly available data was analysed. LncAtlas contains RNA-Seq data performed on nuclear and cytoplasmic fractions of multiple cell types and produces a Relative Concentration Index (RCI) of cytoplasmic FPKM/ nuclear FPKM for a specific gene (Mas-Ponte et al., 2017). Therefore, RNAs enriched in the cytoplasm will have a positive RCI value. Only 28 of the 35 translated lncRNAs were present in this database.

RCI data was extracted from the SH-SY5Y precursor cell line, SK-N-SH, for 28 translated lncRNAs and the 88 polysome associated but not translated lncRNAs, which act as a control for cytoplasmic lncRNAs, without nuclear lncRNAs to skew the average RCI. Translated lncRNAs had a mean RCI of 0.3 compared to non-translated lncRNAs with a mean RCI of -1.3 (Figure 3.4). This indicates translated lncRNAs are on average more cytoplasmic than non-translated lncRNAs. Individual RCI values were considered when selecting lncRNAs for further investigation.

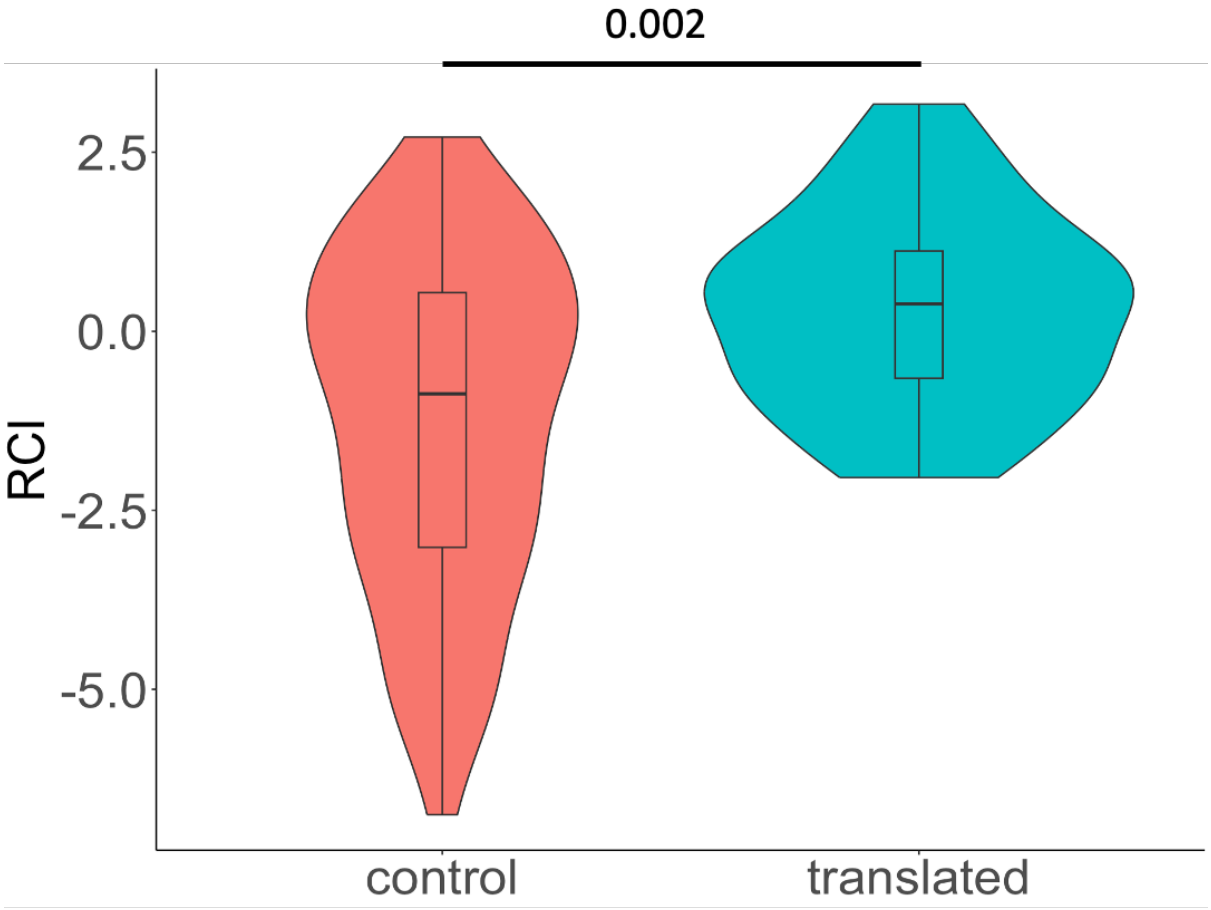


Figure 3.4 Translated lncRNAs are more cytoplasmic than lncRNAs SK-N-SH in general. *RCI data collected from lncAtlas on the cell type SK-N-SH. RCI score for 28 translated lncRNAs and 88 polysome associated but not translated lncRNAs were collected from the SK-N-SH cell line. The mean RCI of translated lncRNAs is significantly higher than non-translated lncRNAs. Significance determined by Mann-Whitney U Test, p value reported on plot.*

3.2.6 Two peptides translated from lncRNAs are predicted to contain protein domains

To investigate potential structural elements and domains in the proteins translated from lncRNAs two tools were used. Firstly, Phyre2 was used to identify structural homology with known proteins. It predicts secondary structures in peptide sequences and matches them to known protein structures (Kelley et al., 2015). Two proteins translated from our candidate lncRNAs, had structural orthologs identified by Phyre2: LIPT2-AS1 and EMSLR. LIPT2-AS1 protein is predicted to contain a helix-turn-helix DNA binding domain, similar to that seen in the canonical DNA binding proteins CenpB and Jerky protein homolog. Phyre2 predicts EMSLR protein to form a structure similar to the C-terminus of the alpha and beta subunits of F1 ATP synthase.

Additionally, predicted amino acid sequences from translated lncRNAs were inputted to EMBL-EBI Sequence similarity search (SSS, (Madeira et al., 2022)) to identify if any had sequence homology to known proteins. Only one high confidence sequence match was identified (E value <1) for LIPT2-AS1, again predicting LIPT2-AS1 protein to contain a DNA binding domain (Table 3.2).

Gene	EMBL-EBI SSS	Phyre2
LIPT2-AS1	HTH domain	CenpB and Jerky protein homolog
EMSLR	N/A	C-terminal domain of alpha and beta subunits of F1 ATP synthase

Table 3.2: Summary of novel peptide structural predictions from Phyre2 and SSS.

3.3 lncRNA-ORF expression from mini-gene reporter results in stable micropeptide production in SH-SY5Y cells

Based on their tissue and subcellular RNA expression profile, relation to neuronal disease or neuronal development 13 candidate translated (s)ORFs were selected for validation of peptide production. The FLAG-tagged expression assay aimed to identify (s)ORFs that produced stable peptide products and if these peptides exhibited specific subcellular localisation indicative of organelle recruitment, and therefore function.

lncRNA-ORFs were cloned into pcDNA3.1 mammalian expression plasmid with a C-terminal 3X FLAG tag, which does not have its own ATG (Figure 3.5). Therefore, any FLAG signal seen will be due to the translation of the novel ORF. A 3X FLAG-tag was selected as highly specific antibodies are available for its detection or purification. Additionally, FLAG is small and hydrophilic so is unlikely to disrupt the micropeptide of interest's localisation or confirmation through steric hinderance or its own localisation. The initial screen was performed in immature SH-SY5Y cells. Of the 13 sORFs tested in this FLAG-transfection assay 12 out of 13 produced FLAG signal in SH-SY5Y cells, indicating these lncRNA-ORFs can produce stable peptides (Figures 3.6, 3.7) (Table 3.3).

Peptides exhibiting a discreet subcellular localisation are more likely to be functional than diffuse peptides. Of the 12 translated lncRNA (s)ORFs that produced a stable FLAG signal, 5 showed specific subcellular localisations. Additionally, one candidate peptide, OLMALINC, was enriched in the nucleus compared to the cytoplasm (Figure 3.6).

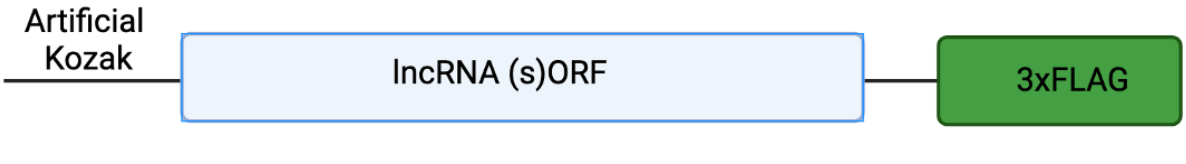


Figure 3.5: Schematic of pcDNA3.1-(s)ORF-FLAG construct. *(s)ORFs were cloned from SH-SY5Y cDNA from their Poly-Ribo-Seq predicted ATG to the end of (s)ORF, excluding their stop codon. An artificial Kozak of GCCACC was used upstream of AUG start codon. A small linker of 2 codons follows the ORF before a 3xFLAG tag that lacks its own ATG.*

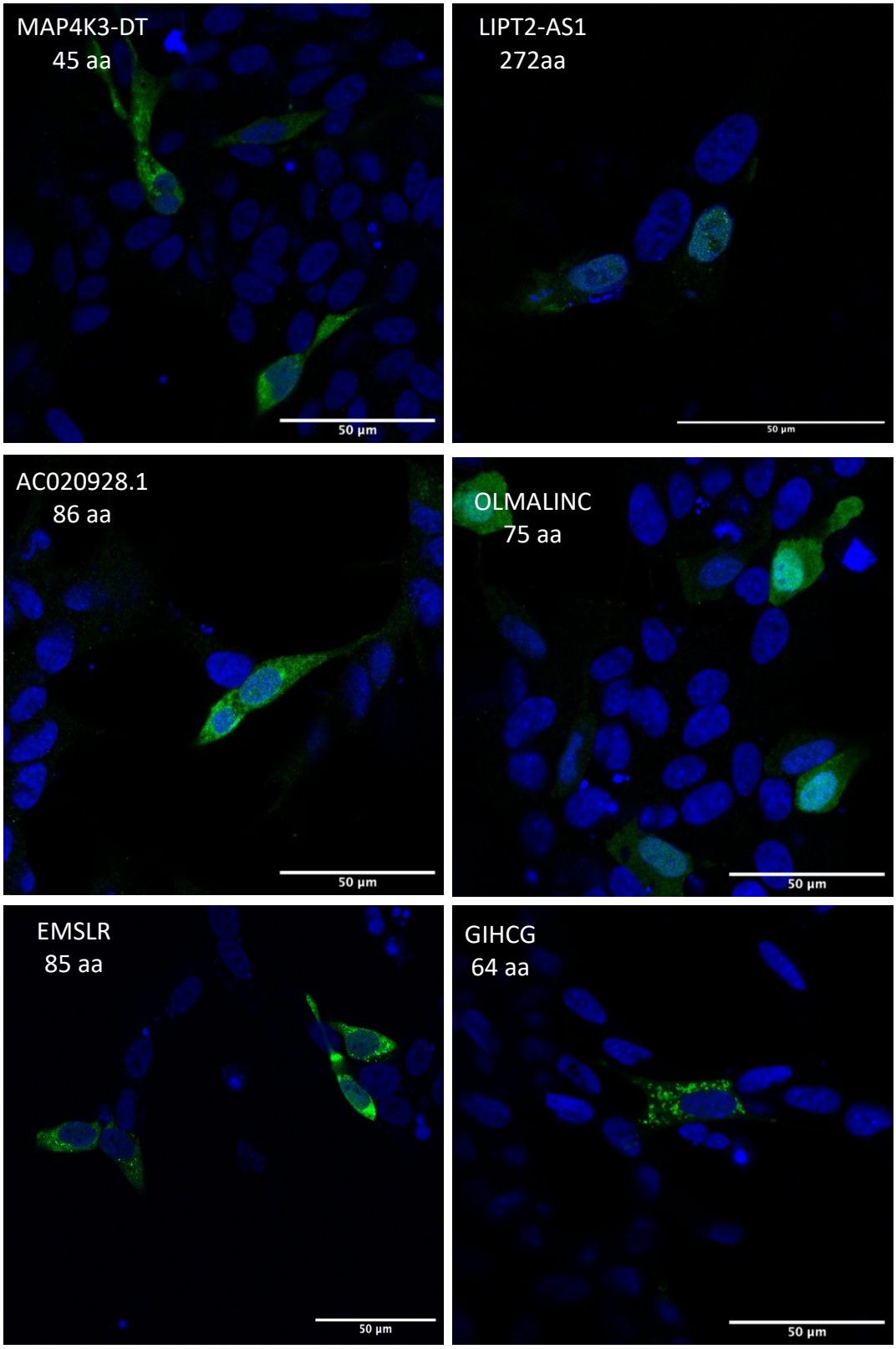


Figure 3.6: FLAG-tagged (s)ORFs from lncRNAs translated in SH-SY5Y cells localise to discreet subcellular localisations. *Representative images of immature SH-SY5Y cells transiently transfected with pcDNA-(s)ORF-FLAG. FLAG was detected with M2 anti-FLAG (Sigma, 1:1000) and labelled with Alexa 488 secondary (Green). Cell nuclei are stained blue with DAPI. Images were obtained with an LSM880 confocal microscope (Zeiss) at 40X magnification using a Pln Apo 40X/1.3 Oil DICIII objective. Scale bar= 20µm.*

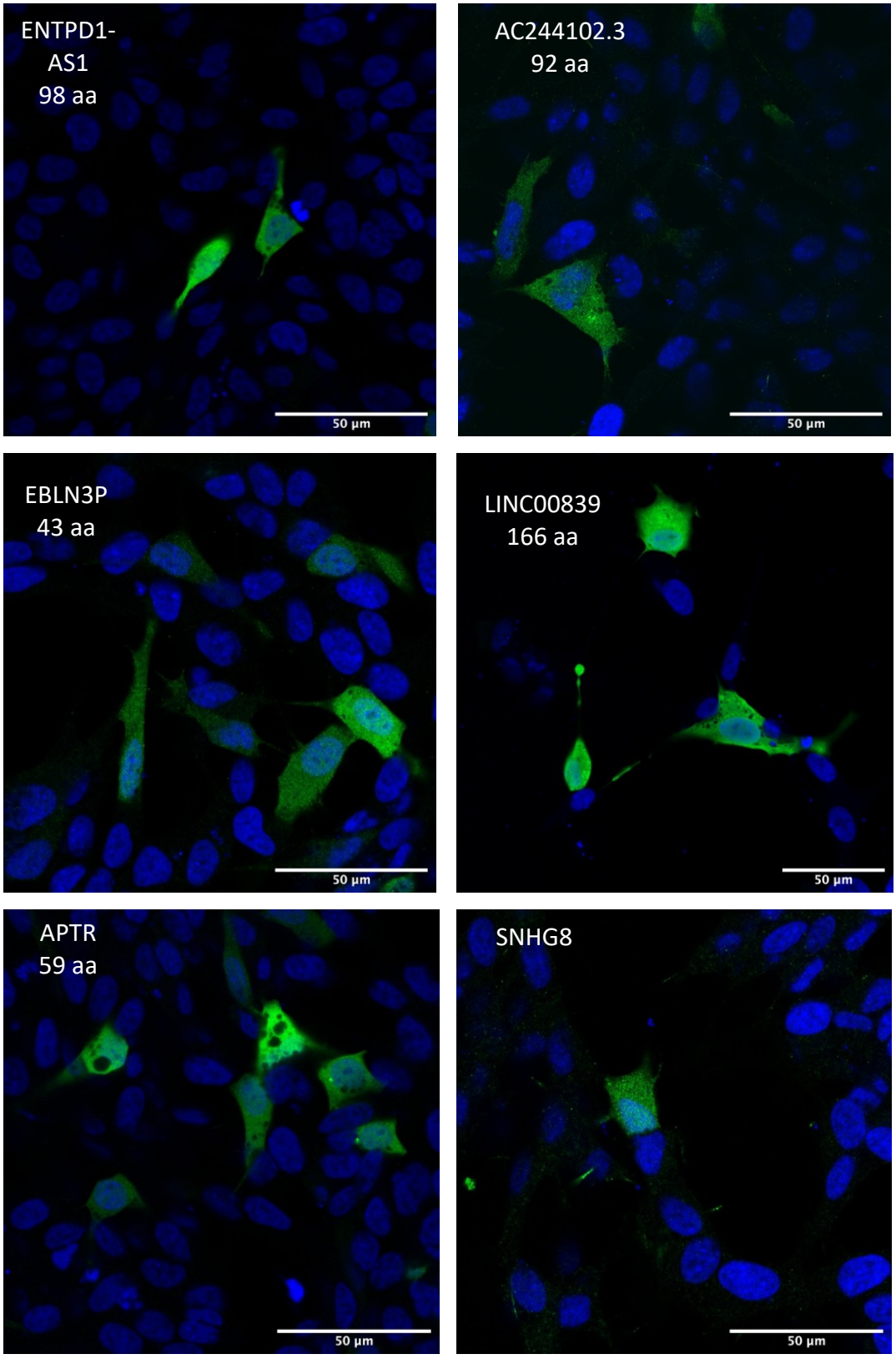


***Figure 3.7*: FLAG-tagged (s)ORFs from lncRNAs translated in differentiated SH-SY5Y cells show a diffuse cellular localisation.** *Representative images of immature SH-SY5Y cells transiently transfected with pcDNA-(s)ORF-FLAG. FLAG was detected with M2 anti-FLAG (Sigma, 1:1000) and labelled with Alexa 488 secondary (Green). Cell nuclei are stained blue with DAPI. Images were obtained with an LSM880 confocal microscope (Zeiss) at 40X magnification using a Pln Apo 40X/1.3 Oil DICIII objective. Scale bar= 20µm.*

Gene name	Transcript ID	Peptide length (amino-acids)	Condition translated	Signal in SH-SY5Y
LINC00839	ENST00000429940	166	Differentiated	Cytosolic and nuclear
APTR	ENST00000440088	59	Differentiated	Cytosolic and nuclear
SNHG8	ENST00000602414	36	Differentiated	Cytosolic and nuclear
GIHCG	ENST00000547492	64	Differentiated	Cytoplasmic puncta
EMSLR	ENST00000419422	85	Differentiated	Cytoplasmic puncta
MAP4K3-DT	ENST00000445520	45	Immature	Cytoplasmic puncta
LIPT2-AS1	ENST00000526036	272	Immature	Nuclear
ENTPD1-AS1	ENST00000449419	98	Immature	Cytosolic and nuclear
OLMALINC	ENST00000454935	75	Immature	Nuclear enriched + cytoplasmic
AC244102.3	ENST00000651163	92	Immature	Cytoplasmic puncta
AC020928.1	ENST00000590750	86	Immature	Cytoplasmic puncta
LINC01515	ENST00000608678	17	Immature	No signal
EBLN3P	ENST00000625445	43	Immature	Cytosolic and nuclear

Table 3.3: Summary of translated lncRNAs included in the FLAG tagged reporter assay and their localisations within immature SH-SY5Y cells.

3.4 Confirmation of translation initiation site for top translated lncRNA-ORF candidates

RiboTaper indentifies ORFs based on Poly-Ribo-Seq reads mapped to the transcriptome to identify regions of RNA between an AUG and STOP that contain Ribo-Seq reads that demonstrate high levels of triplet periodicity (Calviello et al., 2016). However, while investigating translated transcript sequences, mutiple candidates possessed upstream, in-frame AUGs that were not detected by RiboTaper to be the start codons and therefore not called as the start site of (s)ORFs.

To determine which of the potential start codons is being used for our top 6 ORFs of interest bioinformatic analysis was performed using the predictive tool Tis-transformer (Clauwaert et al., 2023). Additionally, (s)ORFs were cloned into the FLAG tagged reporter with their endogenous 5'UTRs (5'UTR-sORF-FLAG plasmids, Figure 3.8). The size of translation products was then assessed by western blot. The localisation of peptides translated from these 5'UTR-sORF-FLAG constructs was also determined by immunostaining.

The 6 candidates selected for further investigation and the data gathered on them are summarised in table 3.4. These candidates were selected as showed robust expression in the mini-gene reporter assay and have literature or RNA expression data suggesting potential roles in neuronal development and/or disease.

3.4.1 Tis-transformer predicts 3/6 (s)ORFs are larger than RiboTaper predicted

Tis-transformer is a deeplearning model that predicts translational start sites from nucleotide sequence alone (Clauwaert et al., 2023). Tis-transformer was used to analysis the sequence of GIHCG, MAP4K3-DT, APTR, LINC00839, OLMALINC and LIPT2-AS1 and predict their translational start sites. The TIS-transformer predicted ATGs for GIHCG, MAP4K3-DT and APTR matched those identified by RiboTaper (Figure 3.9). Additionally, Tis transformer also predicts a potential downstream ORF in MAP4K3-DT (Figure 3.9 B).

For LIPT2-AS1 Tis-transformer predictes that the main ORF starts 360 nucleotides upstream of the RiboTaper called start site. Additionally, Tis-transformer highlights a potential uORF, which may explain the low translation efficience seen in the Poly-Ribo-Seq (Douka et al.,

2021) (Figure 3.10 A). Additionally, both LINC00839 and OLMALINC ORFs were predicted to be larger than origionally thought by 66 and 144 nucleotides, respectivly (Figure 3.10 B-C).

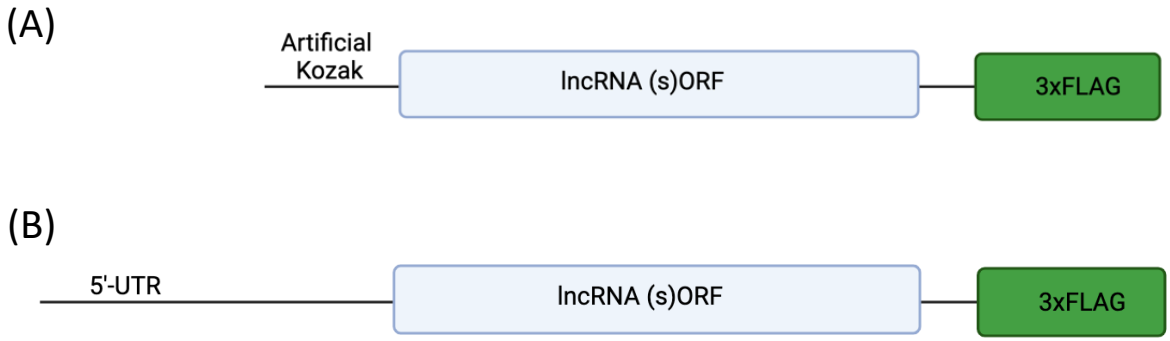


Figure 3.8: Schematic of FLAG tagged expression constructs. *(A) (s)ORF only constructs starting from RiboTaper predicted ATG with an artificial Kozak (referred to as pcDNA-(s)ORF-FLAG). (B) (s)ORFs of interest were cloned again to include their endogenous 5'-UTR and again inserted into pcDNA 3.1 – FLAG for transient transfection into SH-SY5Y cells (referred to as 5'UTR-sORF-FLAG).*

Gene name	length (aa)	Conservation	MS	Disease relation	Developmental dynamics (Organ)	Peptide localisation in SH-SY5Y cells	TE	Predicted domain	Publication
LIPT2-AS1	**392**	**x**	**x**	**Dysregulated in non-neuronal cancers including breast and colon**	**Not dynamic**	**Nuclear**	**0.29**	**DNA-binding (HTH) domain**	(Douka et al., 2021; X. Liu et al., 2022)
APTR	**59**			**Multiple cancers and dysregulated in Alzheimer's disease**	**Not dynamic**	**Cytoplasmic puncta and diffuse throughout cell**	**1.59**		(Guan et al., 2019; Ren et al., 2022)
GIHCG	**64**	**x**	**x**	**Multiple cancers including Glioma and medulloblastoma**	**Testis**	**Mitochondrial co-localisation**	**0.22**		(Jiang et al., 2019; Sandmann et al., 2023a; S.-Y. Zhu et al., 2022)
LINC00839	**188**	**x**	**x**	**Cancers including neuroblastoma**	**Not dynamic**	**Cell periphery**	**2.07**		(Xie et al., 2022; Q. Zhang et al., 2022, Douka et al., 2021)
MAP4K3-DT	**45**			**Dysregulated in Glioblastoma and Alzheimer's disease**	**Brain - adult**	**Cytoplasmic puncta**	**1.04**		(Zheng et al., 2023)
OLMALINC	**123**	**x**	**x**	**Multiple cancers including medulloblastoma**	**Not dynamic but role in oligodendrocyte maturation**	**Nuclear**	**0.91**		(W. J. Lee et al., 2021; S. Liu et al., 2020; Mills et al., 2015, Prensner et al., 2021)

Table 3.4: Summary of 6 translated lncRNAs selected for further investigation and siRNA screen. Proteins were marked as conserved if sequence homology identified outside of humans. Disease relation and MS status were determined from literature search. Translation efficiency (TE) calculated by Dr Isabel Birds from SH-SY5Y Poly-Ribo-Seq data (Douka et al., 2021).

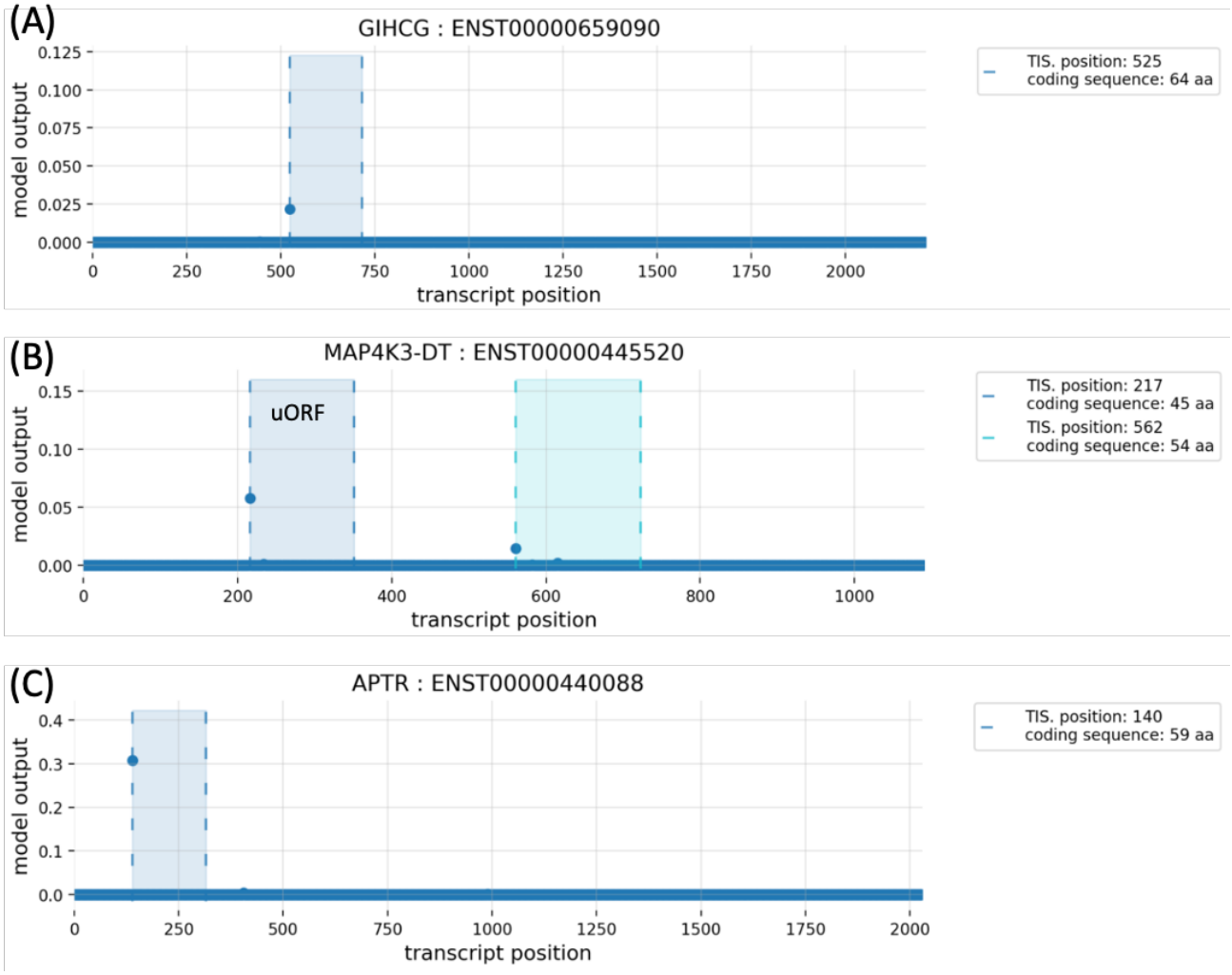


Figure 3.9: Tis-Transformer plots showing predicted strength of translational start sites of translated lncRNAs. *The candidates GIHCG, MAP4K3-DT and APTR are all predicted to be the same size as determined by RiboTaper. Dotted lines indicate in frame start and stop sites for a given ORF and dots represent Tis score (predicted strength) of AUGs in sequence.*

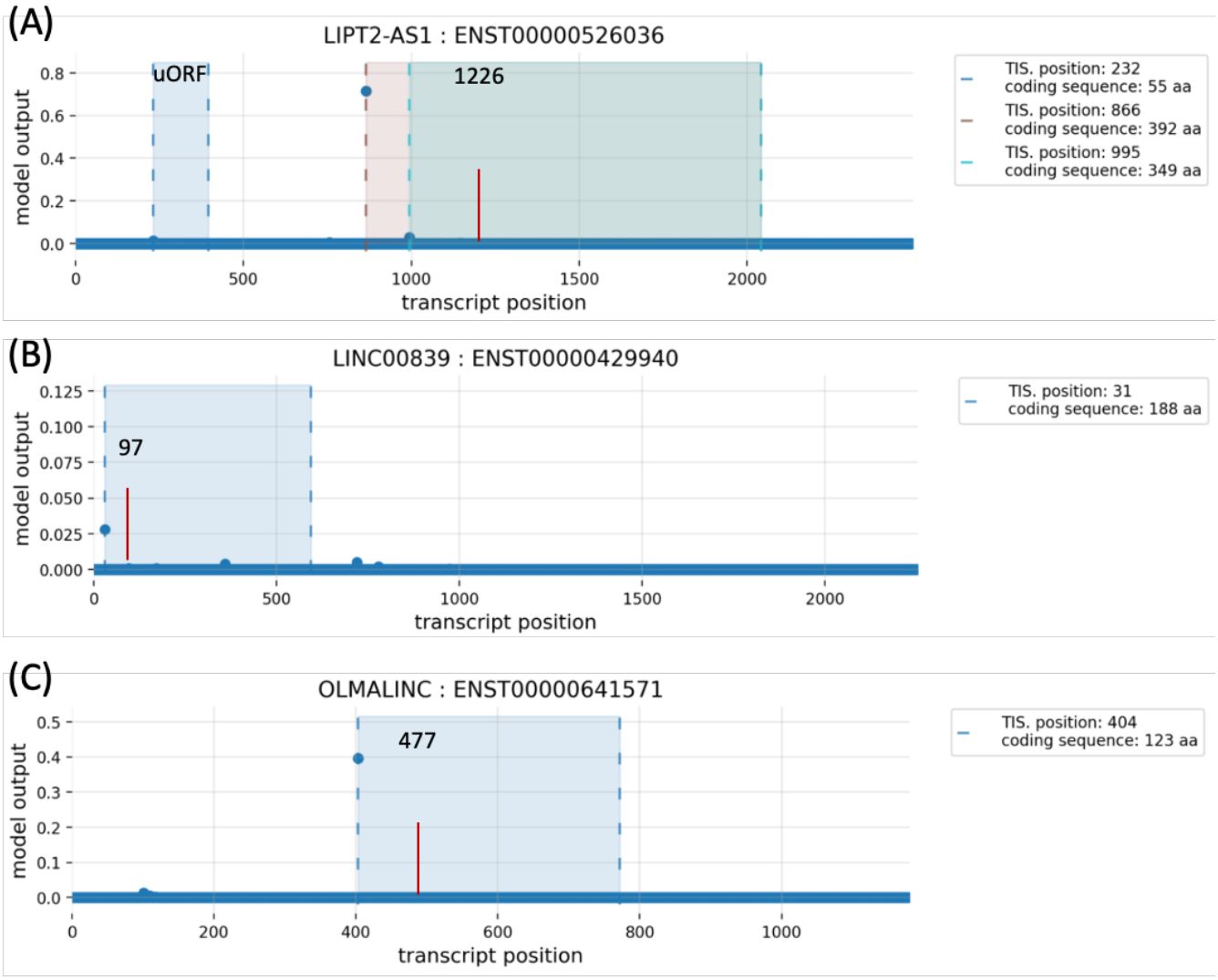


Figure 3.10: Tis-Transformer plots showing predicted strength of translational start sites of translated lncRNAs. *LIPT2-AS1, LINC00839 and OLMALINC ORFs are predicted to start at upstream of originally determined AUG (annotated with red line). This results in an increase in ORF length of 120 codons for LIPT2-AS1, 22 codons for LINC00839 and 48 codons for OLMALINC.. Dotted lines indicate in frame start and stop sites for a given ORF and dots represent Tis score (predicted strength) of AUGs in sequence.*

3.4.2 Inclusion of sORFs endogenous 5'-UTR supports alternative start sites for multiple sORFs

All six candidates were cloned to include their endogenous 5'UTR. This was particularly paramout for LIPT2-AS1, OLMALINC and LINC00839 following the identification of upstream inframe ATGs in these translated lncRNAs. ATG prediction by Tis-transformer suggested LINC00839s ORF to be 22 aa longer than origionally predicted. When expressing the ORF with its endogenous 5'UTR (5'UTR-(s)ORF-FLAG) a clear difference in localisation can be seen compared to the ORF alone (e.g. N-terminal truncated). LINC00839s full length peptide localised to the periphery of SH-SY5Y cells rather than diffuse throughout as the truncated version does (Figure 3.11A).

The other ORFs that were predicted to be longer by TIS-Transformer showed a similar but more specific localisation compared to N-terminal truncated versons of the proteins from ORF alone constructs. LIPT2-AS1 and OLMALINC were predicted to be 120 aa and 48 aa longer, respectively. LIPT2-AS1 ORF-only-FLAG demonstrated a primarily nuclear localisation (Figure 3.11 C) whilst 5'-UTR LIPT2-AS1 formed descrete puncta within the nucleus. Inclusion of OLMALINC's 5'UTR resulted in a peptide that only localised to the nucleus with no cytoplasmic signal (Figure 3.11 B), whilst the truncated ORF shows both nuclear and cytoplasmic localisation (Figure 3.11 B). Table 3.5 summarises the differences between shorter ORFs (ORF-only-FLAG) compared to the longer ORFs (5'UTR-ORF-FLAG) for LIPT2-AS1, OLMALINC and LINC00839. 5'UTR-(s)ORF-FLAG reporters for GIHCG, MAP4K3-DT and APTR showed similar FLAG signal localisation to the ORF only constructs. This was as expected as Tis-Transformer and RiboTaper were in agreement on site of initiation.

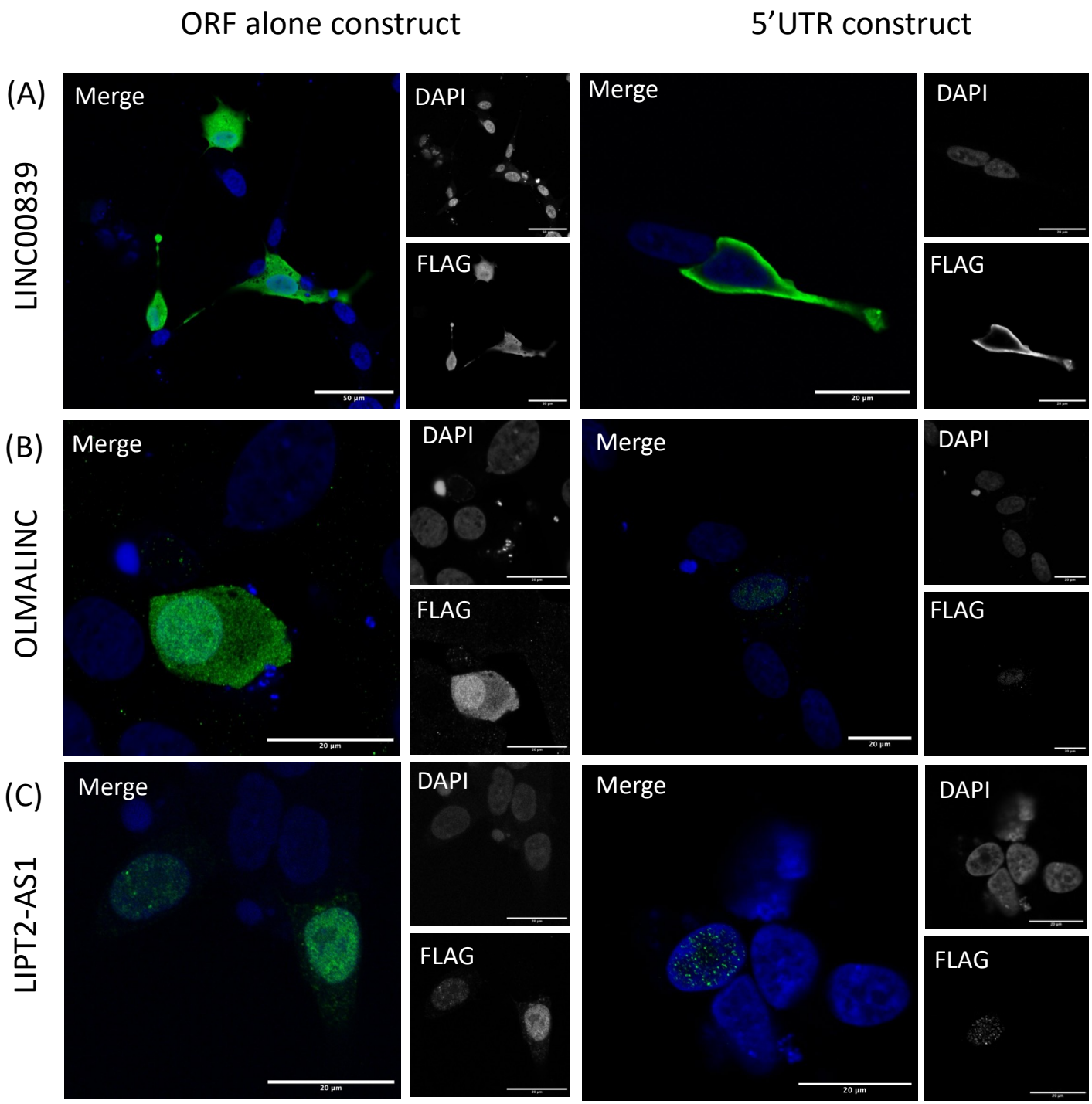


Figure 3.11: Full length ORFs show different localisation to N-terminal truncated versions. *Representative images of immature SH-SY5Y cells transiently transfected with either pcDNA-5'-(s)ORF-FLAG or (s)ORF-alone-FLAG. FLAG was detected with M2 anti-FLAG (Sigma, 1:1000) and labelled with Alexa 488 secondary (Green). Cell nuclei are stained blue with DAPI. Images were obtained with an LSM880 confocal microscope (Zeiss) at 40X magnification using a Pln Apo 40X/1.3 Oil DICIII objective. Scale bar= 20µm.*

Translated lncRNA	ORF size (RiboTaper)	Tis Transformer – ORF size	Truncated localization	Full length localization
LINC00839	166	188	Diffuse	Cell periphery
LIPT2-AS1	272	392 (also predicts potential uORF)	Nuclear (diffuse)	Nuclear foci
OLMALINC	75	123	Nuclear enriched + cytoplasmic	Nuclear (no cytoplasmic)

Table 3.5: Summary of predicted ORF and actual ORF size. Details of peptides and localisation from alternative start sites.

3.11 lncRNA derived peptides show consistent localisation between immature and differentiated SH-SY5Ys

For the 6 candidates selected for further investigation an in-depth localisation analysis was performed including expressing proteins in differentiated SH-SY5Y cells and co-localisation analysis where possible. All candidates showed consistent localisation between immature and differentiated cells (Figures 3.12 – 3.17).

3.11.1 GIHCG, OLMALINC and LIPT2-AS1 show clear localisation to specific organelles

FLAG-tagged GIHCG-peptide shows clear co-localisation with mitochondria stained with Mitotracker CMXRos in both in immature and differentiated SH-SY5Y cells (Figure 3.12). The degree of co-localisation was quantified using Just another co-localization plugin in FIJI (BOLTE & CORDELIÈRES, 2006) and revealed a Mander's coefficient for GIHCG localizing to mitochondria of 0.999. However, no mitochondrial localisation signal detected with TargetP-2.0, so the mechanism by which the peptide is trafficking GIHCG towards the mitochondria is unclear (Almagro Armenteros et al., 2019). The full-length products of LIPT2-AS1 and OLMALINC localises to the nucleus of SH-SY5Y cells and is consistent between immature and differentiated SH-SY5Y cells (Figure 3.13/3.14).

LINC00839 and MAP4K3-DT also show precise subcellular localisations, but their exact co-localisation to an organelle marker has not been determined. LINC00839 is localised to the periphery of immature and differentiated SH-SH5Y cells (Figure 3.15). To elucidate if LINC00839 is present on the exterior of SH-SY5Y cells IF was performed without detergent and therefore antibodies would not be able to enter the fixed cells. Therefore, as LINC00839-FLAG signal could still be seen, this suggests that a proportion of LINC00839 protein is found on the exterior of SH-SY5Y cells (Figure 3.16). MAP4K3-DT peptide forms cytoplasmic puncta in both immature and differentiated cells (Figure 3.17). However, does not co-localise with mitochondria, Golgi, or ER markers. Finally, APTR is diffuse throughout the cell (cytoplasm and nucleus) forming small densities at unknown locations in both immature and differentiated cells (Figure 3.18).

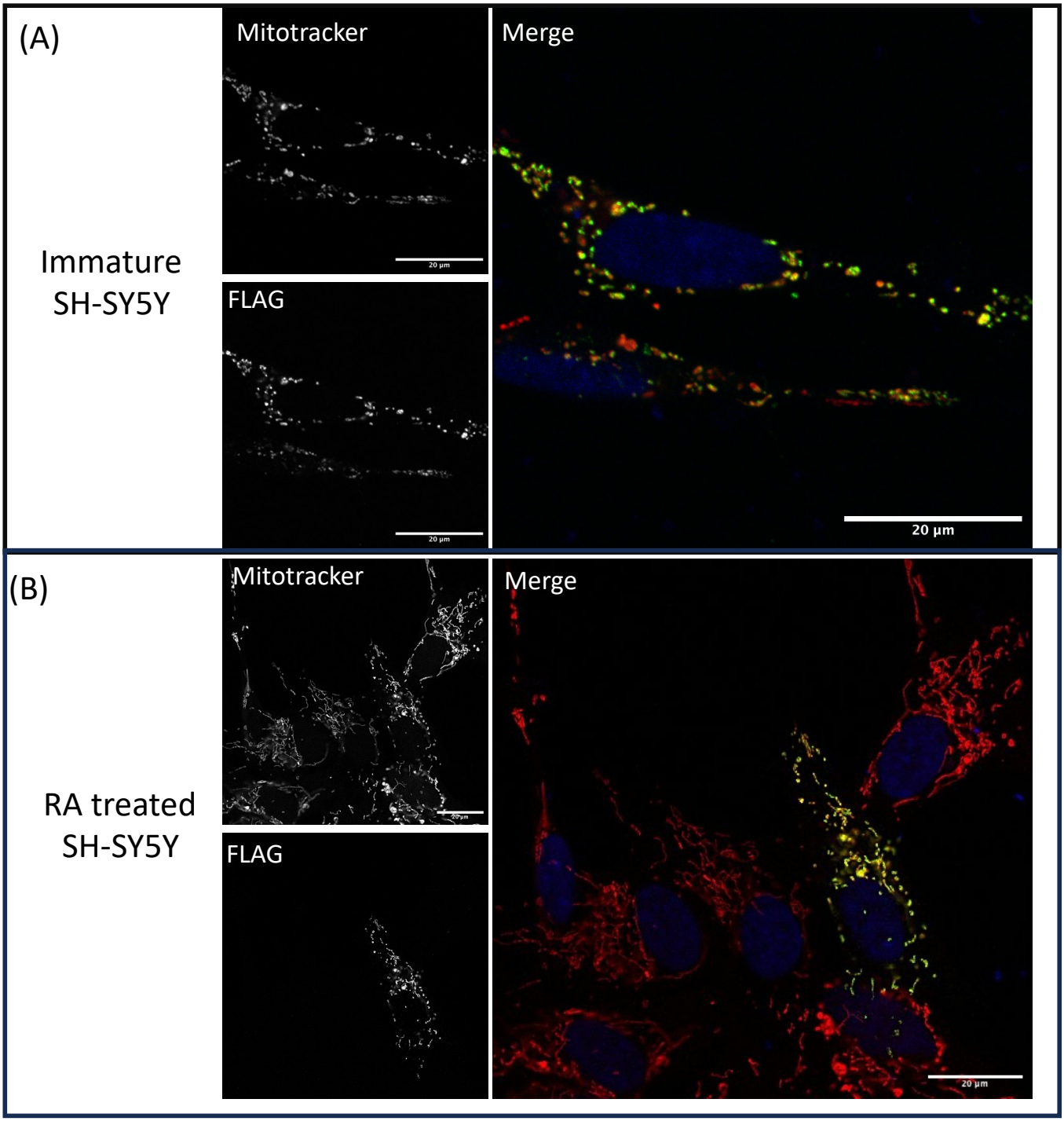


Figure 3.12: GIHCG sORF is translated into a peptide that localizes to the mitochondria of immature and differentiated SH-SY5Y cells. *Representative images of immature (A) and RA treated (B) SH-SY5Y cells transiently transfected with pcDNA-5'-GIHCG-FLAG. FLAG was detected with M2 anti-FLAG (Sigma, 1:1000) and labelled with Alexa 488 secondary (Green). FLAG signal co-localises with mitochondria (stained red with Mitotracker CMXRos). Cell nuclei are stained blue with DAPI. Images were obtained with an LSM880 confocal microscope (Zeiss) at 40X magnification using a Pln Apo 40X/1.3 Oil DICIII objective. Scale bar= 20µm.*

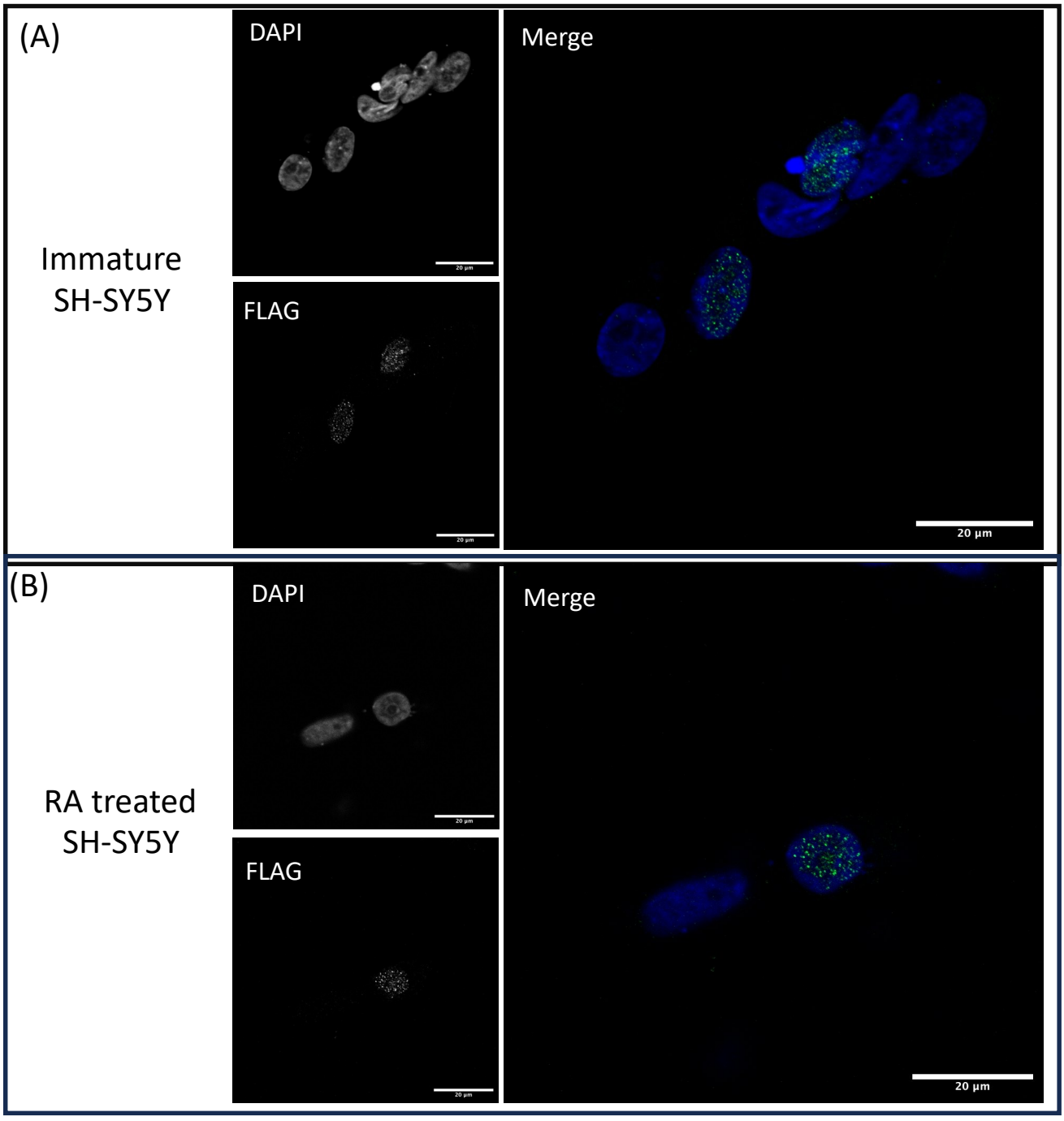


Figure 3.13: LIPT2-AS1 ORF is translated into a peptide that localizes to the nucleus of immature and differentiated SH-SY5Y cells*. Representative images of immature (A) and RA treated (B) SH-SY5Y cells transiently transfected with pcDNA-5'-LIPT2-AS1-FLAG. FLAG was detected with M2 anti-FLAG (Sigma, 1:1000) and labelled with Alexa 488 secondary (Green). FLAG signal co-localises with DNA (stained blue with DAPI). Images were obtained with an LSM880 confocal microscope (Zeiss) at 40X magnification using a Pln Apo 40X/1.3 Oil DICIII objective. Scale bar= 20µm.*

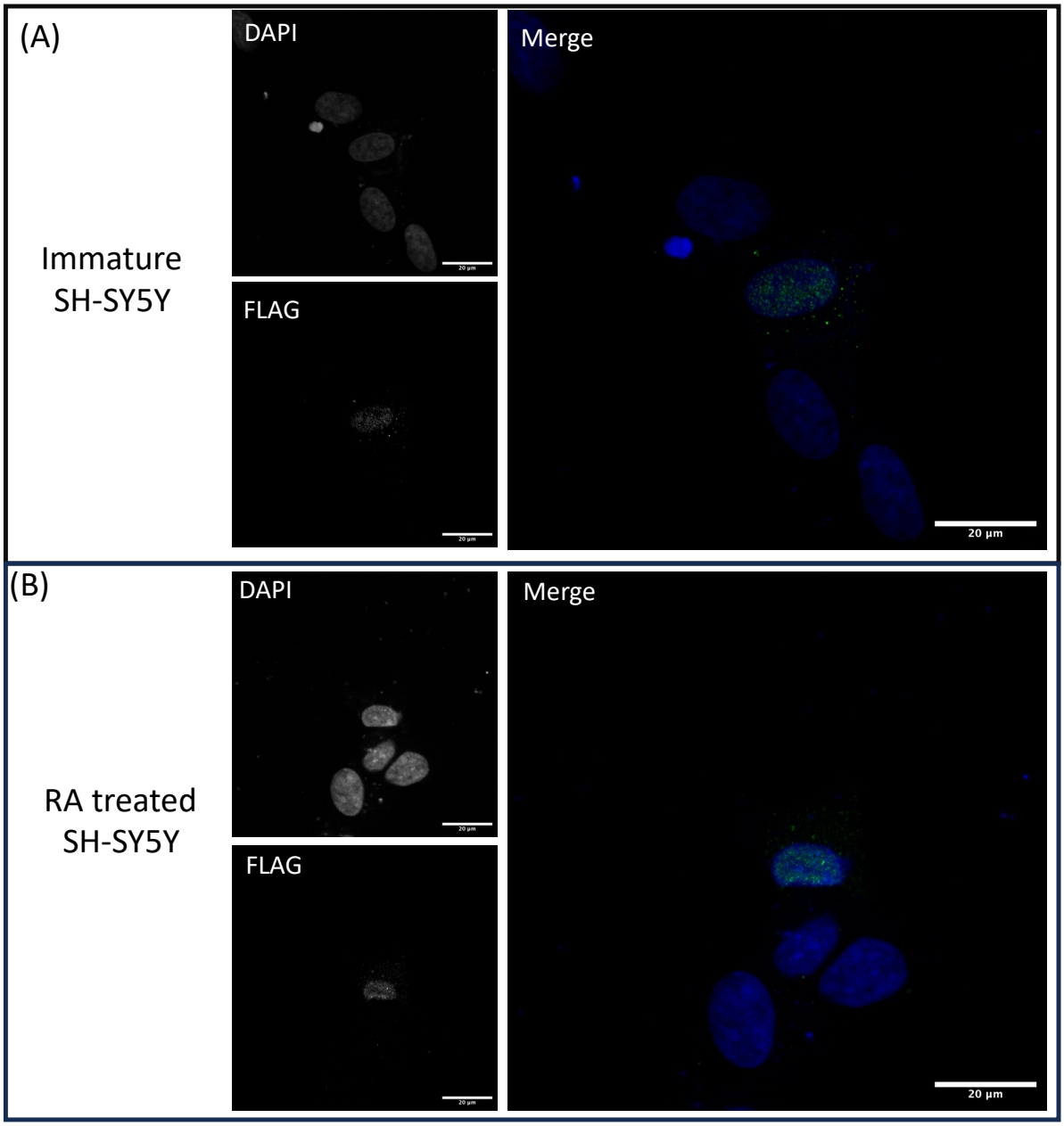


Figure 3.14: OLMALINC sORF is translated into a peptide that localizes to the nucleus of immature and differentiated SH-SY5Y cells. *Representative images of immature (A) and RA treated (B) SH-SY5Y cells transiently transfected with pcDNA-5'-OLMALINC-FLAG. FLAG was detected with M2 anti-FLAG (Sigma, 1:1000) and labelled with Alexa 488 secondary (Green). FLAG signal co-localises with DNA (stained blue with DAPI). Images were obtained with an LSM880 confocal microscope (Zeiss) at 40X magnification using a Pln Apo 40X/1.3 Oil DICIII objective. Scale bar= 20µm.*

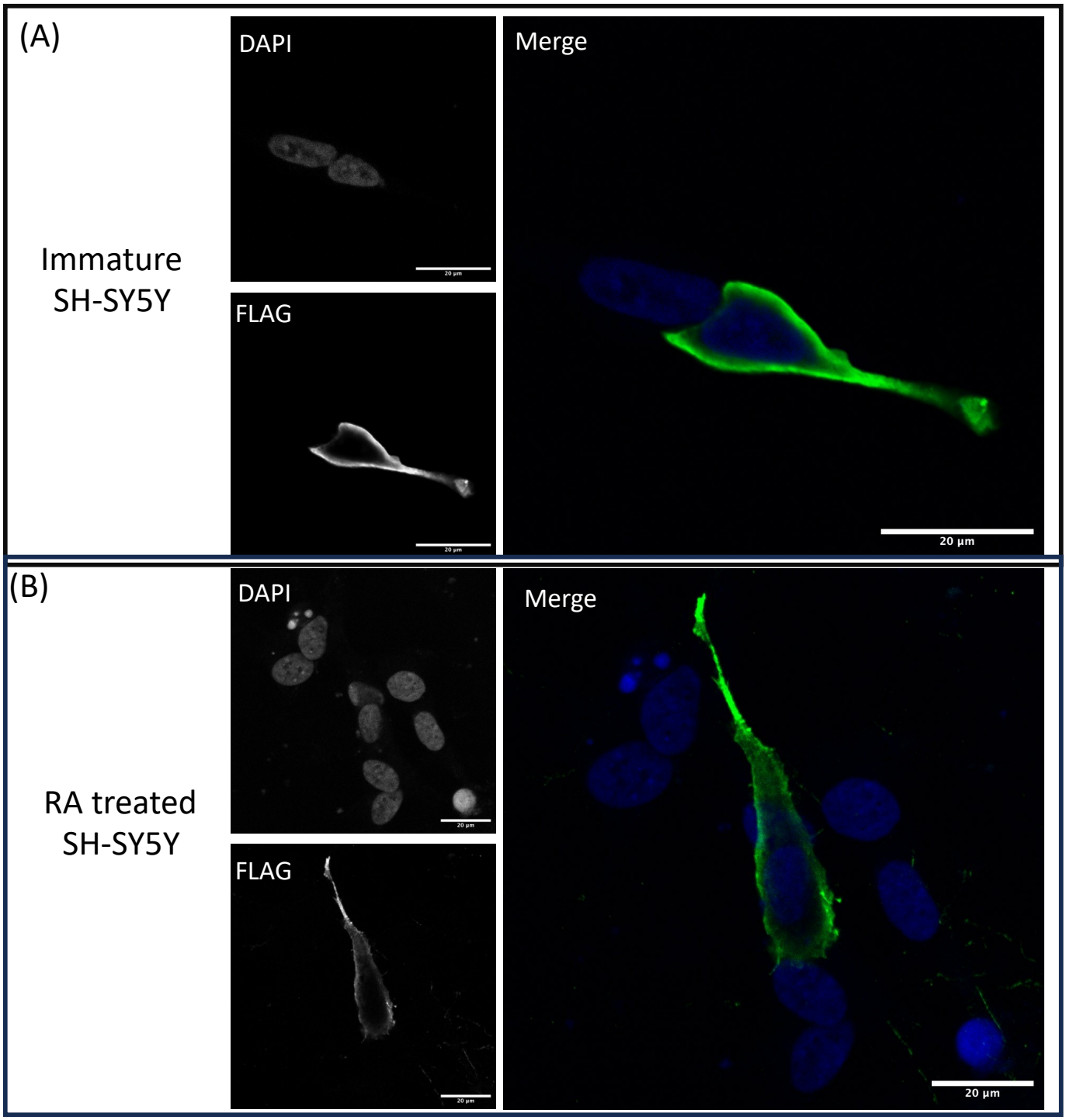


Figure 3.15: LINC00839 ORF is translated into a peptide that localizes to the periphery of immature and differentiated SH-SY5Y cells. *Representative images of immature (A) and RA treated (B) SH-SY5Y cells transiently transfected with pcDNA-5'-LINC00839-FLAG. FLAG was detected with M2 anti-FLAG (Sigma, 1:1000) and labelled with Alexa 488 secondary (Green). Cell nuclei are stained blue with DAPI. Images were obtained with an LSM880 confocal microscope (Zeiss) at 40X magnification using a Pln Apo 40X/1.3 Oil DICIII objective. Scale bar= 20µm.*

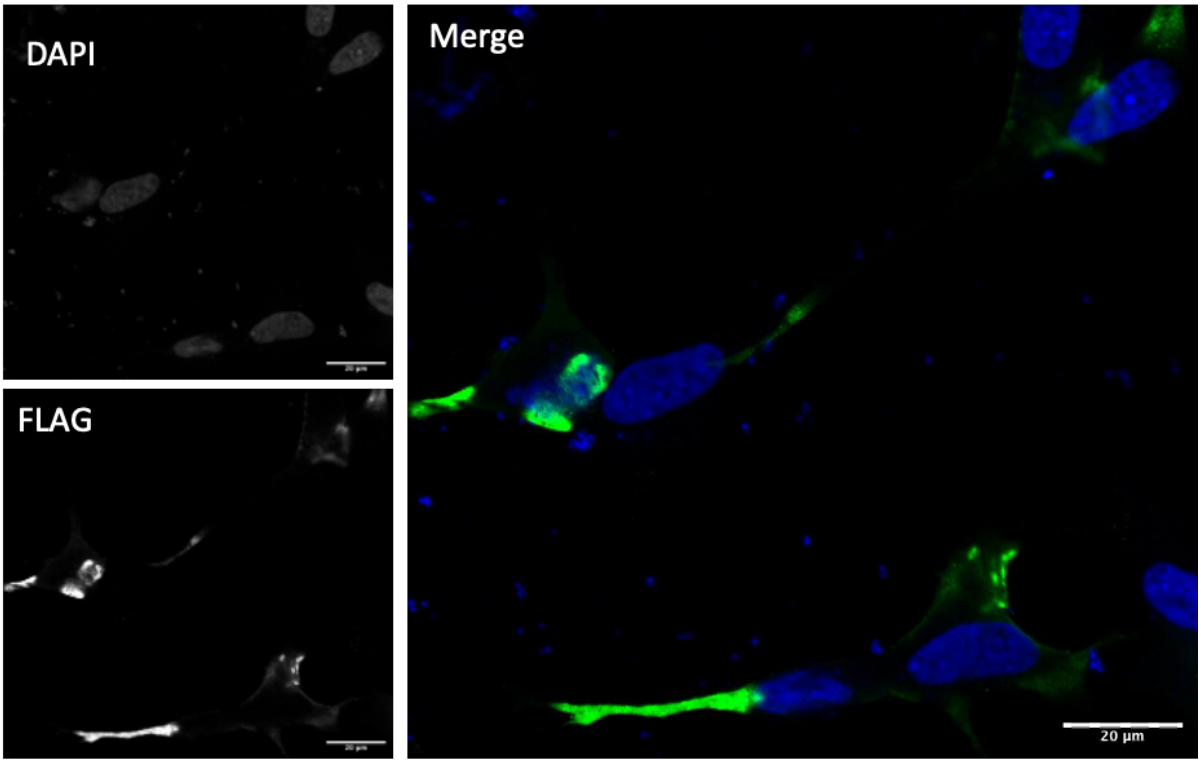


Figure 3.16: LINC00839 ORF is translated into a peptide that can be detected on the periphery of immature SH-SY5Y cells. *Representative images of immature SH-SY5Y cells transiently transfected with pcDNA-5'-LINC00839-FLAG. No detergent (Triton-X) was used and at any point of the IF process meaning the cells will not be permeabilised. FLAG was detected with M2 anti-FLAG (Sigma, 1:1000) and labelled with Alexa 488 secondary (Green). Cell nuclei are stained blue with DAPI. Images were obtained with an LSM880 confocal microscope (Zeiss) at 40X magnification using a Pln Apo 40X/1.3 Oil DICIII objective. Scale bar= 20µm*

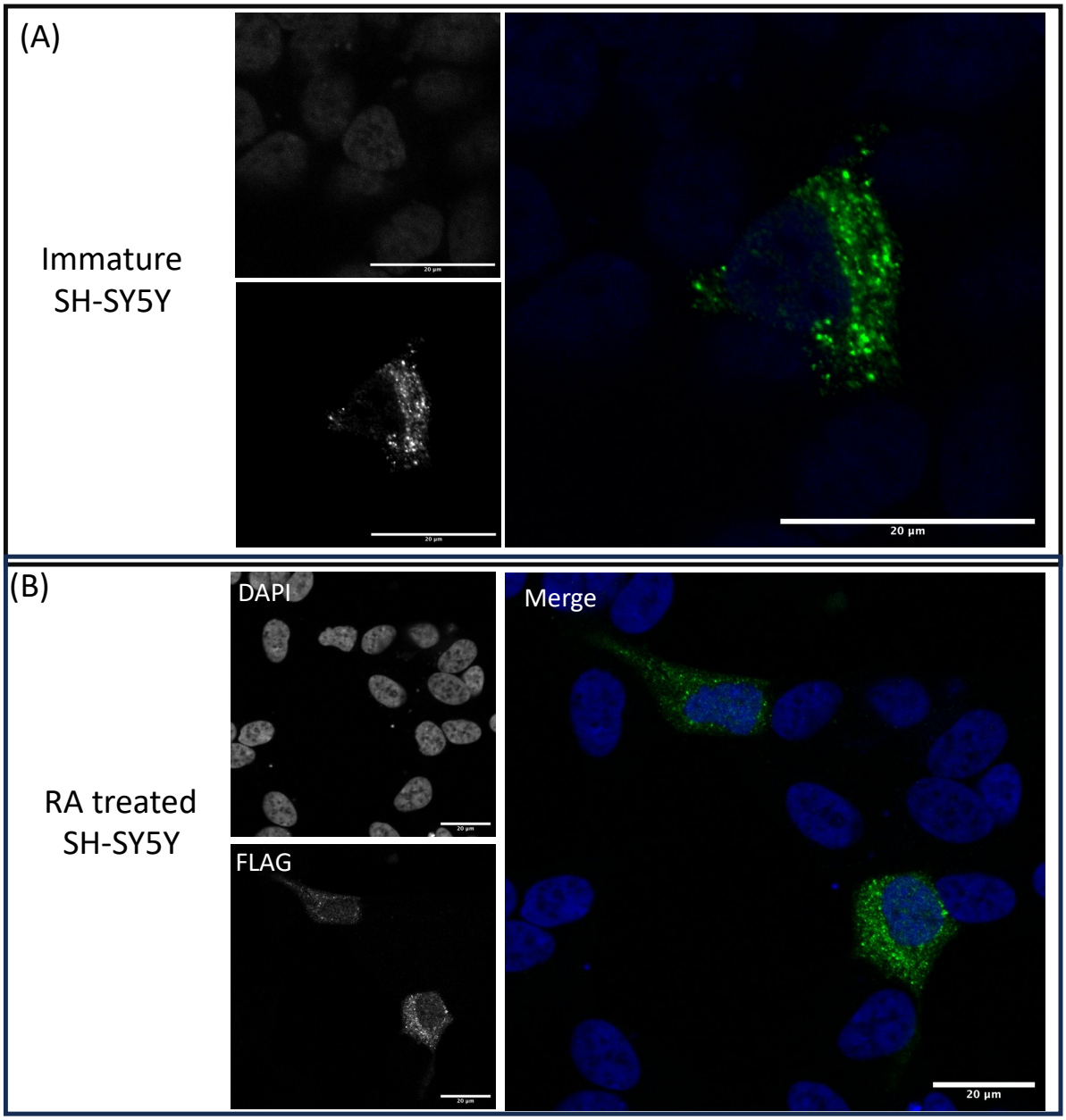


Figure 3.17: MAP4K3-DT ORF is translated into a peptide that forms cytoplasmic puncta in immature and differentiated SH-SY5Y cells. *Representative images of immature (A) and RA treated (B) SH-SY5Y cells transiently transfected with pcDNA-5'-MAP4K3-DT-FLAG. FLAG was detected with M2 anti-FLAG (Sigma, 1:1000) and labelled with Alexa 488 secondary (Green). Cell nuclei are stained blue with DAPI. Images were obtained with an LSM880 confocal microscope (Zeiss) at 40X magnification using a Pln Apo 40X/1.3 Oil DICIII objective. Scale bar= 20µm*

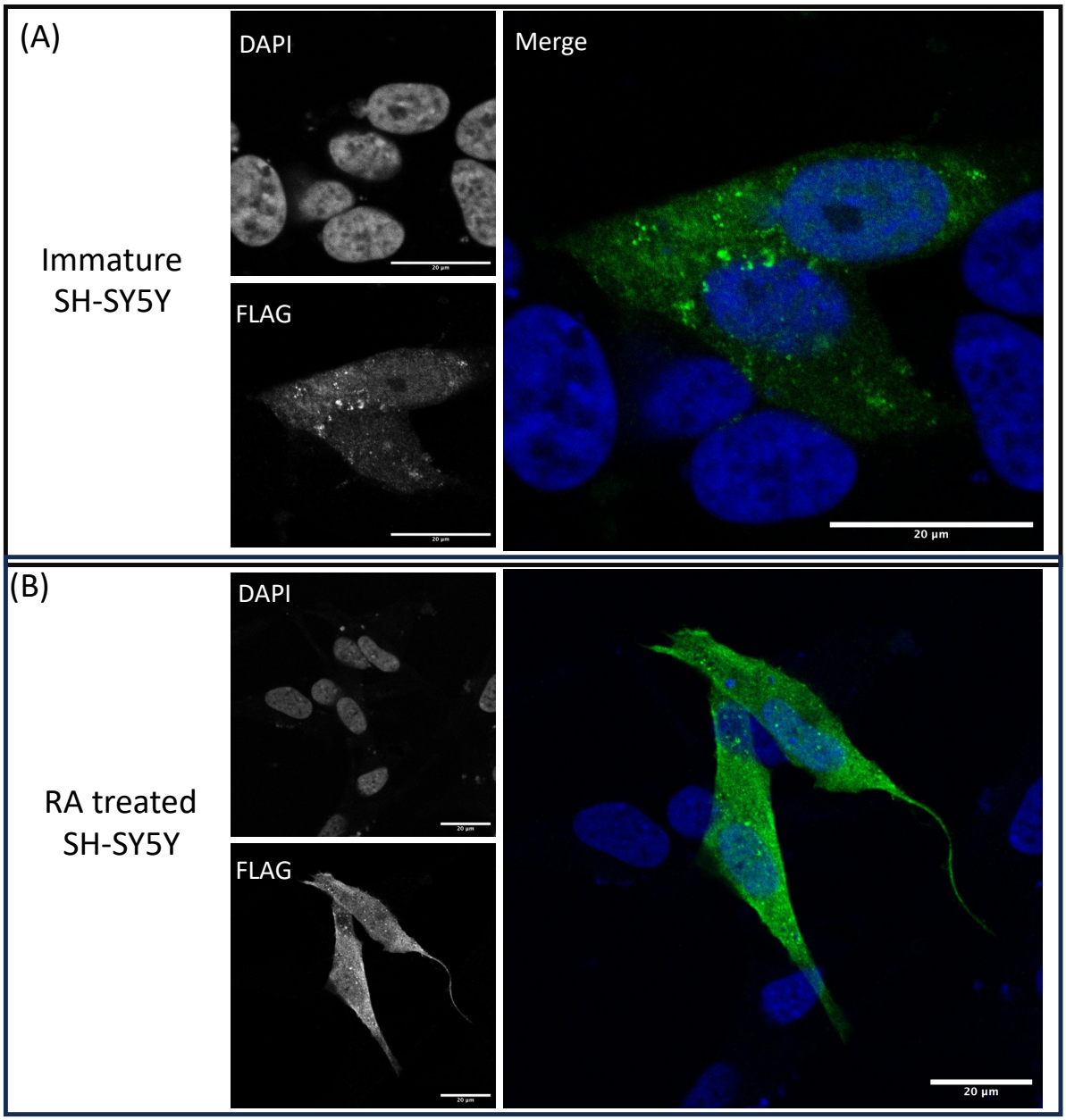


Figure 3.18: APTR ORF is translated into a peptide that localises throughout immature and differentiated SH-SY5Y cells. *Representative images of immature (A) and RA treated (B) SH-SY5Y cells transiently transfected with pcDNA-5'-APTR-FLAG. FLAG was detected with M2 anti-FLAG (Sigma, 1:1000) and labelled with Alexa 488 secondary (Green). Cell nuclei are stained blue with DAPI. Images were obtained with an LSM880 confocal microscope (Zeiss) at 40X magnification using a Pln Apo 40X/1.3 Oil DICIII objective. Scale bar= 20µm.*

3.13 FLAG-tagged novel proteins are detected by western blot

To determine the sizes of the peptides translated from the lncRNA-ORFs, western blot was performed on SH-SY5Y cells transiently transfected with FLAG-tagged (s)ORFs. This was performed for GIHCG, LINC00839, OLMALINC, LIPT2-AS1, APTR and MAP4K3-DT using 5'-UTR-(s)ORF-FLAG reporters (Figure 3.8 B) and detected with FLAG antibody. The size for each novel peptide was predicted based on amino acid length and composition (Stothard, 2000) (Table 3.6). All novel proteins except for MAP4K3-DT could be detected, suggesting the novel peptides are stably expressed. The sizes of the bands matched their predicted kDas (Figure 3.19).

Novel peptide	Predicted kDa (+ 3x FLAG)
GIHCG	10.5
LINC00839	23.1
OLMALINC	16.7
LIPT2-AS1	47.9
APTR	9.4
MAP4K3-DT	8.4

Table 3.6: Predicted novel peptide size (kDa) for 6 translated lncRNA derived peptides.

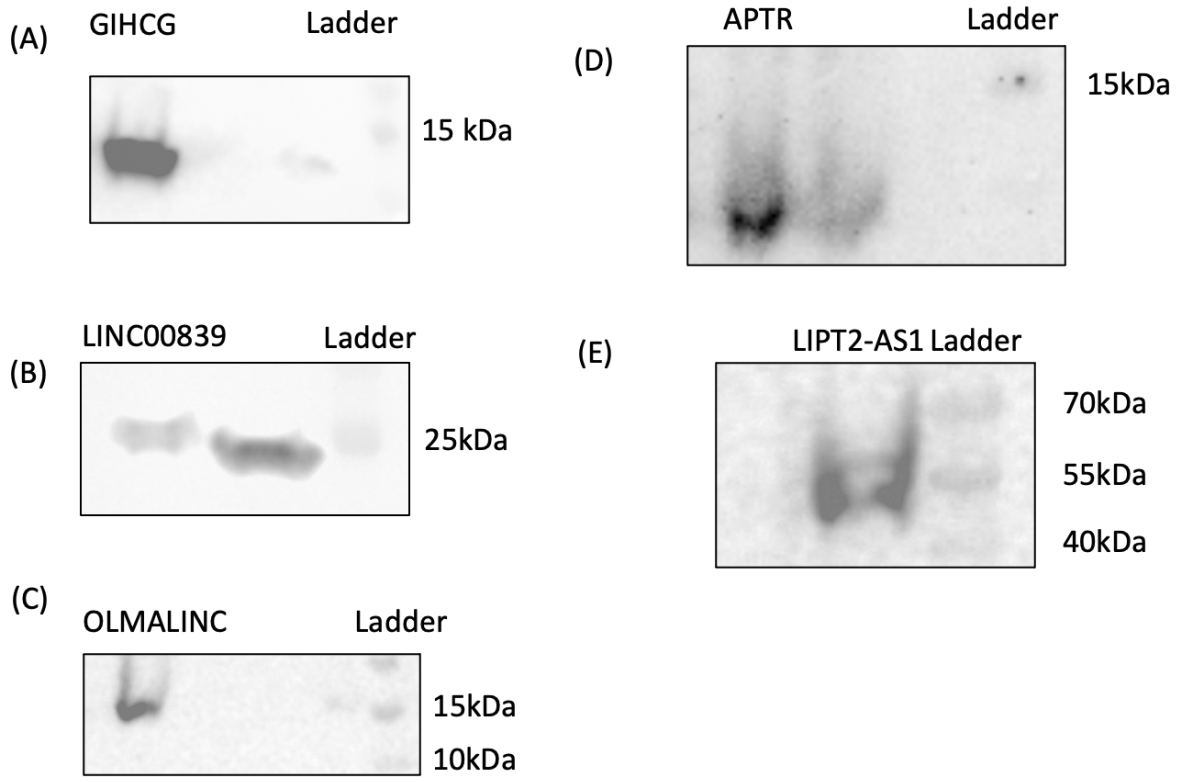


Figure 3.19: Western blot of transiently overexpressed FLAG-tagged novel peptides from SH-SY5Y cell lysate. *~2 million SH-SY5Y cells transfected with 5'-UTR-(s)ORF-FLAG constructs were harvested and lysed in 4x laemmli buffer. Lysates were run on a 4%-20% gradient gel and transferred onto 0.2μm nitrocellulose membrane before immunoblotting (Primary – M2 anti-FLAG, 1:10,000. Secondary, 680nm anti-mouse, 1:10,000) and imaging with a LI-COR. Bands corresponding to novel peptides can be seen at the correct size for GIHCG, LINC00839, OLMALINC, APTR and LIPT2-AS1. No signal was detected for MAP4K3-DT.*

3.14 Novel protein structural prediction with ColabFold

Candidates for further study had already been selected before the release of AlphaFold therefore it was not used in the selection process. However, once available ColabFold (a browser based Alphafold) was used to predict the structure of 6 candidates lncRNA-peptides (Mirdita et al., 2022). Despite the small size of the lncRNA derived peptides only LINC00839 and OLMALINC were predicted to be entirely unstructured (Figure 3.20). GIHCG and MAP4K3-DT are both predicted to form alpha helixes whereas APTR contains a region of β-pleated sheet. For LIPT2-AS1, the HTH DNA binding domain predicted by Phyre2 and SSS is also predicted by ColabFold (Figure 3.21).

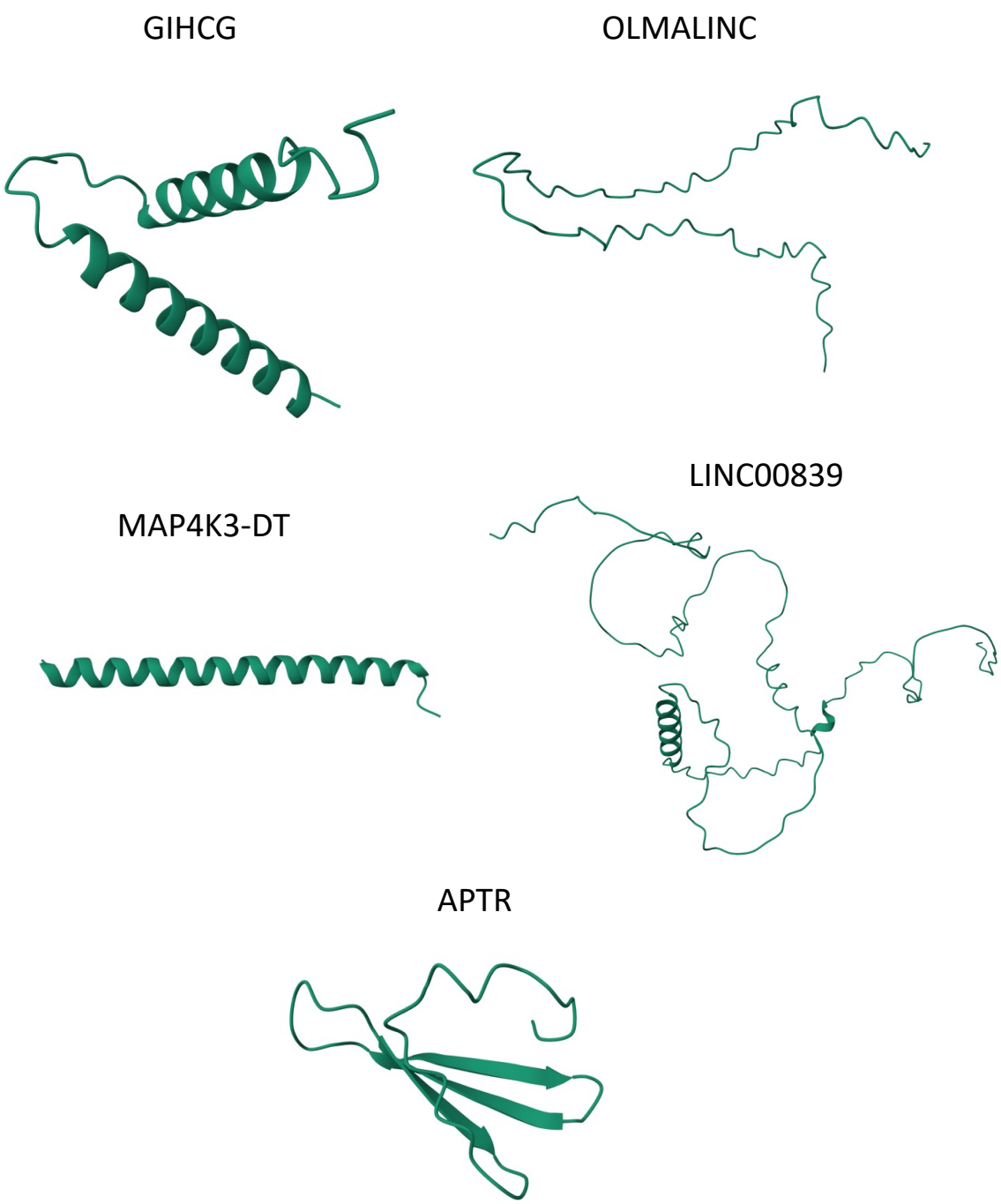


Figure 3.20: Colabfold models for GIHCG, OLMALINC, LINC00839, MAP4K3-DT and APTR. *OLMALINC and LINC00839 are all predicted to be largely disordered. GIHCG, MAP4K3-DT are predicted to fold into alpha-helixes and APTR to contain a region of β-pleated sheet.*

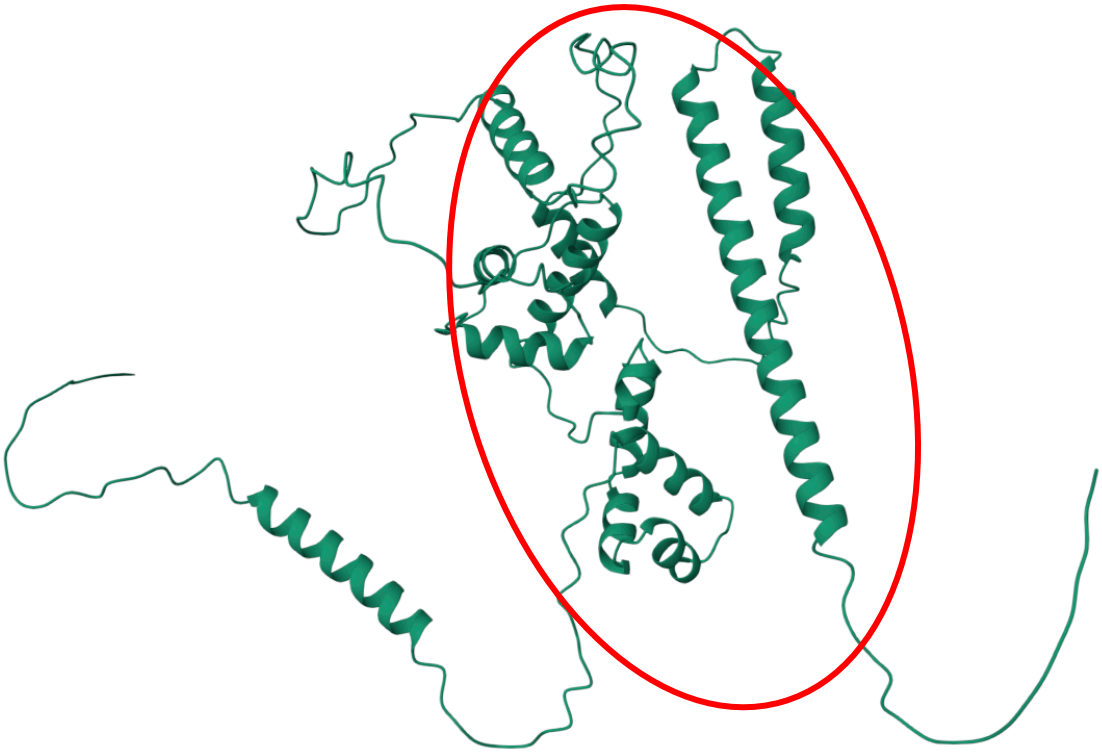


Figure 3.21: ColabFold model of LIPT2-AS1*. Prediction of LIPT2-AS1 protein shows complex structure and the presence of a HTH CenpB-type DNA-binding domain (circled in red).*

3.15 Discussion

3.15.1 Translated lncRNAs show distinct characteristics to total lncRNAs

This chapter investigates the physiological and pathological relevance of translated lncRNAs. Publicly available data revealed that the translated lncRNAs exhibit distinct attributes when compared with all lncRNAs. For example, translated lncRNAs are more likely to show developmentally dynamic expression, indicating they are being actively regulated and therefore potentially functional. The majority (73-90% depending on the species) of protein coding genes show developmentally dynamic expression compared to just 19% of lncRNAs in humans. While still much less likely to show developmentally dynamic expression to canonical protein-coding genes, 39% of translated lncRNAs were dynamically expressed. This suggests an enrichment for physiologically relevant genes in the translated lncRNA population, when compared to lncRNAs in general (Figure 3.2)

Translated lncRNAs are also dysregulated in disease, particularly CNS cancers (Figure 3.3) where 68% of translated lncRNAs are dysregulated in at least one type of CNS cancer. This is perhaps not surprising as they were detected as translated in SH-SY5Ys, an immortalised cell line derived from a neuroblastoma biopsy (Kovalevich & Langford, 2013). RNA-Seq data from the Illumina Body Map shows translated lncRNAs are expressed in the human brain (and other tissues) of healthy humans (Figure 3.1). This along with translated lncRNAs developmentally dynamic expression suggests some translated lncRNAs have physiological roles, which are dysregulated in disease states, in particular cancer.

The final distinction between translated lncRNAs and total lncRNAs is that translated lncRNAs are enriched in the cytoplasm compared to non-translated lncRNAs (Figure 3.3). This feature suggests translated lncRNAs detected by Poly-Ribo-Seq are more likely to efficiently be exported from the nucleus into the cytoplasm where they can be translated into novel proteins. This likely indicates these RNAs are efficiently spliced as inefficient splicing is linked to nuclear localisation (Ntini et al., 2018).

3.15.2 Translated lncRNAs produce stable peptides

Translated lncRNAs ability to produce proteins was investigated with a mini-gene reporter assay. Protein localisation is strongly associated with function; therefore, the first indication

of novel protein function is its specific subcellular localisation. There was no obvious bias to a particular subcellular localisation from the 13 (s)ORFs expressed (Figure 3.5, 3.6), which is consistent with the literature where characterised human microproteins show diverse subcellular localisation (J. Deng et al., 2023). Additionally, novel peptide sizes were validated through western blot for 5 out of the 6 of the candidates tested (Figure 3.19). This data allowed us to confirm the true start site, and therefore size, of the novel proteins.

Structural prediction performed on the novel peptides by Phyre2 and SSS only yielded 2 hits, in LIPT2-AS1 and EMSLR. However, this does not indicate that the remained of novel peptides are not functional. Firstly, this may indicate a limitation in structural modelling tools which are known to work poorly on short peptides (McDonald et al., 2023). Deep-learning based tools, such as AlphaFold, outperform dedicated peptide structure prediction tools, especially when structurally modelling small peptides that contain α-helical, β-hairpin, and disulfide-rich regions. Despite performing better than other methods AlphaFold still showed significant limitations in predicting the structure of small peptides. This is likely due to AlphaFold's training set not including small proteins (McDonald et al., 2023). Therefore, the predicted lack of structure indicative of function for the majority of novel peptides maybe due to limitations of current tools rather than candidates truly lacking structure.

Secondly, ~10% of canonical proteins are completely disordered and ~50% of all proteins contain a disordered region. Unstructured or flexible proteins may adopt multiple confirmations dependant on interactions with other proteins or lipids and therefore pose a challenge for computational predictions (McDonald et al., 2023). Therefore, a lack of structure does not categorise a protein as lacking function. De novo genes tend to encode proteins which are disordered, transmembrane or secreted (Peng & Zhao, 2024). This disorder is primarily due to amino acid composition and on average disorder is reduced in "old" more conserved genes. The intrinsic disorder of novel protein has been argued to support the preadaptation hypothesis where de novo proteins are initially are designed to have minimal risk of being harmful and can then develop more subtle characteristics (Wilson et al., 2017).

LIPT2-AS1 stands out from the other candidates due to the size of its ORF (392 aa), which is substantially above the 100 or 150 aa cut-offs often used for defining "micropeptides" (Guillén et al., 2013; Makarewich & Olson, 2017). Additionally, LIPT2-AS1 is predicted to

contain a HTH DNA binding domain, suggesting it can directly interact with DNA. The protein LIPT2-AS1 shows structural homology to CENP-B, a DNA binding protein that binds a 17-bp sequence in the centromeric alpha satellite DNA. CENP-B has established roles in maintaining centromere stability and modulating chromatin structure (Nagpal et al., 2023). It is unclear if LIPT2-AS1 protein will also bind the same DNA sequence but does demonstrate clear nuclear localisation forming discreet puncta suggesting specific interactions within the nucleus (Figure 3.13).

3.16 Conclusion

In this chapter key differences were identified between translated and untranslated lncRNAs, including translated lncRNAs being enriched for signs of physiological and pathological relevance. Of the 13 lncRNA-ORFs tested 12 showed evidence of translation by FLAG tag reporter. The size and subcellular localisation of 6 key candidates (Table 3.4) was characterised in more detail. In the following chapter these 6 candidates will be investigated for potential functional roles in neuronal differentiation.

Chapter 4: Investigating the functional role of translated lncRNAs in neuronal differentiation

4.1 Introduction

Conservative estimates put the number of noncanonical (s)ORFs to be ~7000, although this estimate varies wildly depending on the source (Prensner et al., 2023). However, only a minor fraction of (s)ORFs identified through Ribo-Seq and MS experiments have had their peptides functionally characterised. It remains a challenge to identify the function of actively translated noncanonical ORFs, partly due to the large numbers identified, but also as these translation products are novel there is minimal information that could suggest a potential function or mechanism. Further, structural predictions tools, such as Alphafold, are trained on canonical proteins and therefore are likely to be poor in their structural predictions of small proteins. Similarly, micropeptide typically lack predictable domains, making the functional predictions impossible (McDonald et al., 2023; Peng & Zhao, 2024).

Despite these challenges, lncRNA derived peptides have been identified as critical factors in physiological development and multiple disease states. The novel micropeptide NEMEP is expressed in response to Nodal signalling in mesendoderm differentiation from mouse embryonic stem cells (mESCs). NEMEP, is a transmembrane protein, has been shown through Co-IP in HEK293 cells to interact with GLUT1 and GLUT3. This association induces glucose uptake in mESC and CRISPR frameshift mutations of NEMEP impairs glucose uptake and consequently inhibits mesendoderm differentiation (Fu et al., 2022). Additionally, a non-annotated ORF within the hominoid specific gene ENSG00000205704, which was previously annotated as a lncRNA named LINC00634, produces a 107 aa protein. This protein is involved in neuronal development with its expression required for cortical organoid development from hESCs. Knockout of ENSG00000205704 resulted in significantly smaller organoids than wild-type controls due to accelerated neuronal maturation (An et al., 2023). Beyond development, lncRNA derived proteins have been identified as having important roles in neurological diseases. The microprotein SHMOOSE is encoded from a mitochondrial gene and is localised in the mitochondria through an interaction with inner mitochondrial membrane mitofilin as shown by Co-IP experiments using SH-SY5Y cell lysate. SHMOOSE protein is expressed throughout the human brain and can be detected in patient CSF. SNPs within SHMOOSE that result in an amino acid substitution are associated with Alzheimer's disease and pathological brain structure suggesting a potential role for SHMOOSE as a biomarker (Miller et al., 2023).

From the 45 (s)ORFs identified as translated but Poly-Ribo-Seq (Douka et al., 2021), 6 candidate ORFs stood out due to a combination of the subcellular localisation of their peptide, showing tightly regulated expression through human brain development or literature indicating roles within neuronal disease. To elucidate if translated lncRNAs are required for neuronal differentiation a small-scale siRNA screen was performed on the strongest 6 translated lncRNAs candidates (Table 3.4).

Finally, cortical organoids were used to investigate the expression and translation of candidates beyond SH-SY5Y cells. SH-SY5Y cells are a common model for neuronal differentiation but are limited their ability to model human brain development due to 2D structure, time-scale limitations, and inability to differentiate into multiple cell types. Therefore, human cortical organoids, with their 3D structure, multiple cell types and transcriptome profiles of hCSs are consistent with the developing human brain (Paşca et al., 2015), are a promising model to investigate the role of translated lncRNAs in neuronal development. Therefore, we wanted to confirm the expression and translation of our candidate lncRNAs in this model.

4.2 Validation of SH-SY5Y as a model of neuronal differentiation

SH-SY5Ys are a widely used model of neuronal differentiation and were previously used for the Poly-Ribo-Seq to identify novel translated lncRNAs (Douka et al., 2021). Therefore, SH-SY5Y were used for siRNA screening of translated lncRNA. SH-SY5Y differentiation has been extensively characterised in literature and includes the hallmarks of changes in gene expression, neurite outgrowth and slowed proliferation (Douka et al., 2021; Forster et al., 2016; Kovalevich & Langford, 2013; Strother et al., 2021).

SH-SY5Ys were induced to differentiate with 30μM RA for 3 days, as has previously been described (Douka et al., 2021; Forster et al., 2016; Korecka et al., 2013). Differentiated SH-SY5Y cells showed significant upregulation of neuronal marker genes NTN4 and RET among others, as is consistent with the literature (Douka et al., 2021; Forster et al., 2016). NTN4 encodes netrin-4, a secreted protein that provide tropic signals for axonal growth and guidance (Dong et al., 2023; Hayano et al., 2014). RT-qPCR was used to quantify the relative expression of NTN4 in immature and RA treated SH-SY5Y cells and a significant 19-fold increase with differentiation (Figure 4.1). RET encodes a TM receptor protein for glial cell line-derived neurotrophic factor (GDNF) family ligands and is essential for CNS development (Mahato & Sidorova, 2020)and is upregulated ~6 fold at the RNA level with RA treatment (Figure 4.1). Due to NTN4 and RETs essential role in neuronal development and robust changes in RNA levels following RA treatment they were selected as markers of differentiation. SOX2 was selected as a marker of pluripotency as it shows significant downregulation upon differentiation in SH-SY5Y cells and is a key regulation of NPC fate (Hutton & Pevny, 2011; Mercurio et al., 2019; Ng et al., 2013). In SH-SY5Y cells a 2-fold decrease in SOX2 RNA abundance is seen between immature and differentiated SH-SY5Y cells, indicating RA treatment drives SH-SY5Ys to adopt a more mature neuronal phenotype (Figure 4.1).

Neurite outgrowth is regularly used to quantify SH-SY5Y differentiation (Dwane et al., 2013). Neurites are the pre-cursors to axons and dendrites but are not functional in terms of action potential transmission and synapse formation (Dwane et al., 2013). RA treatment of SH-SY5Ys rapidly induces extension of neurites and can be easily quantified to show neurite

length is significantly greater in differentiated SH-SY5Y cells compared to immature cells (Figure 4.2).

However, SH-SY5Ys are not terminally differentiated after 3 days of RA treatment, with the process more accurately being seen as the first stage of change from NPC to mature neuron. Further differentiation is possible with some differentiation protocols continuing beyond 3 days, resulting in a more mature neuron expressing catecholaminergic markers and capable of forming synapses and show electrical activity (Strother et al., 2021).

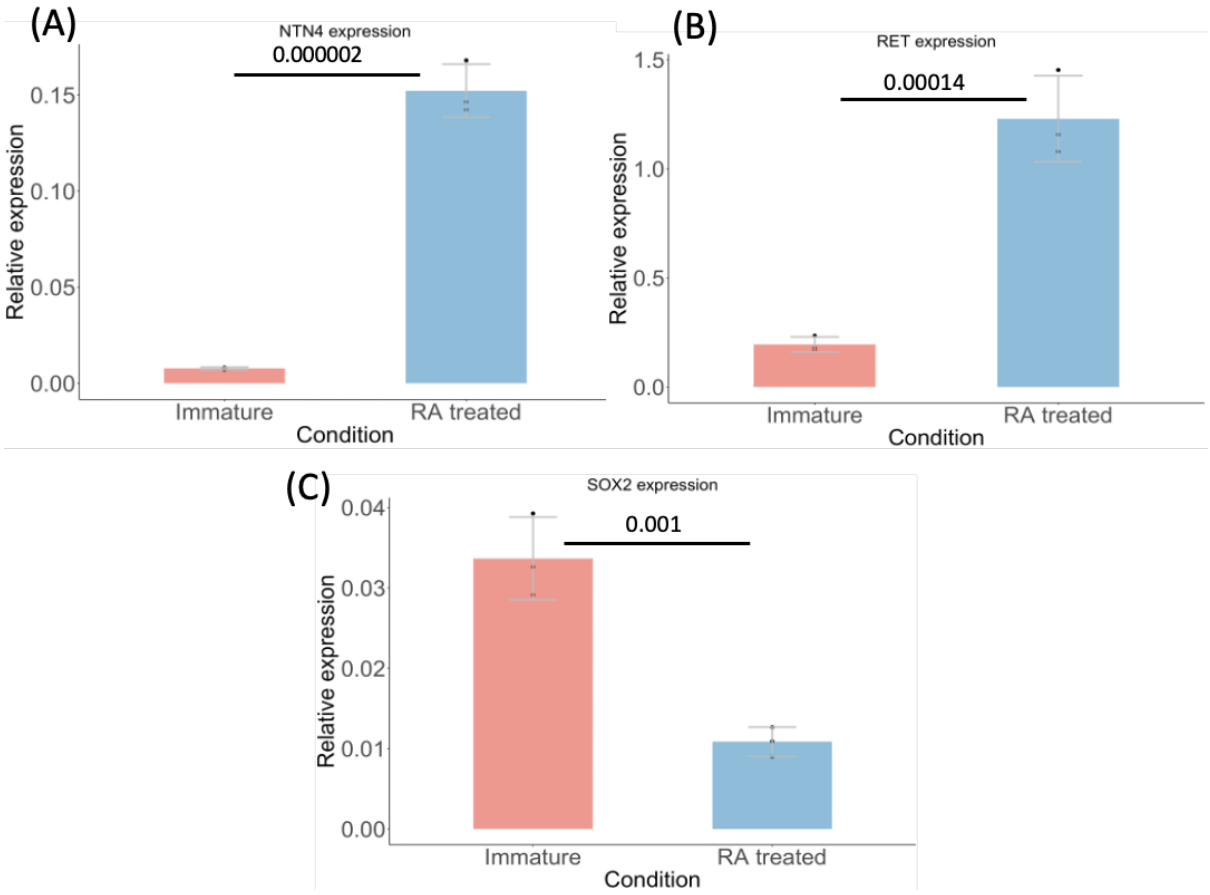


Figure 4.1: Successful differentiation of SH-SY5Y cells can be measured with mRNA levels of marker genes. *RNA abundance of (A) NTN4 and (B) RET significantly increases with 3 days of RA treatment whereas RNA abundance of the pluripotency marker (C) SOX2 show significant downregulation indicating that the SH-SY5Ys have been differentiated into a more neuronal like phenotype. Relative quantification was performed using housekeeping genes NDUFV2 and HAUS8 and calculated with the ΔCq method. Two-way paired t-test, N=3, *p<0.05. Individual biological repeats denoted with black dots on bar charts.*

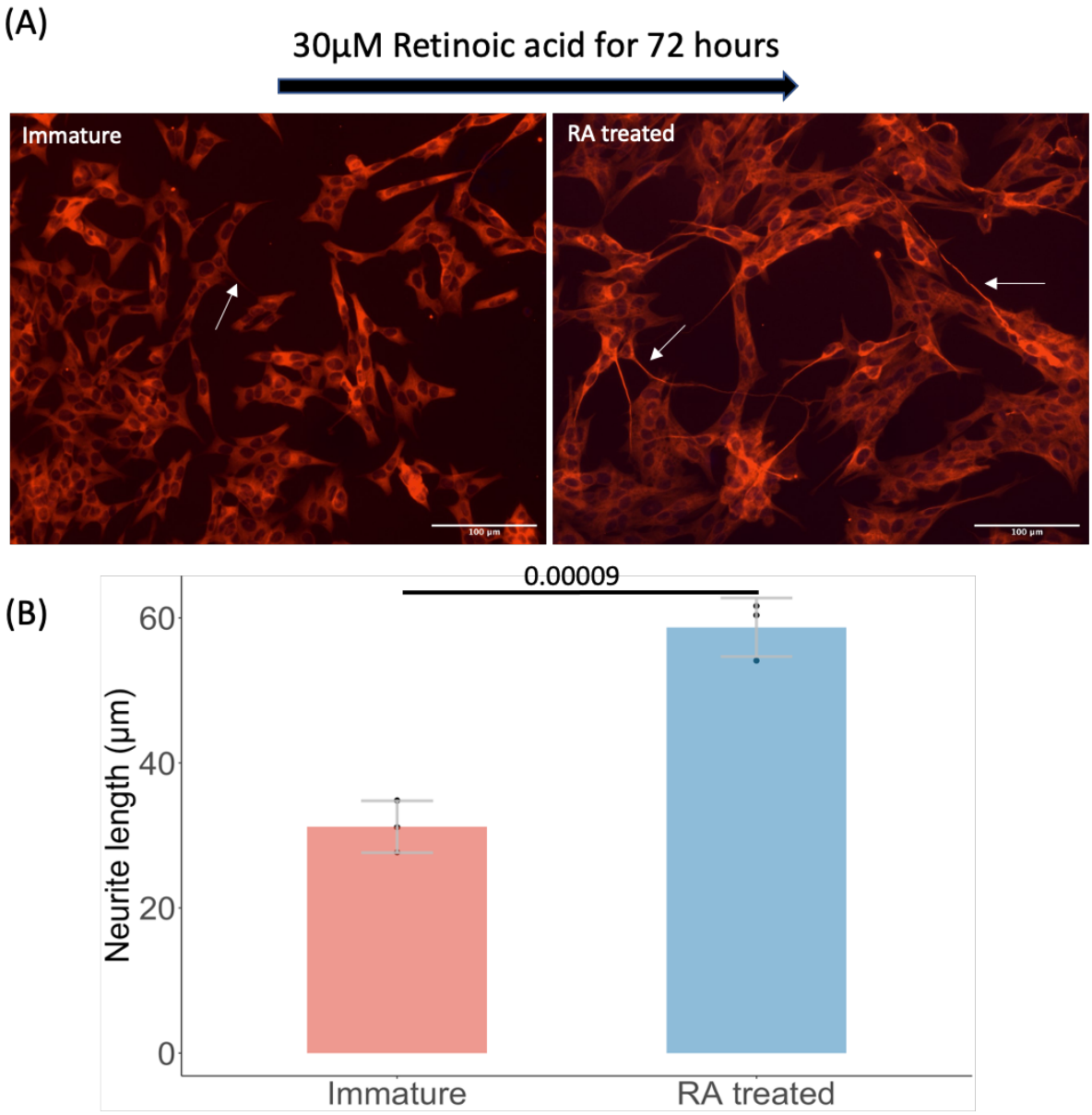


Figure 4.2: Differentiated SH-SY5Y have significantly longer neurites that immature cells. *(A) Representative images of SH-SY5Ys treated with either DMSO (vehicle control) or 30µM RA for 72 hours. Scale bar = 100µm. Red = Tuj1. Blue = DAPI. (B) Quantification of neurite length was performed, and RA treated SH-SY5Y cells have significantly longer neurites than immature SH-SY5Ys. N=3, >100 neurites measured in FIJI per condition, Two-way paired t-test. P value reported on graph.*

4.3 Translated lncRNAs show consistent RNA abundance throughout SH-SY5Y differentiation

Large changes in RNA abundance are observed with SH-SY5Y differentiation with RNA-seq showing 1628 differentially expressed protein coding genes in immature vs differentiated SH-SY5Ys (Douka et al., 2021). To assess any changes in translated lncRNA abundance with differentiation, RT-qPCR was used to assess the relative levels of translated lncRNAs in immature and differentiated SH-SY5Y cells. Of the 6 lncRNAs investigated 5 showed no significant difference in RNA levels between the two conditions (Figure 4.3). This was surprising as all candidates were only detected as translated in one condition by Poly-Ribo-Seq (Douka et al., 2021). This could be due to the high bioinformatic threshold required to be called as translated by RiboTaper or translation control where lncRNAs are transcribed in both immature and differentiated SH-SY5Ys but only translated, and therefore detected by Poly-Ribo-Seq, in one condition. MAP4K3-DT was the only lncRNA tested to significant changes in RNA level with a ~2-fold change from immature to differentiated SH-SY5Y cells (Figure 4.3).

4.4 Translated lncRNAs are required for normal neuronal differentiation

To assess functional importance of translated lncRNAs in SH-SY5Y differentiation, siRNA knockdown was performed. SH-SY5Y cells were treated with either siRNAs targeting a translated lncRNA or a pool of scrambled siRNA complexes. siRNA was selected as the method of knockdown for this small-scale screen due to their ease of use, minimal preparation time and well-established controls. Where available pools of 4 siRNAs specific to a target were used (Lincode-SMARTpool, DharmaconTM) and compared to cells treated with a pool of 4 scrambled siRNAs. Where a pool was not available, 2 siRNAs targeting different regions of the translated lncRNA transcript were selected and combined (Thermo). Cells were allowed to recover for 24 hours following siRNA treatment cells then were either induced to differentiate for 3 days with RA or treated with a vehicle control (remaining immature). After 3 days, cells were harvested to assess if a) lncRNA knockdown had been effective and b) there was any effect on neuronal differentiation. This was primarily assessed by RT-qPCR to determine relative levels of neuronal marker genes, NTN4 and RET,

and the pluripotency marker SOX2. Additionally, cells were fixed, stained with the neuronal specific cytoskeletal protein Tuj1, and imaged to measure neurite length (Figure 4.4).

4.4.1 siRNA effectively knocks down translated lncRNAs of interest

RNA was extracted from knockdown and controls cells 4 days after siRNA transfection and knockdown levels assessed through the ΔΔCq method (Haimes & Kelley, 2015) with NDUFV2 and HAUS8 used as reference genes. All lncRNAs targeted were effectively knocked down with reductions in expression of between 57%-91% (Figure 4.5).

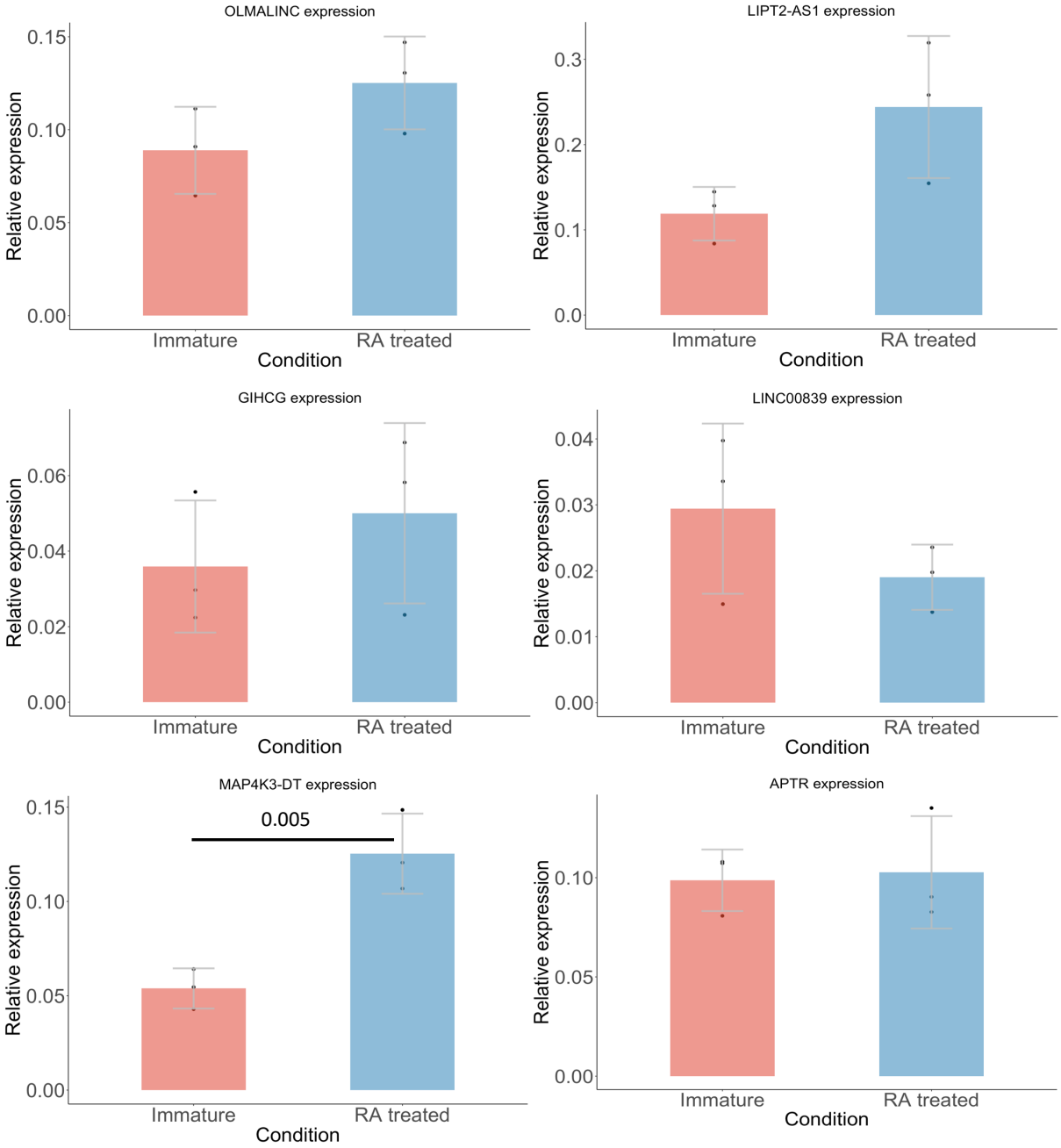


Figure 4.3: Total RNA levels of translated lncRNAs OLMALINC, LIPT2-AS1, GIHCG and LINC00839 do not significantly change with SH-SY5Y differentiation. *SH-SY5Y cells were differentiated with 30μM RA over 72 hours (or treated with DMSO as a vehicle control). Relative RNA abundance was then measured with RT-qPCR in immature and differentiated SH-SY5Y cells. MAP4K3-DT is the only translated lncRNA to show a significant difference in RNA abundance between immature and differentiated conditions. Relative quantification was performed housekeeping genes NDUFV2 and HAUS8 and calculated with the ΔCq method. Two-way paired t-test, N=3. Individual biological repeats denoted with dots on bar charts. P value reported on graphs.*

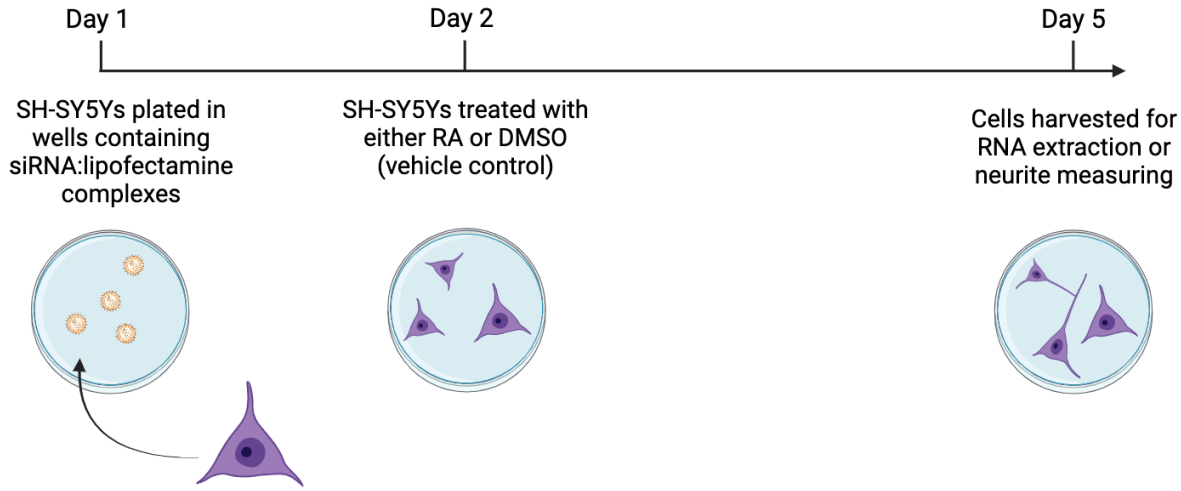


Figure 4.4: Schematic showing siRNA knockdown timeline*. SH-SY5Y cells are reverse transfected with siRNA, allowed to recover for 24 hours before being induced to differentiate with RA or left as immature and treated with DMSO as a vehicle control. After 3 days cells are harvested for RNA extractions to perform RT-qPCR or fixed and stained for neurite measuring.*

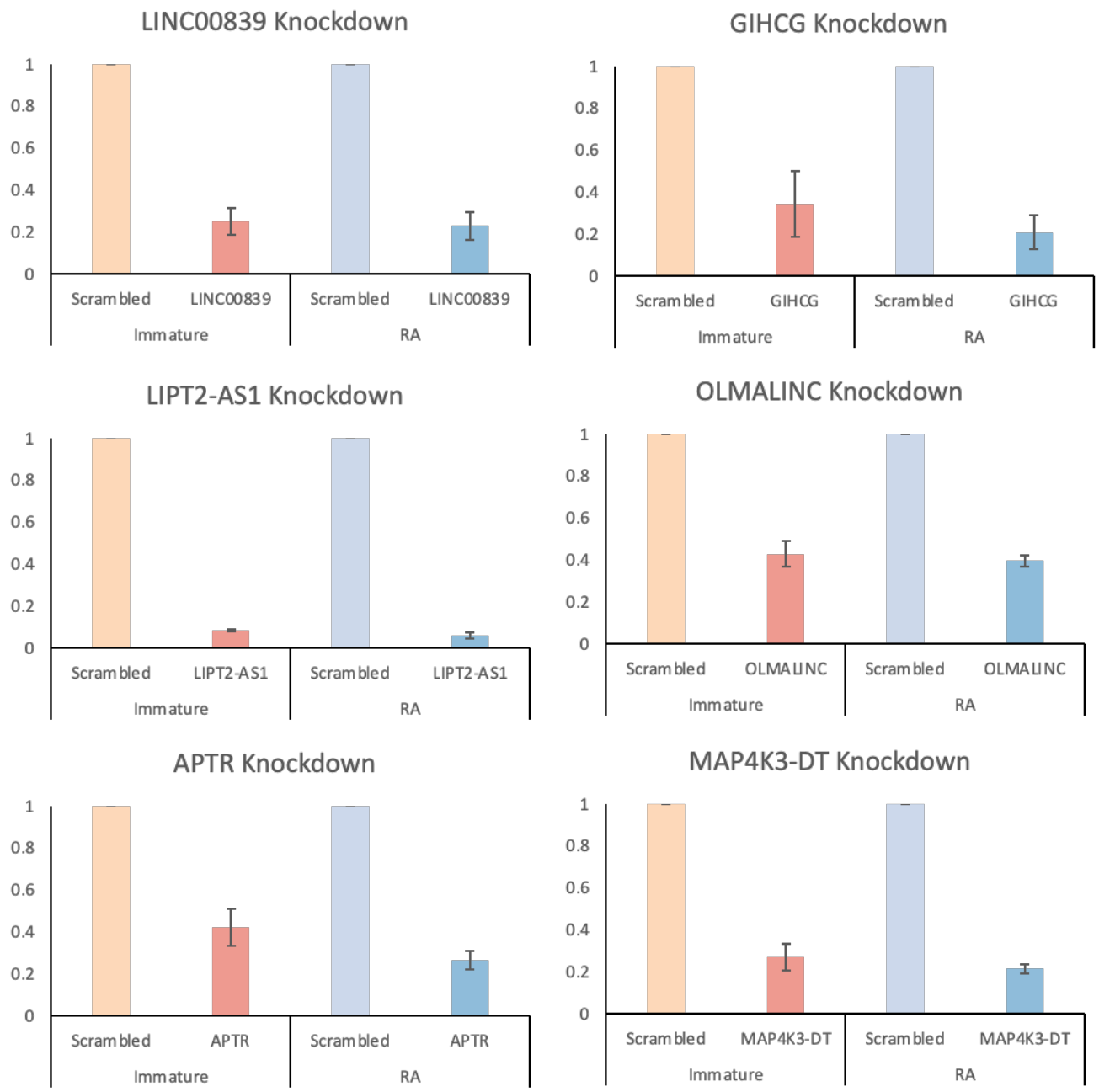


Figure 4.5: Translated lncRNAs are successfully silenced by siRNA throughout SH-SY5Y differentiation. *(A) RNA abundance of LINC00839 is reduced by 75% and 77% in immature and differentiated SH-SY5Ys, respectively. (B) RNA abundance of GIHCG is reduced by 66% and 80% in immature and differentiated SH-SY5Ys, respectively. (C) RNA abundance of LIPT2-AS1 is reduced by 89% and 91% in immature and differentiated SH-SY5Ys, respectively. (D) RNA abundance of OLMALINC is reduced by 57% and 60% in immature and differentiated SH-SY5Ys, respectively. (E) RNA abundance of APTR is reduced by 58% and 74% in immature and differentiated SH-SY5Ys, respectively. (F) RNA abundance of OLMALINC is reduced by 73% and 79% in immature and differentiated SH-SY5Ys, respectively. Relative quantification was performed using housekeeping genes NDUFV2 and HAUS8 and calculated with the ΔΔCq method, N=3.*

4.4.2 LINC00839 and GIHCG are required for neuronal marker gene upregulation

Although LINC00839 has been extensively characterised as an RNA in cancer (Table 4.1), its role in neuronal differentiation has not been characterised. siRNA knockdown of LINC00839 in differentiated SH-SY5Y cells significantly reduced the RNA levels of mature marker genes NTN4 and RET compared to scrambled controls, suggesting that the expression of LINC00839 is required for normal neuronal differentiation (Figure 4.6). LINC00839 knockdown also dysregulated SOX2 levels in immature, but not differentiated, SH-SY5Ys. SOX2 levels were significantly reduced in knockdown cells, indicating a disruption in the stemness of immature SH-SY5Y cells upon loss of LINC00839. This is supported with a mild upregulation of RET in immature SH-SY5Y cells suggesting in immature SH-SY5Ys loss of LINC00839 induces SH-SY5Y cells to differentiate (Figure 4.6)

Similarly, to data GIHCG has not been found to have a role in the differentiation of any cell type but has been proposed to function in multiple cancers (Table 4.1). siRNA knockdown of GIHCG resulted in significantly reduced levels of NTN4 in differentiated SH-SY5Y cells compared to scrambled controls. However, GIHCG knockdown had no effect on RET or SOX2 abundance (Figure 4.7).

Beyond changes in RNA abundance, knockdown of LINC00839 attenuates SH-SY5Y neurite extension in response to RA with the neurites of LINC00839 knockdown SH-SY5Ys being significantly shorter than scrambled controls. Neurite length in GIHCG silenced cells were also an average of 26μm shorter that controls but did not meet the threshold for significance of $p < 0.05$ (Figure 4.8).

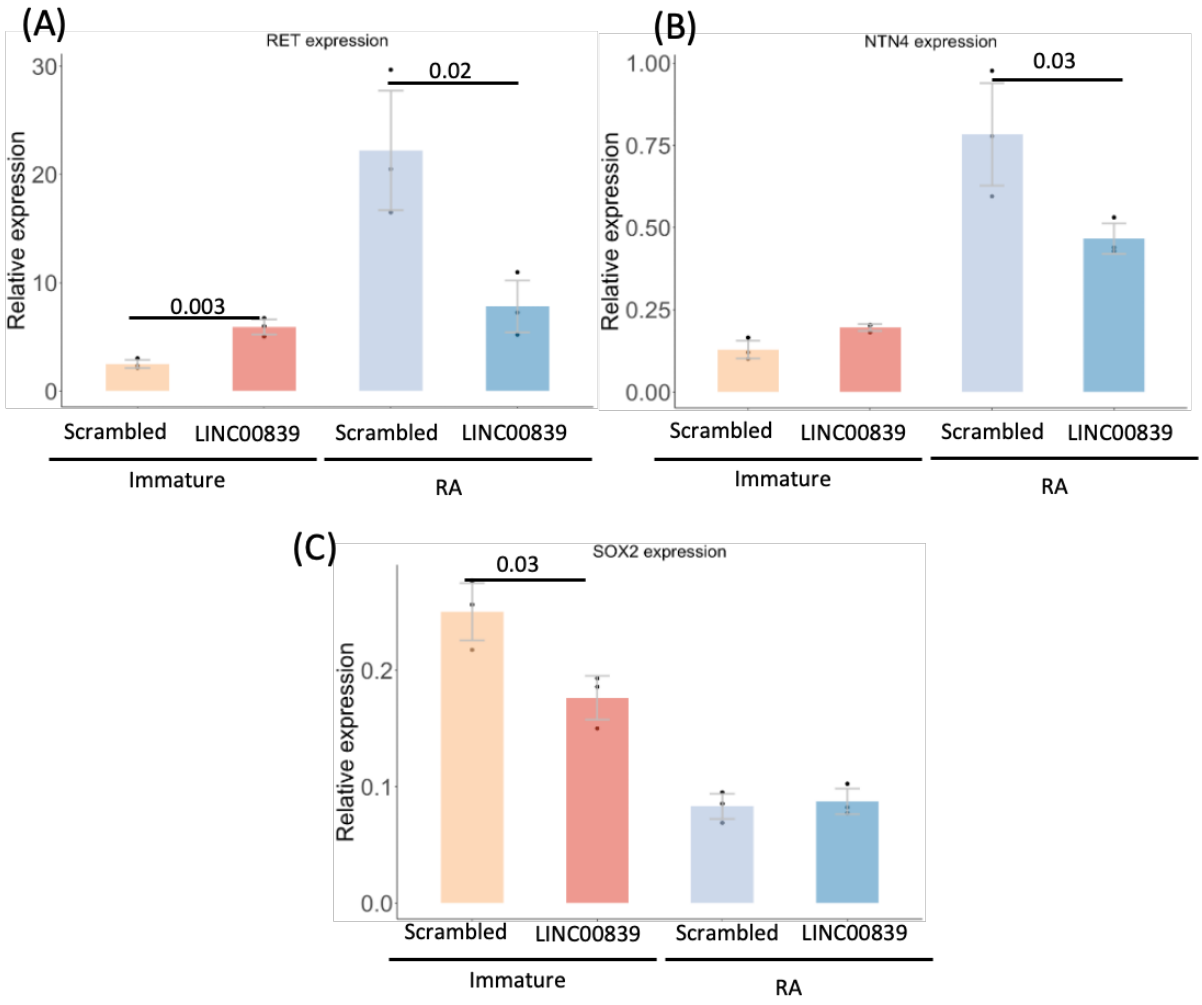


Figure 4.6: LINC00839 knockdown attenuates neuronal marker gene upregulation in response to RA and dysregulates SOX2 levels. *Relative quantification of mature neuronal markers RET and NTN4 and pluripotency marker SOX2 in SH-SY5Y treated with either siRNAs targeting LINC00839 or a pool of scrambled siRNAs. LINC00839 knockdown resulted in a significant reduction of (A) RET and (B) NTN4 in differentiated SH-SY5Y cells. (C) SOX2 levels were significantly reduced in immature SH-SY5Y cells with LINC00839 knocked down. Relative quantification was performed housekeeping genes NDUFV2 and HAUS8 and calculated with the ΔCq method. Two-way paired t-test, N=3. Individual biological repeats denoted with dots on bar charts. P value reported on graphs.*

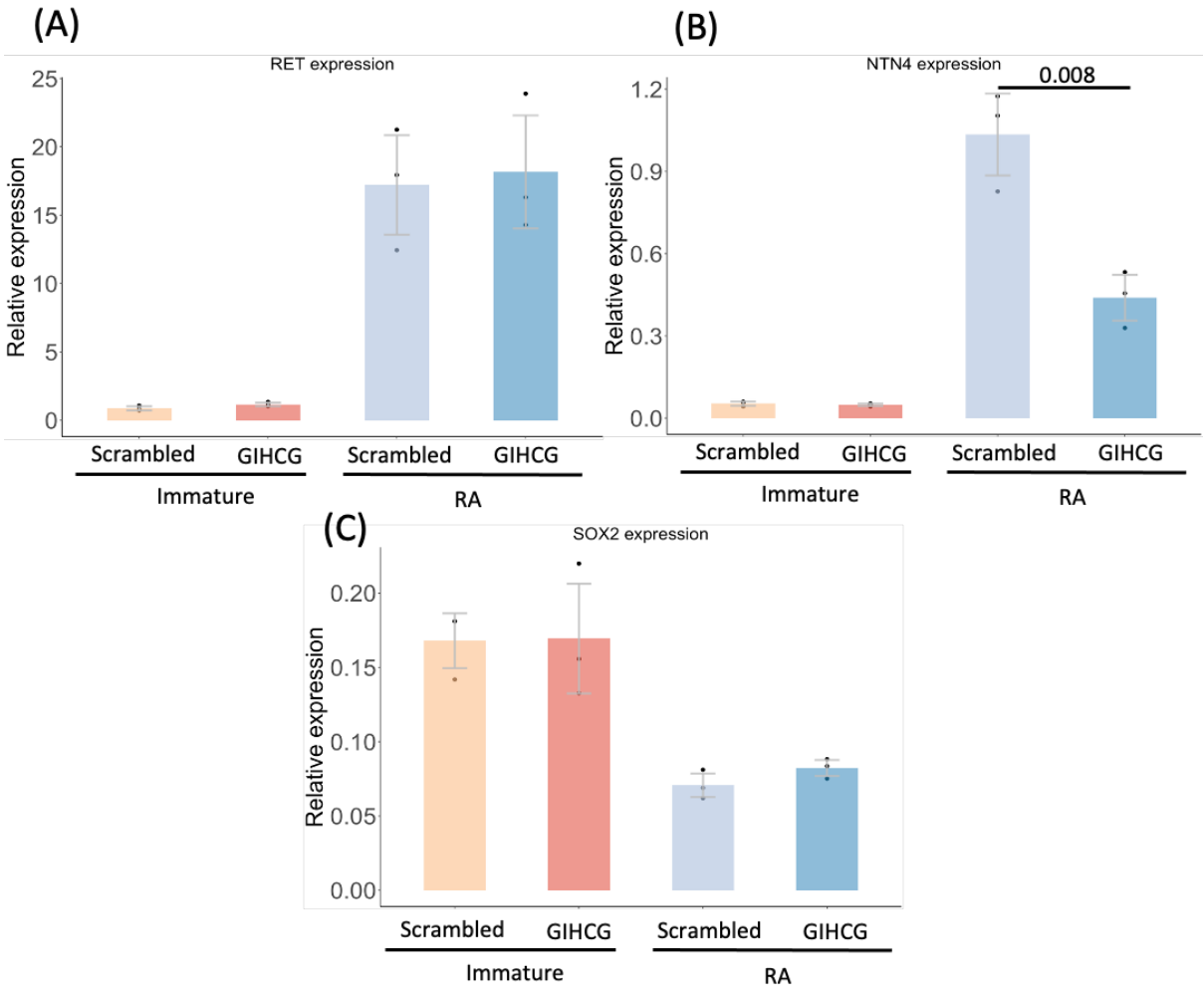


Figure 4.7: GIHCG knockdown attenuates NTN4 upregulation in response to RA but has no impact on RET or SOX2. *Relative quantification of mature neuronal markers RET and NTN4 and pluripotency marker SOX2 in SH-SY5Y treated with either siRNAs targeting GIHCG or a pool of scrambled siRNAs. GIHCG knockdown resulted in a significant reduction of (B) NTN4, but not (A) RET or (C) SOX2, in differentiated SH-SY5Y cells. GIHCG knockdown resulted in no significant changes in immature SH-SY5Y cells. Relative quantification was performed housekeeping genes NDUFV2 and HAUS8 and calculated with the ΔCq method. Two-way paired t-test, N=3. Individual biological repeats denoted with dots on bar charts. P value reported on graphs.*

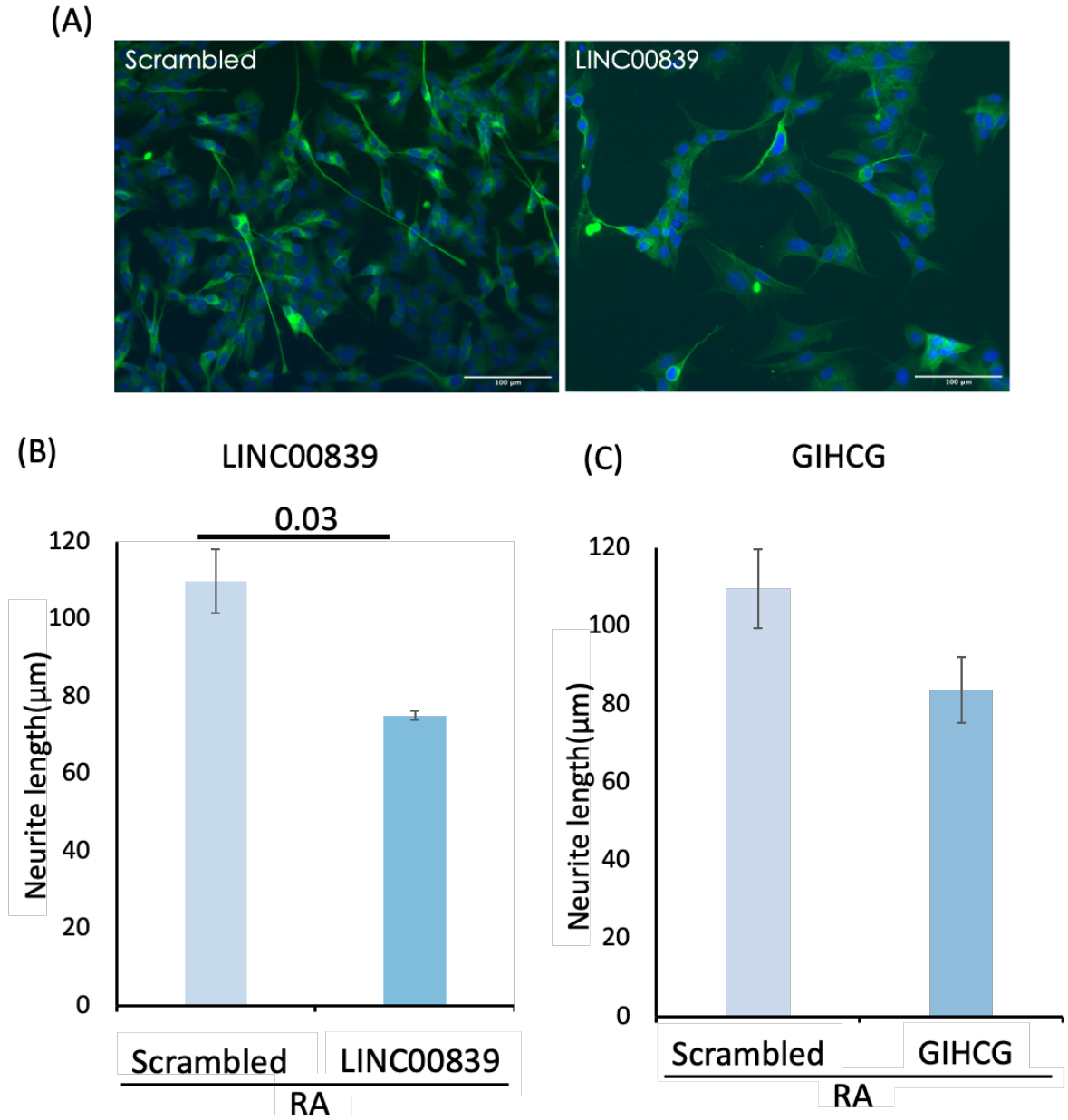


Figure 4.8: siRNA knockdown of LINC00839, but not GIHCG, results in significantly shorter neurites in differentiated SH-SY5Y cells vs control cells. *(A) Representative images of SH-SY5Ys treated with either a scrambled siRNA pool or a siRNAs against LINC00839. Scale bar = 100µm. Green = Tuj1. Blue = DAPI. (B) Quantification of neurite length in differentiated SH-SY5Y cells treated with either scrambled siRNA or siRNA targeting LINC00839. Neurites of LINC00839 knockdown cells are significantly shorter than scrambled controls. (C) Quantification of neurite length in differentiated SH-SY5Y cells treated with either scrambled siRNA or siRNA targeting GIHCG. Neurites of GIHCG knockdown cells are shorter than scrambled controls but do not reach the threshold for significance. N=3, >100 neurites measured per condition, Two-way paired t-test.*

4.4.3 OLMALINC and LIPT2-AS1 knockdown results in significant upregulation of neuronal marker genes

OLMALINC RNA is highly expressed in the human brain, specifically in cortical white matter and has been show as essential to oligodendrocyte and neuronal cell maturation (Mills et al., 2015). Specifically, when OLMALINC was silenced by siRNA knockdown in the SH-SY5Y precursor cell-line SK-N-SH, a significant upregulation of SOX2 and SOX4 was observed (Mills et al., 2015). This was not observed with knockdown in SH-SY5Ys, with SOX2 levels showing no difference in RNA level between knockdown and control. However, a robust upregulation of both RET and NTN4 compared to scrambled control was seen in both immature and differentiated SH-SY5Y cells (Figure 4.9). This suggests that in SH-SY5Y cells, but potentially not SK-N-SH cells, loss of OLMALINC promotes neuronal differentiation.

LIPT2-AS1 has not been extensively characterised but shows high expression in the human brain and FLAG tagging experiments indicate it is translated into a nuclear protein (Table 4.1). siRNA knockdown of LIPT2-AS1 results in significantly higher RNA levels of neuronal markers NTN4 and RET in differentiated SH-SY5Y compared to the scrambled control. This suggests LIPT2-AS1 plays a regulatory role in SH-SY5Y differentiation. LIPT2-AS1 knockdown has no effect on the expression of pluripotency marker SOX2 suggesting LIPT2-AS1 is not required to maintain stemness of NPCs (Figure 4.10).

Silencing of LIPT2-AS1 results in significantly reduced neurite length of differentiated SH-SY5Y cells compared to scrambled controls, indicating the loss of LIPT2-AS1 is disrupting SH-SY5Ys ability to differentiate (Figure 4.11). OLMALINC knockdown did not have a consistent effect on neurite length, as shown by large error bars. This could be reflecting the inconsistent knockdown achieved by the OLMALINC siRNA pool, leading to high variability in biological repeats (Figure 4.11).

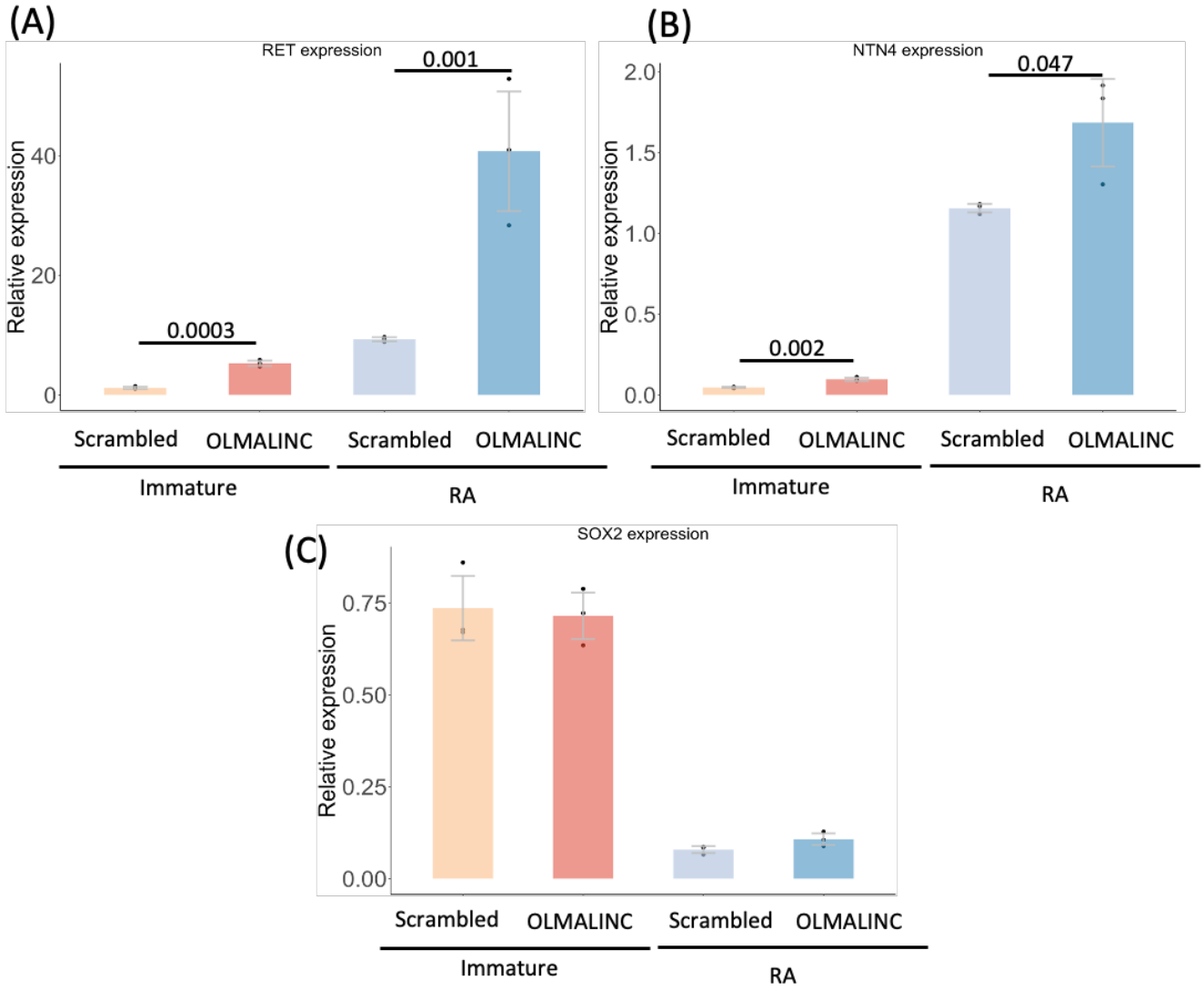


Figure 4.9: OLMALINC knockdown results in NTN4 and RET upregulation in response to RA but has no impact on SOX2. *Relative quantification of mature neuronal markers RET and NTN4 and pluripotency marker SOX2 in SH-SY5Y treated with either siRNAs targeting OLMALINC or a pool of scrambled siRNAs. OLMALINC knockdown resulted in a significant increase of (A) RET and (B) NTN4 compared to scrambled treated samples. OLMALINC knockdown resulted in no significant changes in immature SH-SY5Y cells and had no effect on SOX2 levels in either condition (C). Relative quantification was performed housekeeping genes NDUFV2 and HAUS8 and calculated with the ΔCq method. Two-way paired t-test, N=3. Individual biological repeats denoted with dots on bar charts. P value reported on graphs.*

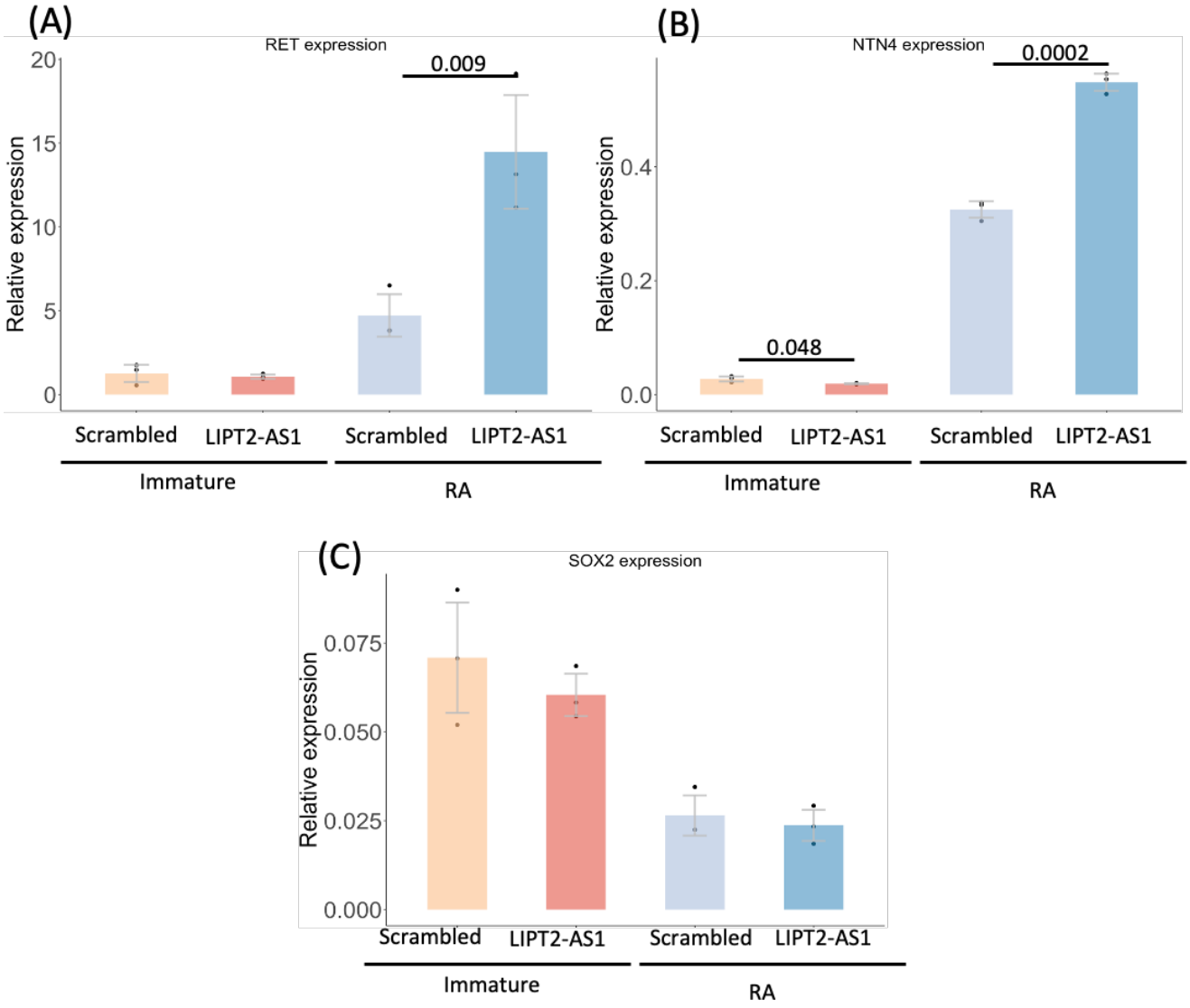


Figure 4.10: LIPT2-AS1 knockdown results in NTN4 and RET upregulation in response to RA but has no impact on SOX2. *Relative quantification of mature neuronal markers RET and NTN4 and pluripotency marker SOX2 in SH-SY5Y treated with either siRNAs targeting LIPT2-AS1 or a pool of scrambled siRNAs. LIPT2-AS1 knockdown resulted in a significant increase of (A) RET and (B) NTN4 compared to scrambled treated samples. LIPT2-AS1 knockdown resulted in no significant changes in immature SH-SY5Y cells. Relative quantification was performed housekeeping genes NDUFV2 and HAUS8 and calculated with the ΔCq method. Two-way paired t-test, N=3. Individual biological repeats denoted with dots on bar charts. P value reported on graphs.*

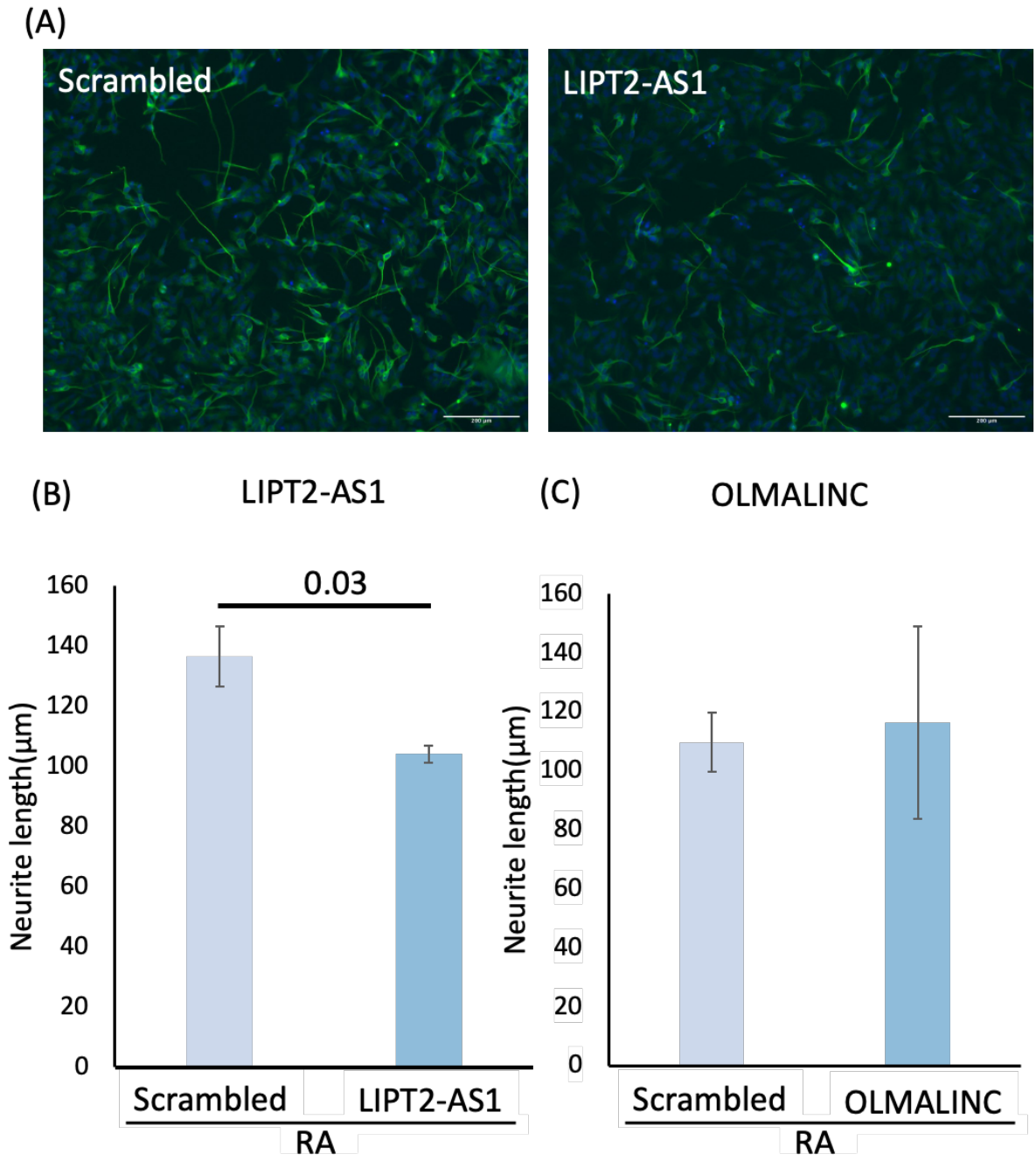


Figure 4.11: siRNA knockdown of LIPT2-AS1, but not OLMALINC, results in significantly shorter neurites in differentiated SH-SY5Y cells vs control cells. *(A) Representative images of SH-SY5Ys treated with either a scrambled siRNA pool or a siRNAs against LIPT2-AS1. Scale bar = 200µm. Green = Tuj1. Blue = DAPI. (B) quantification of neurite length in differentiated SH-SY5Y cells treated with either scrambled siRNA or siRNA targeting LIPT2-AS1. Neurites of LIPT2-AS1 knockdown cells are significantly shorter than scrambled controls. (C) quantification of neurite length in differentiated SH-SY5Y cells treated with either scrambled siRNA or siRNA targeting OLMALINC. N=3, >100 neurites measured per condition, Two-way paired t-test. P value reported on graphs. Error bars show SEM.*

4.4.4 APTR and MAP4K3-DT siRNA knockdown has no effect on neuronal differentiation

Neither APTR nor MAP4K3-DT has previously been characterised in the context of cell differentiation. siRNA knockdown of both translated lncRNAs had no significant impact on the marker genes of NTN4, RET and SOX2 (Figure 4.12 and 4.13). This indicates that APTR and MAP4K3-DT are not involved in either NPC maintenance or neuronal differentiation. Due to the lack of RNA phenotype, effects on neurite extension were not assessed.

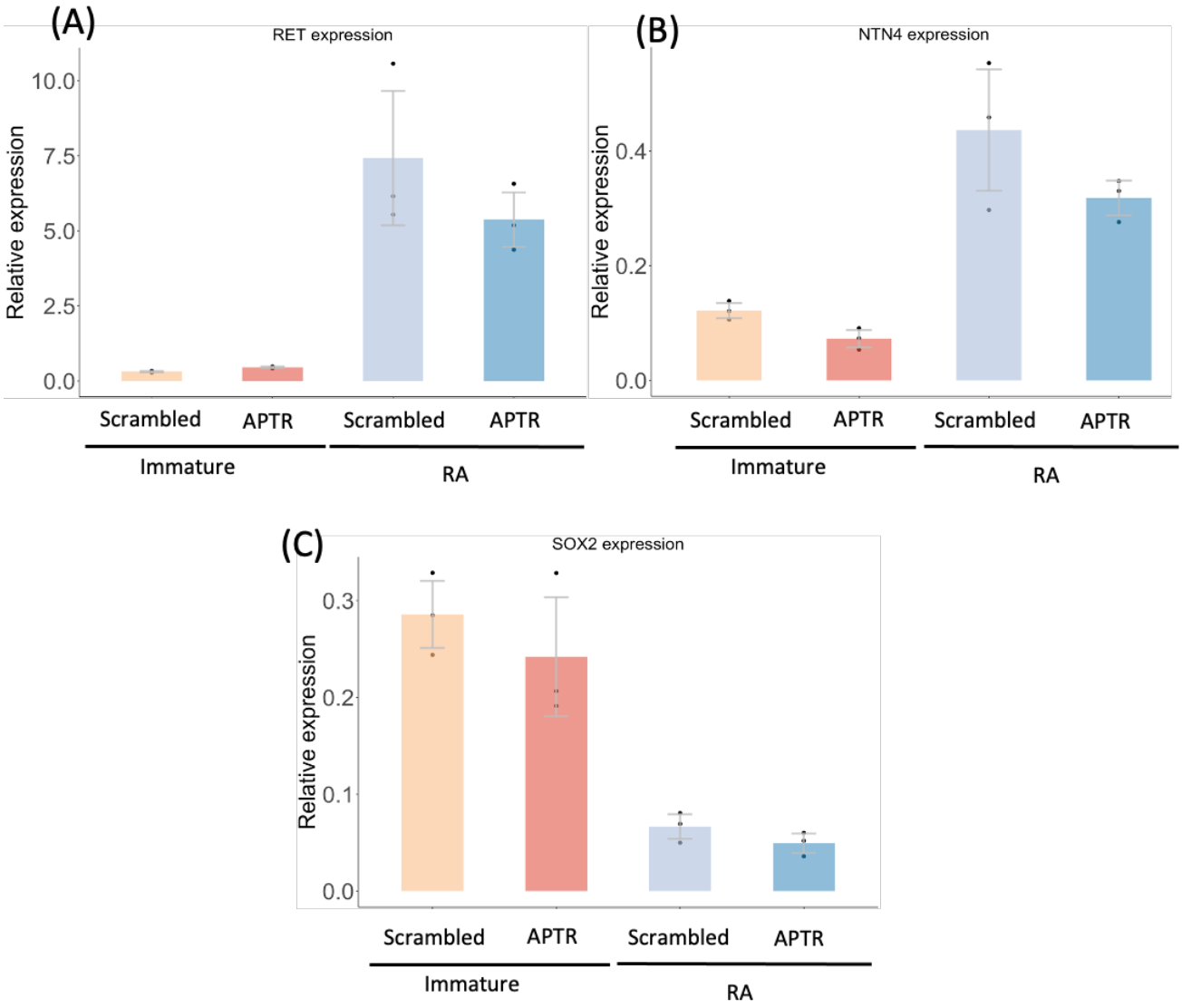


Figure 4.12: Knockdown of APTR has no effect on SH-SY5Y neuronal differentiation. *Relative quantification of mature neuronal markers (A) RET and (B) NTN4 and pluripotency marker (B) SOX2 in SH-SY5Y treated with either siRNAs targeting APTR and were compared to cells treated with a pool of scrambled siRNAs. Relative quantification was performed housekeeping genes NDUFV2 and HAUS8 and calculated with the ΔCq method. Two-way paired t-test, N=3. Individual biological repeats denoted with dots on bar charts. P value reported on graphs.*

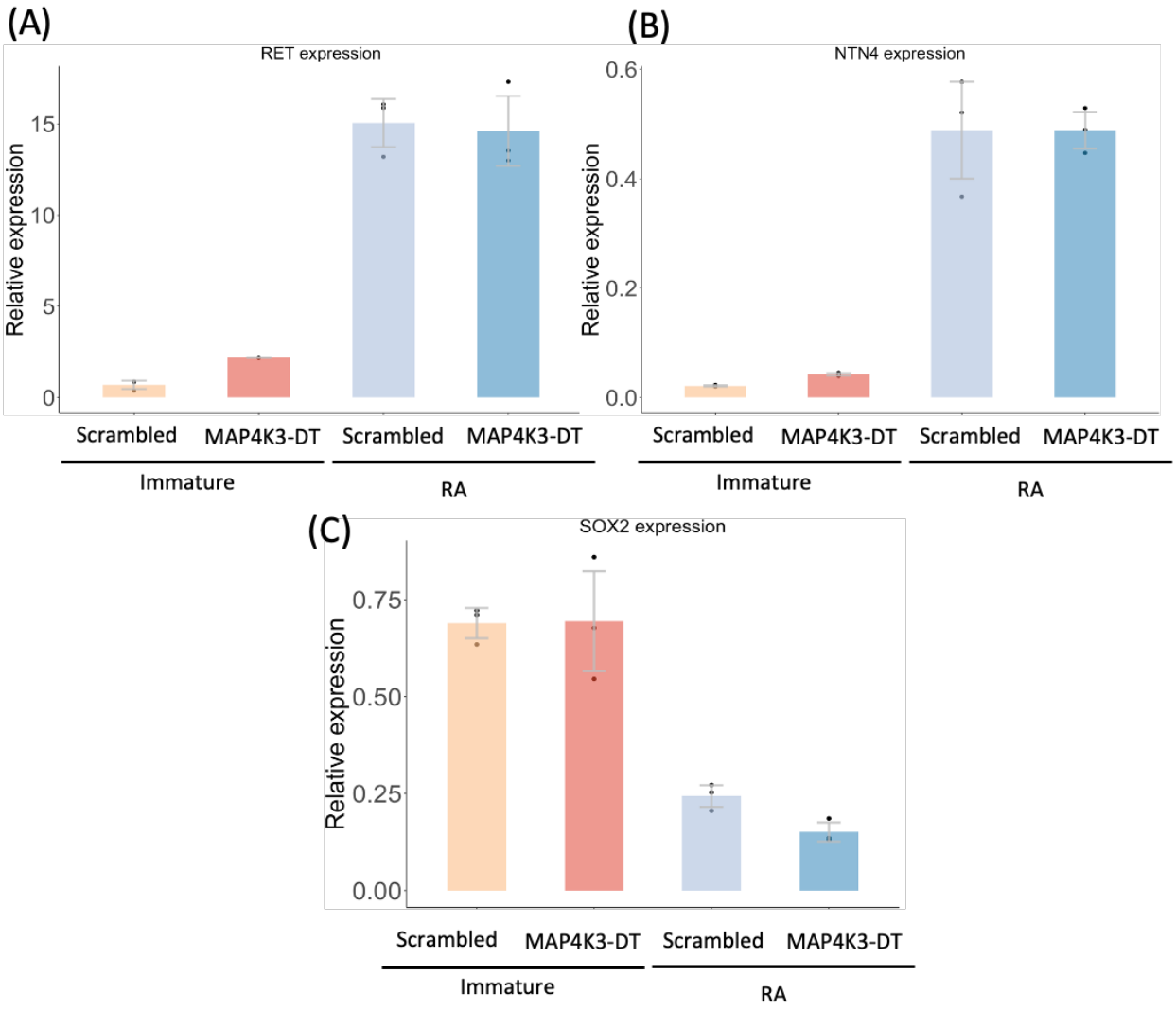


Figure 4.13: siRNA knockdown of MAP4K3-DT has no effect on SH-SY5Y marker gene expression. *Relative quantification of mature neuronal markers (A) RET and (B) NTN4 and pluripotency marker (B) SOX2 in SH-SY5Y treated with either siRNAs targeting MAP4K3-DT and were compared to cells treated with a pool of scrambled siRNAs. Relative quantification was performed housekeeping genes NDUFV2 and HAUS8 and calculated with the ΔCq method. Two-way paired t-test, N=3. Individual biological repeats denoted with dots on bar charts. P value reported on graphs.*

4.5 GIHCG overexpression results in changes to mitochondrial structure in immature SH-SY5Y cells

Mitochondrial network dynamics, the balance of mitochondrial fusion and fission, is essential in cell cycle, metabolic regulation and is known to underpin multiple diseases (Chen et al., 2023). The importance of mitochondrial dynamics in neuronal differentiation and development is well documented with a transition from branched to a more fragmented state being associated with NSC fate commitment and differentiation (Khacho et al., 2016). Additionally, neurodegenerative conditions including HD and AD display excessive or dysregulated mitochondrial fission (Flippo & Strack, 2017; W. Wang et al., 2020)

From the FLAG tagging screen (Chapter 3.12) it was apparent that GIHCG peptide localises to the mitochondria in SH-SY5Y cells. Transfection of GIHCG-sORF seemed to affect the shape of the mitochondria, with mitochondria in cells overexpressing GIHCG-FLAG appearing to have a more rounded and less connected morphology compared to non-transfected controls (Figure 4.14a).

To quantify the effect of GIHCG-peptide's effect on SH-SY5Y mitochondrial morphology the Fiji plugin MiNA (Valente et al., 2017)was used to measure the mitochondrial morphology in cells overexpressing GIHCG-peptide compared to control cells. Overexpression of GIHCG causes a significant reduction in mitochondrial branch length in transfected SH-SY5Ys compared to controls (Figure 4.14b). This effect is present in both immature and differentiated SH-SY5Y cells and indicates GIHCG overexpression leads to an increase in the division (fission) of large mitochondrial networks into smaller more spherical units.

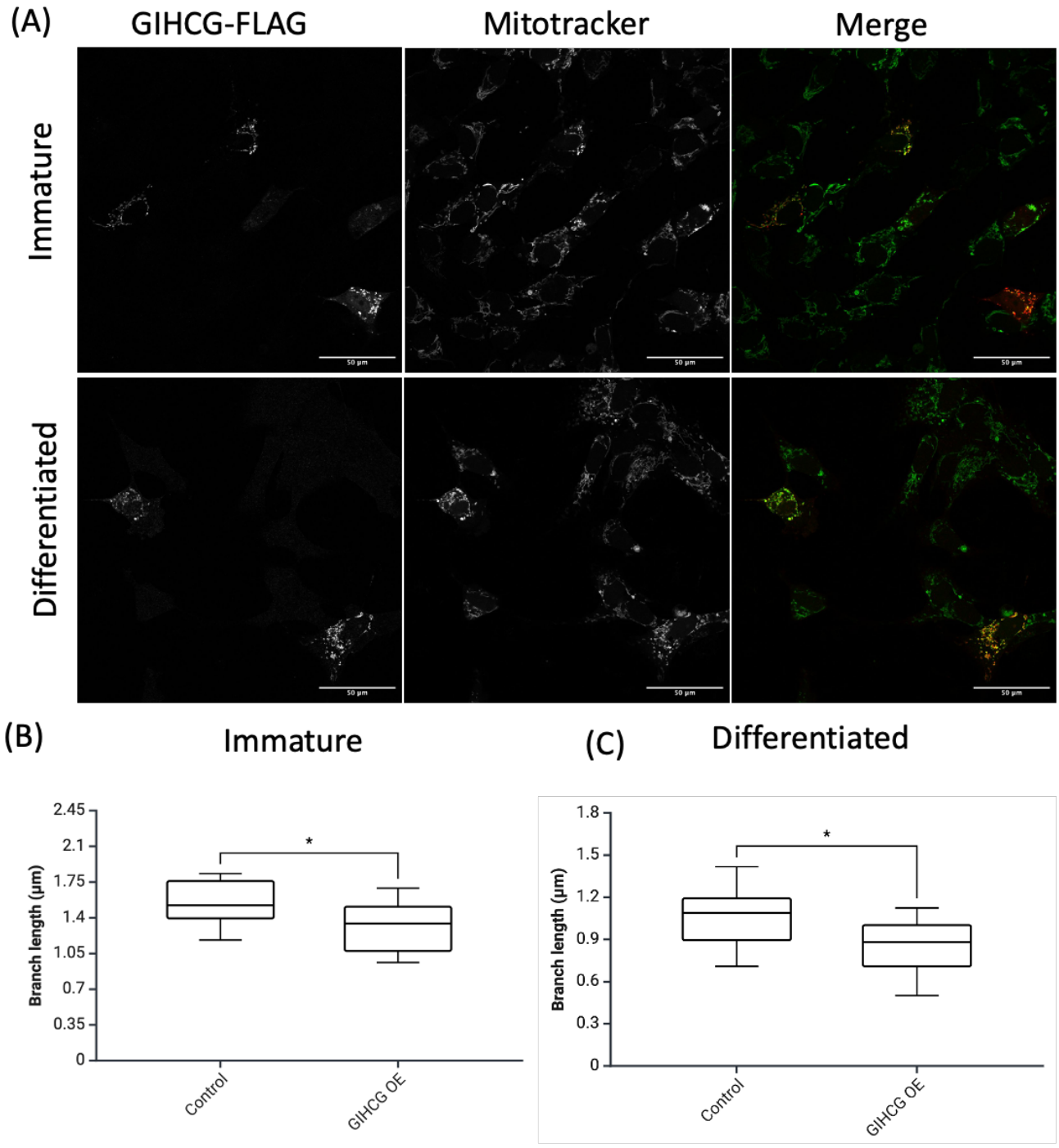


Figure 4.14: Overexpression of GIHCG results in shortening branch length of mitochondria in immature and differentiated SH-SY5Y cells. *(A) representative images of immature and differentiated SH-SY5Y cells overexpressing GIHCG-FLAG (red) and mitochondria are stained with Mitotracker CMXRos (green). Cells overexpressing GIHCG-FLAG appear to have a more rounded mitochondria morphology. (B) Quantification of mitochondrial branch length with the Mitochondrial Network Analysis (MiNA) toolset. Mitochondrial mean branch length of cells overexpressing GIHCG is significantly shorter than controls in both immature and differentiated SH-SY5Y cells. N= 3, with 28 total cells measured per condition. Kruskal-Wallis test with Dunn's multiple comparisons test performed as suggested by MiNA creators* (Valente et al., 2017). ** denotes p < 0.05.*

4.6 Translated lncRNAs detected in SH-SY5Ys are expressed in and enriched in the polysomes of cortical organoids

Many of the translated lncRNAs detected in SH-SY5Y cells are expressed during normal human development and disease states according to publicly available datasets (Chapter 3.2). While SH-SY5Y are a powerful and extensively used model for neuronal differentiation they are limited in the differentiation pathway they can undertake. Cortical organoids/spheroids (hCSs) are the current gold standard for modelling organ development with hCSc better modelling the CNS organisation, cell-cell contacts and can differentiate further than 2D culture (Eigenhuis et al., 2023; Paşca et al., 2015).

Using organoids kindly generated by Erica Harris in Dr James Poulter's group (Faculty of Medicine, University of Leeds) I investigated both the expression and translation of translated lncRNAs in cortical organoids.

4.6.1 Translated lncRNAs are dynamically expressed in organoid development

The first question I sought to answer was do cortical organoids express the translated lncRNAs identified in SH-SY5Y cells and how do their levels change during organoid development. Per biological replicate, 6 Cortical organoids were grown by Erica Harris and collected at ~ 3-day intervals from 3 day after iPSC seeding to day 26. This covers the phases of neural induction, neuronal precursor cell expansion and the start of differentiation (Figure 4.15). I extracted RNA from the organoids, generated cDNA and lncRNA abundance quantified using ddPCR. Absolute quantification with ddPCR was performed due to the difficulty in finding a stable housekeeping gene across a long-time course. Levels of the mRNA NDUFV2 were measured as a reference to compare the abundance of our translated lncRNAs to, allowing us to identify if lncRNAs were highly or lowly abundant in organoids (Figure 4.15).

All 4 lncRNAs tested were detected in cortical organoids. LIPT2-AS1 is highly expressed (~twice as highly expressed as NDUFV2) throughout the time course showing a stable level of RNA beyond day 6. Whilst GIHCG RNA exhibited low levels at early time points (Day 3) it steadily increased throughout the NPC expansion stage but never reach the level of NDUFV2

mRNA abundance. LINC00839 while detected throughout the time-course but at low levels (Figure 4.15).

4.6.2 Translated lncRNAs are enriched in the polysomal fraction of cortical organoids

To determine whether the translated lncRNAs expressed in cortical organoids were also translated in the organoids, day 25 organoids were collected and polysome fractionation was performed with the help of Dr Karl Norris (Figure 4.16). This enabled the purification of the 40S, 60S, 80S and polysomal fractions. RNA extractions were performed on these fractions so the abundance of lncRNA associated with each ribosomal component could be determined (Figure 4.16). For actively translated RNAs a clear enrichment in the polysomal fraction is expected, as can be see for the mRNA NDUFV2 (Figure 4.17). Of the 5 translated lncRNAs tested, 4 showed a clear enrichment in the polysomal fraction compared to the other ribosomal fractions (Figure 4.17). This indicates active translation of the lncRNAs in cortical organoids.

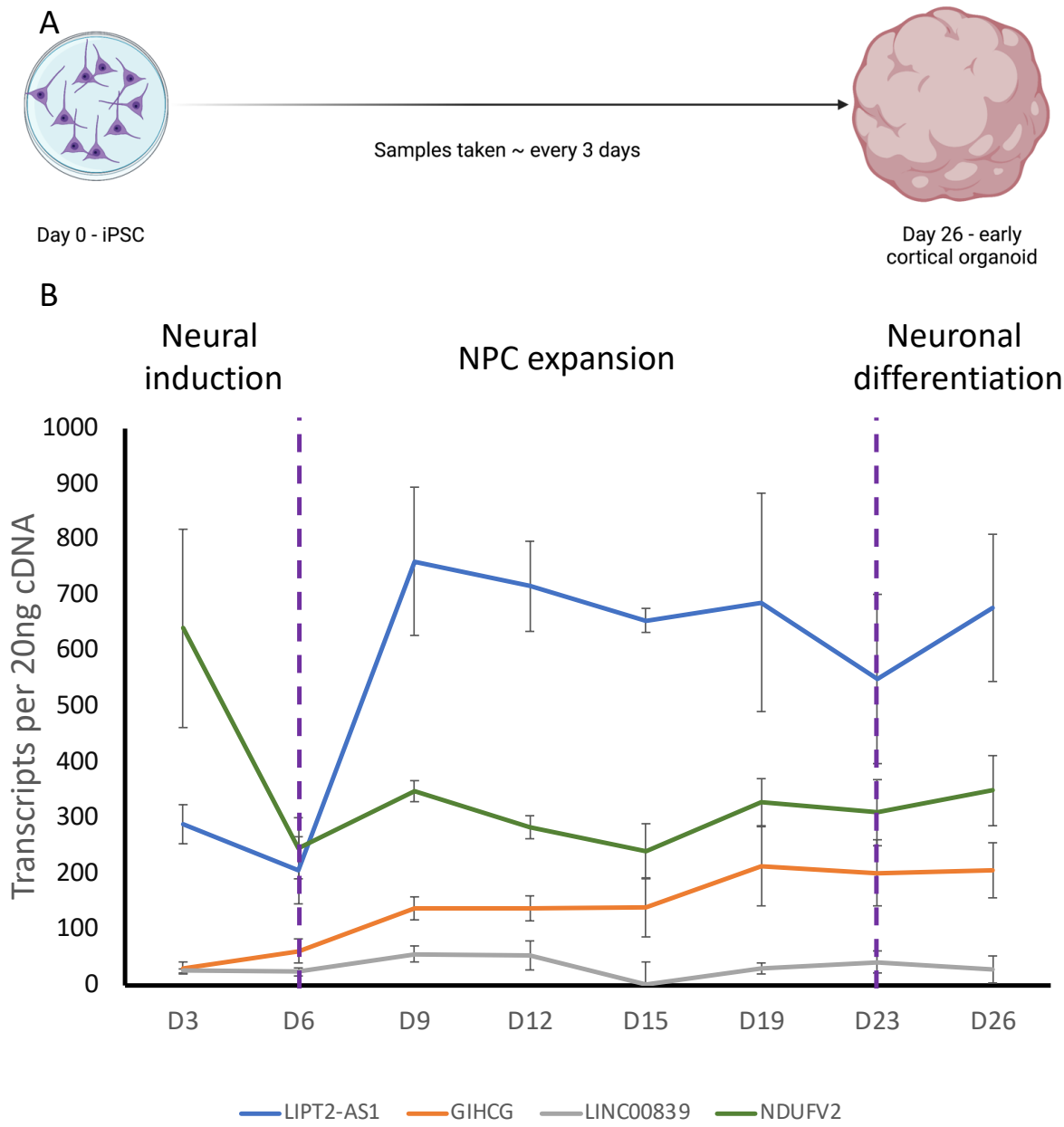


Figure 4.15: lncRNAs abundance is comparable to that of mRNA throughout early stages of cortical organoid development. *(A) schematic of organoid growth over 26 days (B) RNA abundance of lncRNAs LIPT2-AS1, GIHCG and LINC00839 as well as mRNA NDUFV2 across the time course of organoids development. Dotted lines denote key stages of organoids development. ddPCR was used to find the absolute abundance of lncRNA transcripts within 20ng of total RNA. N=3. Error bars show SD.*

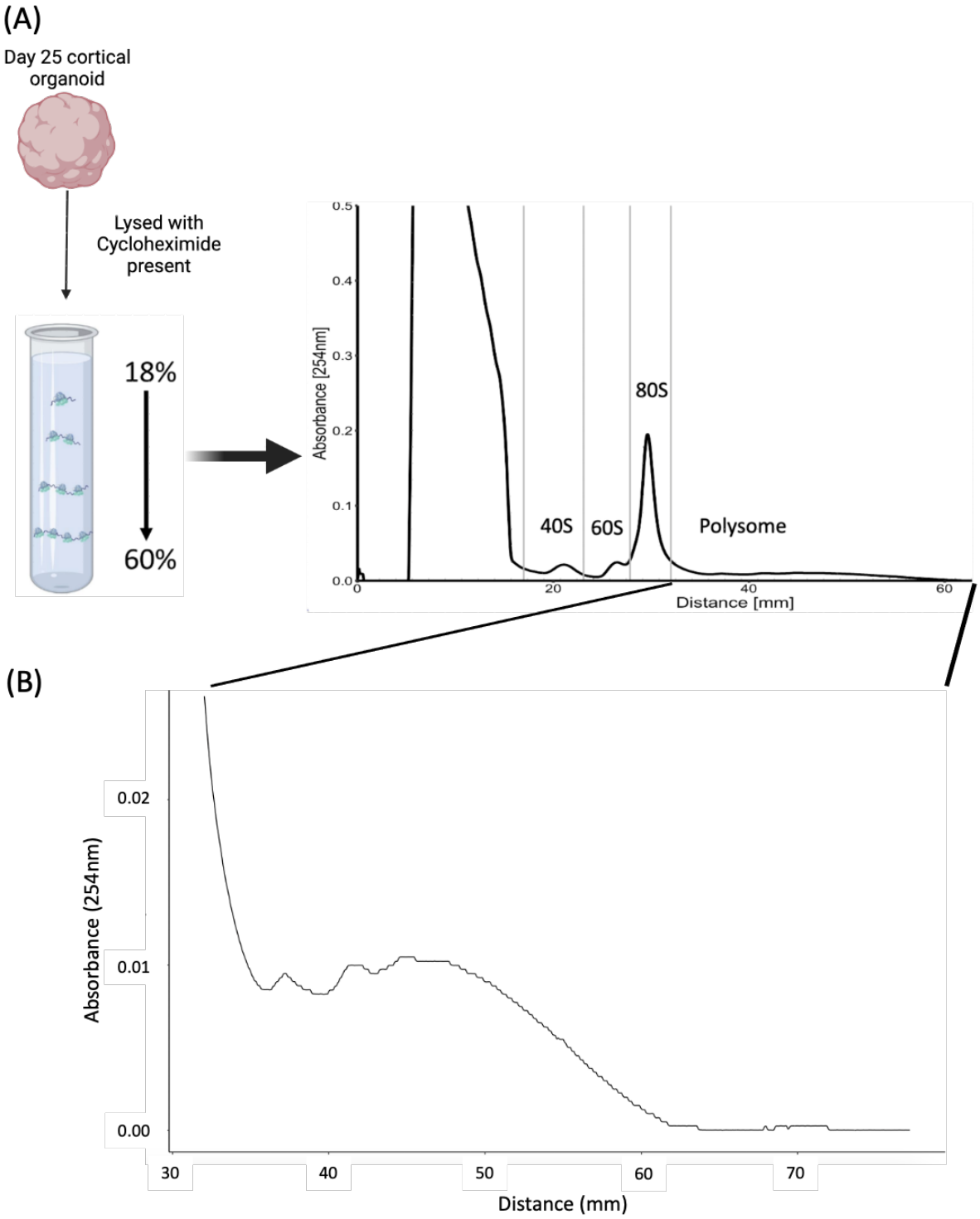


Figure 4.16: Polysome trace from day 25 cortical organoids shows clear 40S, 60S, 80S and polysome fractions. *(A) Organoid lysates from ~80 individual organoids were loaded onto a 18%-60% sucrose gradient. Gradients were spun at 121,355 x g_{avg} for 3.5 hours at 4 ºC in ultracentrifuge. Trace displays absorbance at 254nm across the gradient (mm). The trace shows peaks corresponding to the 40S, 60S, 80S and polysomal fractions. Fractions, indicated by grey lines, were collected for RNA extractions. (B) Zoom in on polysome fraction showing peaks of differently sized polysomes.*

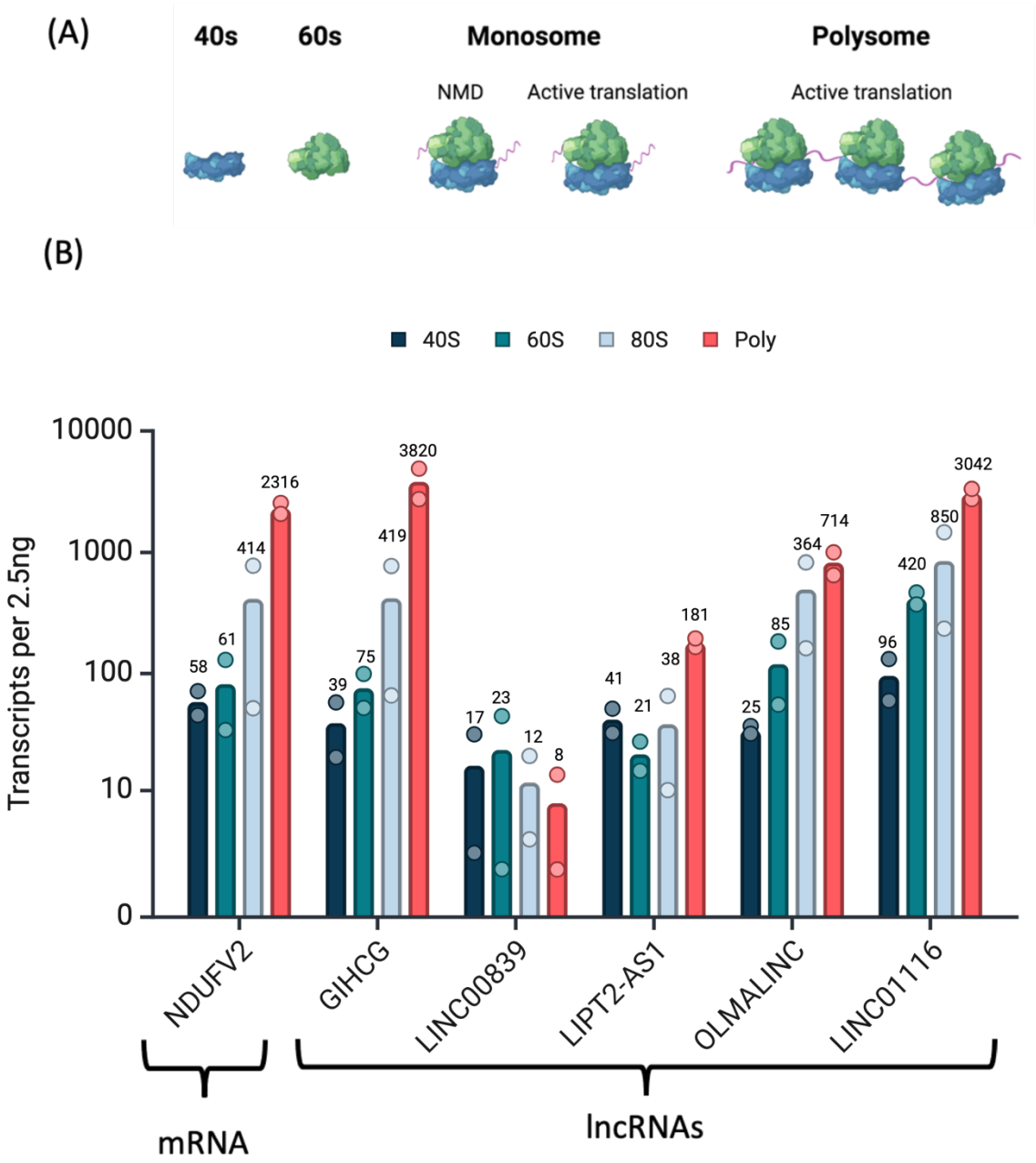


Figure 4.17: All lncRNAs except LINC00839 are enriched in polysomes of early-stage cortical organoids. *(A) Schematic of different ribosomal fractions collected. (B) RNA from 40S, 60S, 80S and Polysome fractions collected from day 25 cortical organoids was isolated and reverse transcribed. ddPCR was used to find the absolute abundance of lncRNA transcripts within 2.5ng of isolated RNA. mRNA NDUFV2 was included as a reference gene. N=2, biological replicates shown by dots. Mean transcripts per 2.5ng RNA listed above bars. Y axis log10.*

4.7 Discussion

4.7.1 Multiple lncRNAs are required for normal neuronal differentiation

This chapter aimed to identify if any of the translated lncRNA discovered in SH-SY5Y cells are required for neuronal differentiation through siRNA knockdown and assessing the knockdown cell's ability to differentiate with RA treatment (Table 4.1). From the 6 candidate translated-lncRNAs tested, 4 produced a phenotype in the screen.

The dysregulation of neuronal marker gene upregulation and reduction in neurite length observed with both GIHCG and LINC00839 knockdown suggests these translated lncRNAs are required for normal neuronal differentiation. Despite similar phenotypes observed GIHCG and LINC00839 are likely affecting different pathways as there are clear differences in their precise effect on neuronal marker genes with GIHCG only effecting NTN4 levels and having a more subtle, and not significant, effect on neurite length. LINC00839 knockdown has a more global effect on SH-SY5Y cells with the disruption of both neuronal and stem marker genes. Interestingly, LINC00839 appears to have opposite effects in immature and differentiated SH-SY5Y cells with immature cells showing a reduction in SOX2 levels and an increase in RET, indicating the cells are beginning to differentiate. Alternatively, RA treated SH-SY5Ys without LINC00839 are not able to differentiate as normal characterised by a poor upregulation of mature neuronal marker genes.

OLAMLINC and LIPT2-AS1 siRNA knockdowns show similar effects on neuronal marker gene expression in differentiated SH-SY5Y cells with significantly upregulated levels of NTN4 and RET. This may suggest that the loss of OLMALINC and LIPT2-AS1 allows for a greater extent of differentiation over 3 days or RA treatment. Alternatively, it may result from the cells not being able to differentiate effectively and therefore an over extensive upregulation of pro-neuronal factors is produced to try and compensate.

While silencing of APTR and MAP4K3-DT had no effect on SH-SY5Y differentiation, it must be considered that these lncRNAs may have a function that is independent of SH-SY5Y neuronal differentiation. As these translation products are novel it is unknown what pathways or biological processes they are likely to be involved in. Previously, APTR has been identified as participating in the progression of osteosarcoma (Guan et al., 2019) but also is thought to be a protective factor in gastric cancer (Ren et al., 2022). MAP4K3-DT is less well characterised,

though has been found to be expressed head and neck squamous cell carcinoma but its molecular function if it has one has yet to be determined (Zheng et al., 2023).

For LIPT2-AS1, GIHCG and LINC00839 this is the first time they have been shown to play cellular role, with current literature describing roles for GIHCG and LINC00839 in disease states. LINC00839 has multiple attributed functions as an RNA in many contexts including in neuroblastoma where LINC00839 RNA is proposed to act as a molecular sponge for multiple miRNAs, allowing LINC00839 expression to regulate neuroblastoma cell proliferation and associating with poor prognosis (Q. Zhang et al., 2022). Additionally, LINC00839 upregulation has been described in response to hypoxia in liver cells, where LINC00839 is through to function via RNA-protein interactions promoting the progression of hypoxia-induced liver cancer (Xie et al., 2022). Similarly, GIHCG has a well characterised associating its expression with cancer progression and poor prognosis, including in glioblastoma (Hazra et al., 2023) and colorectal carcinoma (Jiang et al., 2019; G. Liu et al., 2019) amongst others. All the proposed mechanisms of function for GIHCG to date are that GIHCG RNA acts as a molecular sponge for miRNAs (G. Liu et al., 2019; S.-Y. Zhu et al., 2022). LIPT2-AS1 has no functional molecular function defined. OLMALINC has previously been identified as required in the differentiation of oligodendrocytes and neuronal-type cells (Mills et al., 2015). However, the phenotypes described following OLMALINC knockdown in SK-N-SH were not recapitulated in SH-SY5Y cells. This could potentially be due subtle differences in the two related cell lines or due to the relatively weak knockdown achieved with OLMALINC siRNA with the efficiency of ~60% knockdown.

Importantly, from this siRNA mini-screen it is not possible to determine if the translated lncRNAs are functioning at the RNA level, as a peptide, or both. This is particularly important for the candidates OLMALINC, GIHCG and LINC00839 that have characterised and functions at the RNA level. The GIHCG-peptides seems to exhibit a function, as evidenced by the overexpression of the sORF alone (not the full-length RNA) altering mitochondrial dynamics in SH-SY5Y cells (Figure 4.14). Due to the poor transfection efficiency of SH-SY5Y rescue experiments through overexpression of lncRNA (s)ORFs alone was not possible in this system. Additionally, the siRNAs would also target the rescue construct for degradation. Therefore, in the next chapter I will be using CRISPR/cas9 to elucidate whether it is the

lncRNA or peptide of these translated lncRNAs that contribute to neuronal differentiation phenotypes upon siRNA knockdown.

Gene name	Translated in	Localisation	RNA phenotype	Neurite phenotype
LINC00839	RA	Cell periphery	Reduced NTN4 and RET	Neurite length reduced
GIHCG	RA	Mitochondria	Reduced NTN4	Neurite length reduced (Not significant)
LIPT2-AS1	Immature	Nuclear	Increased NTN4 and RET	Neurite length reduced
OLMALINC	Immature	Nuclear	Increased NTN4 and RET	No phenotype
APTR	RA	Cytosolic	No effect	N/A
MAP4K3-DT	Immature	Cytoplasmic puncta	No effect	N/A

Table 4.1: Summary of translated lncRNAs included in siRNA screen and effect of silencing on SH-SY5Y differentiation

4.7.2 Translated lncRNAs are expressed and enriched in the polysomes of cortical organoids

All 5 candidates investigated could be detected in cortical organoids at the RNA level throughout the early stages of development (Figure 4.15). LIPT2-AS1, showed high expression throughout the entirety of NPC expansion stage, with approximately twice the level of RNA than the mRNA NDUFV2. GIHCG also showed robust expression throughout the time-course with levels increasing with time. This regulation of RNA levels suggests these two translated lncRNAs are involved in cortical organoid maturation (Cardoso-Moreira et al., 2019; Sarropoulos et al., 2019).

Again, all translated lncRNAs investigated were found within the polysomal faction of day 25 cortical organoids (figure 4.17). GIHCG was the most highly abundant RNA tested in the polysomal fraction of cortical organoids, despite being present at lower levels than NDUFV2 and LIPT2-AS1 in total RNA. This may highlight a preferential translation of certain lncRNAs in cortical organoids and indicates at day 25 of cortical organoid development GIHCG is being actively translated. LINC00839 is lowly abundant in total and polysomal RNA fractions. However, this could be due to the timeframe we investigated as neuronal differentiation only begins at ~ day 23 so LINC00839 may be more highly expressed at later time points.

Further, Poly-Ribo-Seq showed that LINC00839 had a greater translation efficiency (TE = 2.07) that GIHCG (TE = 0.21), LIPT2-AS1 (TE = 0.29) and OLMALINC (TE = 0.90) in SH-SY5Y cells (Douka et al., 2021). Therefore, despite relatively low RNA levels of LINC00839 in cortical organoids proportionally large levels of LINC00839 protein could be generated.

Combining the strong enrichment of translated lncRNAs in the polysomal fraction of cortical organoids with the stable protein products observed in the FLAG-tagged assay suggests that lncRNA derived peptides will be present within cortical organoids. Additionally, as NMD targets are typically depleted from polysomal fractions compared to total RNA and monosome fractions, the enrichment in the polysomal fraction seen for 4/5 translated lncRNAs supports they are undergoing active translation rather than having their levels regulated by NMD (Lloyd et al., 2020; Zeng et al., 2018). However, as total organoids were harvested it is not possible to comment on which sub-type of cells the translated lncRNAs are expressed and translated in. Single-cell RNA-Seq datasets from organoids could be investigated to determine if translated lncRNAs are globally expressed in organoids or restricted to a subset of cell-types (Velasco et al., 2019).

4.8 Conclusion

In this chapter I investigated the requirement of 6 translated lncRNAs for normal neuronal development, using SH-SY5Y cells as a model for early neuronal differentiation. I identified 4 translated lncRNAs, which are essential for SH-SY5Y differentiation. Additionally, I validated the expression and polysome association of the functional translated lncRNAs in an advanced model of neuronal development, cortical organoids. All translated lncRNAs tested were detected in total RNA with GIHCG, OLMALINC and LIPT2-AS1 showing enrichment in the polysomal fraction, comparable to that of mRNA, suggesting active translation.

Chapter 5: Characterisation of novel ‘lncRNA’ encoded proteins

5.1 Introduction

Cytoplasmic lncRNAs are known to function at an RNA level through a diverse range of mechanisms (Figure 1.2). Additionally, it is now accepted that some lncRNA genes function as an RNA and are translated into a functional protein. These genes are referred to as either dual functional (T. Liu et al., 2023), binary functional RNA (bifunctional RNA) or coding and noncoding RNA (cncRNA) (Huang et al., 2021). Examples of dual function RNAs can be seen from both genes annotated as lncRNAs and mRNAs. From the gene locus of LINC00961 both a functional RNA (LINC0096) and a micropeptide SPAAR are produced, with both showing high conservation between humans and mice (Spencer et al., 2020). These gene products have opposing roles on angiogenesis as demonstrated with the overexpression of SPAAR protein alone in the endothelial cell line HUVECs resulting in increased tubule formation whereas overexpression of the full transcript (including SPAARs ORF) did not. Additionally, overexpression of the full-length transcript with SPAAR ATG mutated had the opposite effect with reduced network formation by HUVECs (Spencer et al., 2020). mRNAs from canonical protein-coding genes have been identified to function at an RNA level, further blurring the line between lncRNA and mRNA. p53 mRNA has been shown by co-IP to interact with Mdm2 protein in human epithelial-like (H1299) cells. This in turn enhances the translation and stability of p53 protein (Candeias et al., 2008). Additionally, from the HIST1H1C locus an RNA and protein with unrelated functions are produced. H1.2 Linker Histone protein, which binds linker DNA between histones and drives the condensation of nucleosome chains. However, *HIST1H1C RNA* functions to negatively regulated telomere length independently of is protein product through direct interaction with human Telomerase RNA (Ivanyi-Nagy et al., 2018). The acceptance that genes can generate both RNAs that function as transcripts and proteins suggests many canonical genes will have undiscovered functions at the RNA level (Huang et al., 2021). Therefore, a key challenge when identifying functional proteins encoded from translated lncRNAs is distinguishing between RNA and protein function. For example, siRNA silencing of translated lncRNAs will knockdown both the lncRNA and its resultant peptide. Therefore, any phenotype observed could be due to degradation of the RNA or resultant loss of the peptide. This is particularly important for translated lncRNAs with established RNA functions.

GIHCG RNA has previously been shown to function as a molecular sponge in multiple disease states (Jiang et al., 2019; G. Liu et al., 2019; S.-Y. Zhu et al., 2022). Most notably, single cell RNA-Seq identified GIHCG is significantly enriched in glioblastoma stem cells (GSC), the cell type responsible for Glioblastoma multiforme treatment resistance. RNA FISH and RT-qPCR revealed that GIHCG RNA localises to both the nucleus and cytoplasmic fractions of patient derived GSC. siRNA knockdown of GIHCG in patient derived GSC resulted in a reduction of stem cell markers and a 44% reduction in GSC migration in a transwell migration assay (Hazra et al., 2023). Similarly, GIHCG was seen to be overexpressed in patient derived gastric cancer samples and artificial overexpression of GIHCG it significantly increased proliferation and migration of two gastric cancer cell lines, HGC-2 and SGC-7901. This effect could be partially reversed by also overexpressing miR-1281 suggesting that part of the GIHCG overexpression phenotype is due to sponging of miR-1281. This interaction was confirmed through a luciferase assay (G. Liu et al., 2019).

LINC00839 RNA has been implemented in promoting neuroblastoma progression through miRNA sponging (Q. Zhang et al., 2022). LINC00839 was identified to be upregulation in neuroblastoma tissue and cell lines. Additionally, LINC00839 expression positively correlated with MYCN amplification which is associated with poor prognosis. siRNA knockdown of LINC00839 in neuroblast cells, IMR32 and BE(2)-C, significantly reduced the cells proliferation rate, invasiveness and ability to form colonies. LINC00839's cancer progression effect is thought to be due to its ability of sponge the anti-tumour miRNA miR-454-3p, potentially leading to an upregulation of NEUROD1 (Q. Zhang et al., 2022). LINC00839 is also though to promote liver cancer progression through a different mechanism. RNA pull down assays of biotinylated LINC00839 in the Liver cancer cell line (Li-7) demonstrated that LINC00839 RNA interacts with multiple protein partners involved in metabolism, RNA transport and migration. Overexpression of LINC00839 promoted the proliferation, migration, and invasion of liver cancer cells under hypoxic culture conditions (Xie et al., 2022).

OLMALINC RNA is highly expressed in the human brain and is expressed at significantly higher levels in white matter than grey. Of the 4 translated lncRNAs that demonstrated an effect on neuronal differentiation OLMLAINC is the only candidate previously identified as having a role in differentiation. siRNA knockdown of OLMALINC in the SH-SY5Y precursor

cell line SK-N-SH and oligodendrocyte cell line (MO3.13) resulted in dysregulated gene expression. This effect was greater in oligodendrocyte cells where OLMALINC knockdown resulted in the upregulation of multiple genes including SOX4 that are incompatible with oligodendrocyte maturation (Mills et al., 2015). Additionally, recent evidence from CRISPR screens indicate mutations in OLMALINCs ORF, but not UTRs, effect multiple cancer cells viability suggesting that translation of OLMALINC may be involved in cancer cell survival (Hofman et al., 2024; Prensner et al., 2021).

Of the 4 translated lncRNAs that showed a phenotype in the siRNA screen only LIPT2-AS1 has no known functions at the RNA level. This includes LIPT2-AS1 although it is overexpressed in cuproptosis-related Hepatocellular carcinoma and head and neck squamous cell cancer (X. Liu et al., 2022; J. Wu et al., 2023; Zheng et al., 2023).

In chapter 3, the sizes and subcellular localisations of the novel lncRNA-derived peptides was determined by western blot and IF respectively. In chapter 4 a siRNA screen was performed on 6 candidates translated lncRNAs and revealed that knockdown of 4/6 translated lncRNAs resulted in a dysregulation of SH-SY5Y neuronal differentiation. This chapter aims to elucidate if the phenotypes exhibited in the siRNA knockdown of the translated lncRNAs GIHCG, LINC00839, OLMALINC and LIPT2-AS1 is the result of loss of protein. CRISPR-Cas9 was used to induce small mutations that will prevent protein expression in SH-SY5Y cells but not impact lncRNA expression and structure, therefore allowing us to determine the phenotype of just from protein knockout. Additionally, further characterisation of LIPT2-AS1 protein is performed to identify global effects of LIPT2-AS1 protein knockout on SH-SY5Y and to identify LIPT2-AS1 protein interaction partners.

5.2 Genotyping CRISPR clones

For each ORF, 2 sgRNAs were used in isolation. Where possible one sgRNA targeted the ATG of the ORF to generate a mutation that will prevent translation initiation, and a second sgRNA targeted downstream of the start site to generate a frameshift mutation. This will allow us to determine if any phenotypes seen from CRISPR mutations are due to the loss of the protein (both ATG and frameshift mutants will show a phenotype) or if phenotypes are dependent on the act of translation (only the ATG mutant will have a phenotype in this scenario).

To identify clonal colonies of SH-SY5Y cells containing mutations disrupting peptide production, either through the disruption of their ATG or introducing a frameshift or early stop codon, gDNA was extracted and the sequence surrounding sgRNA target sites sequenced. The effect on peptide/protein levels could not be determined as no antibodies are available for these novel peptides. CRISPR-Cas9 editing was performed on all 4 translated lncRNAs that showed a phenotype in the siRNA knockdown screen (Table 4.1)

5.2.1 GIHCG

Three clones containing suitable mutations were generated for GIHCG knockout. Two of these were heterozygous with one copy of GIHCG being mutated to disrupt the start site and a second mutation on the other gene copy resulting in a frameshift (Figure 5.1 A, B). The third clone contains a homozygous insertion of a single nucleotide that disrupts GIHCGs ORF start site preventing translation initiation (Figure 5.1 C).

5.2.2 LINC00839

Knockout of LINC00839 was unsuccessful. Multiple clones were generated with mutations; however, none of them were suitable for further investigation as 3 clones were heterozygous and still contained a WT sequence and therefore will be able to generate LINC00839 protein. Additionally, 2 clones were generated with both alleles mutated but the mutations were deletions or insertions being a multiple of 3 and therefore not generating a frameshift as desired (Figure 5.2).

5.2.3 OLMALINC

Three successful clones were generated for OLMALINC knockout in SH-SY5Y cells. Two showed identical genotypes with both containing a homozygous insertion of a single nucleotide resulting in a frameshift and truncation of the OLMALINC protein. The third clone is heterozygous with one copy of the gene containing the same single nucleotide insertion and the second copy containing a 14-nucleotide deletion which results in a frameshift and truncation (Figure 5.3)

5.2.4 CRISPR of LIPT2-AS1 generated a homozygous frameshift mutant

Genotyping of LIPT2-AS1 CRISPR clonal colonies revealed one successful clone. This clone (now referred to as LIPT2-AS1$^{-/-}$) had a homozygous insertion of a single nucleotide after the ORFs ATG resulting in a frameshift mutation in both copies of the gene (Figure 5.4 A/B). As LIPT2-AS1's ORF is still present a truncated protein will be made, however it will not contain the DNA binding domain of LIPT2-AS1. This can clearly be seen in the Alphafold model of wild type and mutant LIPT2-AS1 (Figure 5.4 C) and from the alignment of WT and mutant LIPT2-AS1 protein (Figure 5.5).

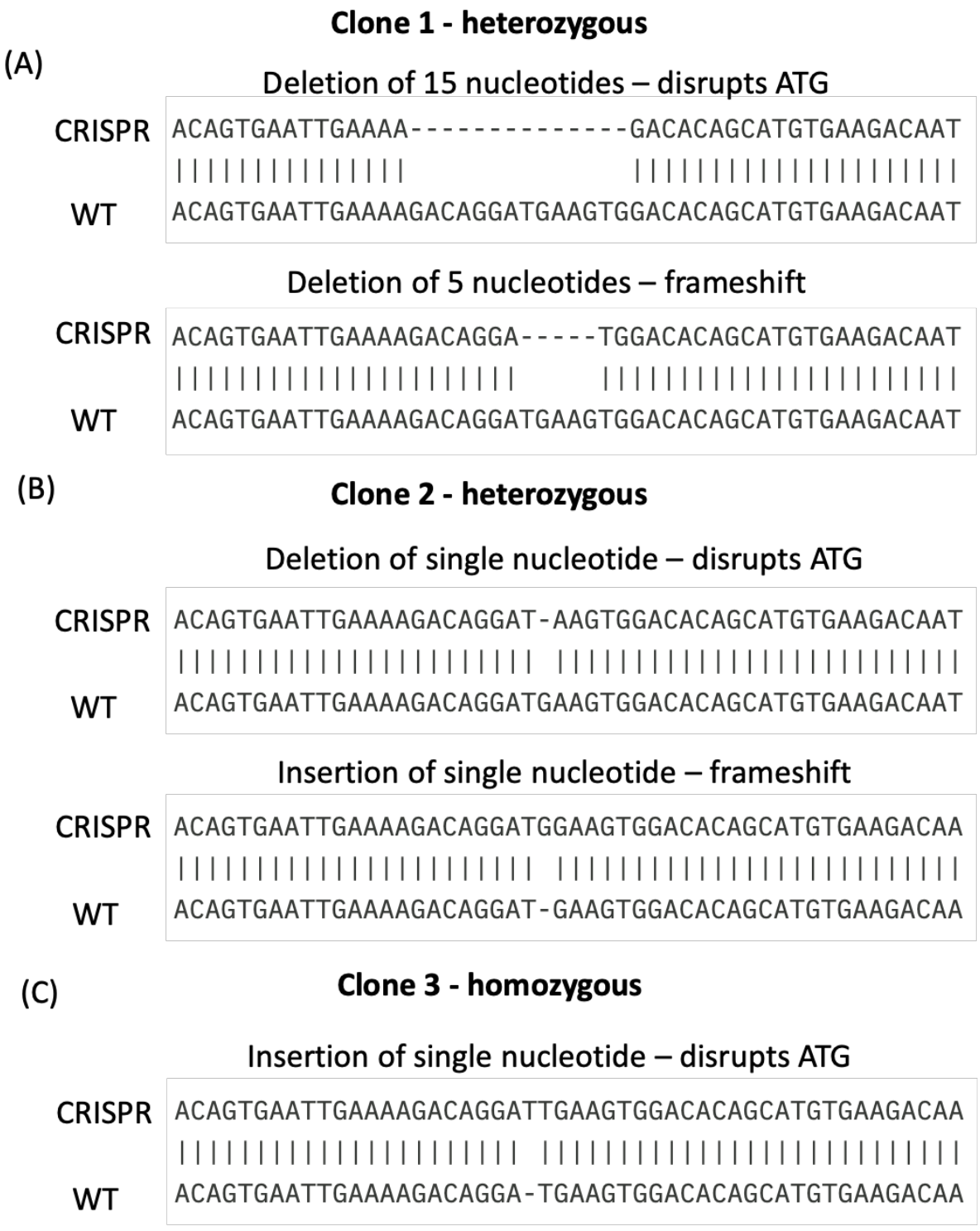


Figure 5.1: Genotype results of 3 successful GIHCG mutants*. (A) clone 1 is a heterozygous mutant with one mutation resulting in the loss of GIHCGs start codon and the second resulting in a frameshift. (B) clone 2 also a heterozygous mutant with one loss of start site and one frameshift mutation. (C) clone 3 is homozygous with a SNP that disrupts the sORFs ATG.*

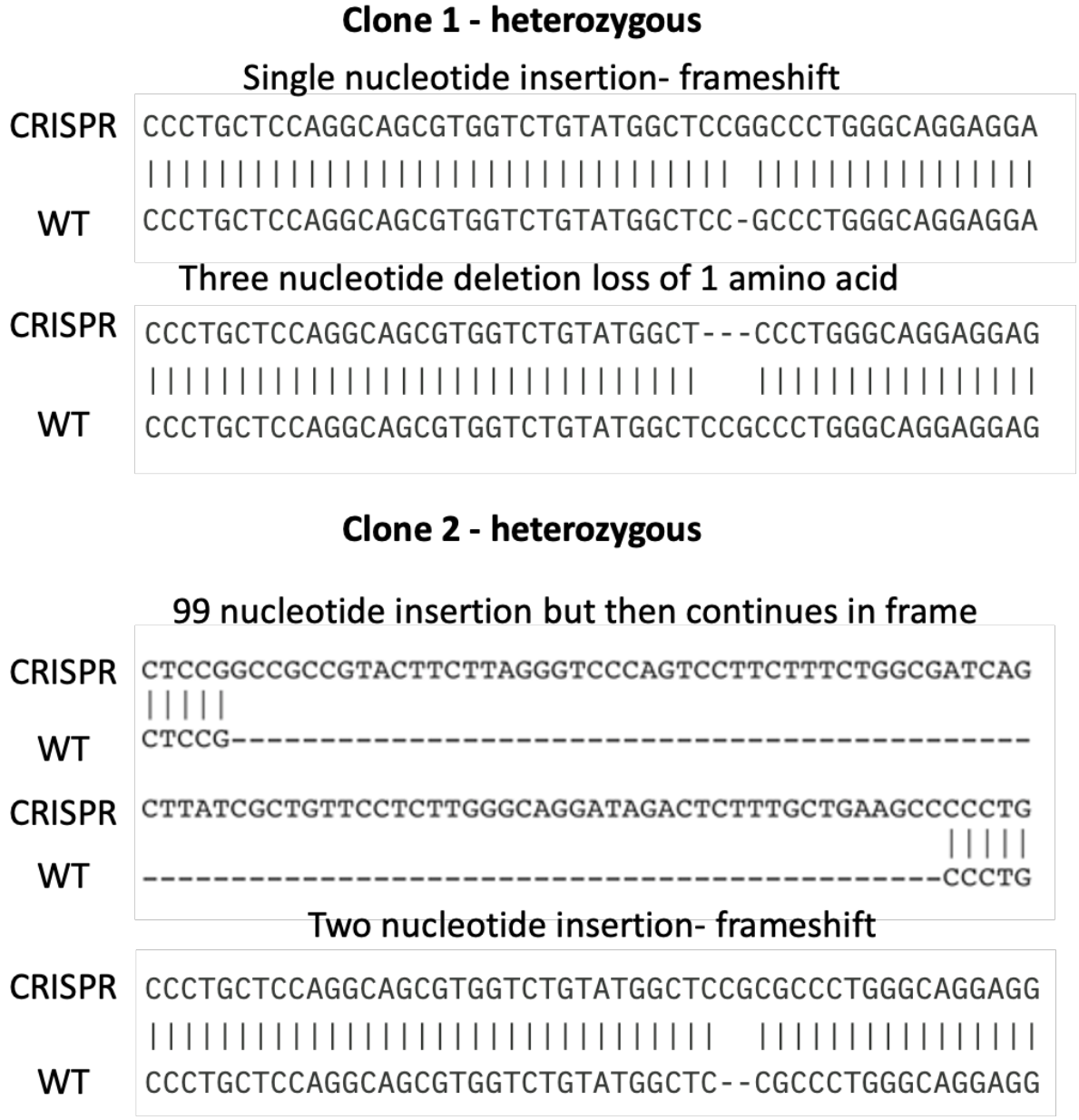


Figure 5.2: Genotype results of 2 unsuccessful LINC00839 mutants. *(A) Clone 1 is a heterozygous mutant with one SNP that results in a frameshift and the second copy containing a 3-nucleotide deletion. (B) clone 2 is a heterozygous mutant with one copy of LINC00839 containing a large insertion of 99 nucleotides and the other containing a 2-nucleotide insertion resulting in a frameshift.*

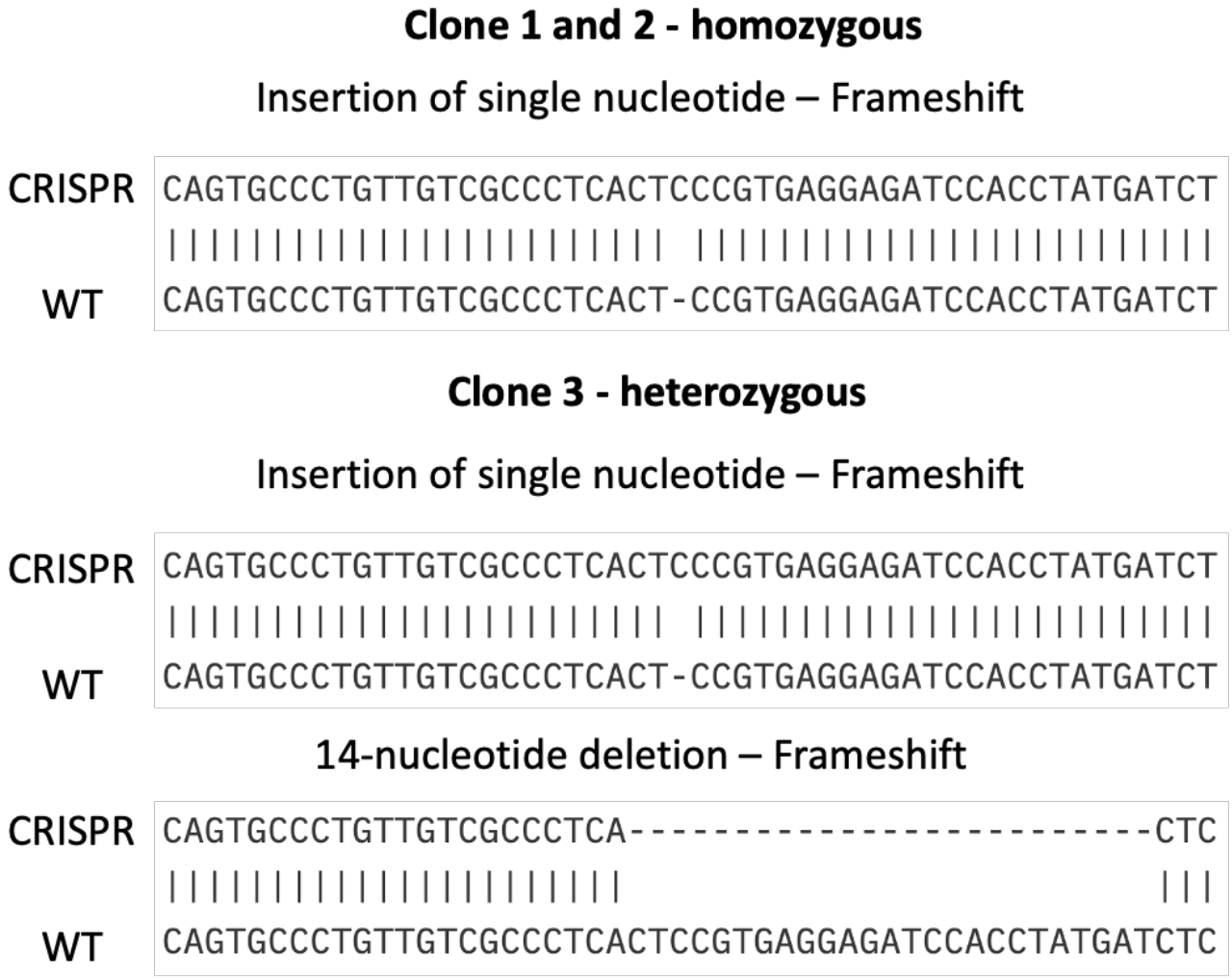


Figure 5.3: Genotype results of 3 successful OLMALINC mutants*. (A) clones 1 and 2 are homozygous mutants with a single nucleotide insertion resulting in a frameshift. (B) clone 3 is a heterozygous mutant one copy of OLMALINC showing the same single nucleotide insertion as clones 1 and 2. The second copy of OLMALINC in clone 3 contains a 14-nucleotide deletion. Both mutations result in a frameshift and truncation of OLMALINC protein.*

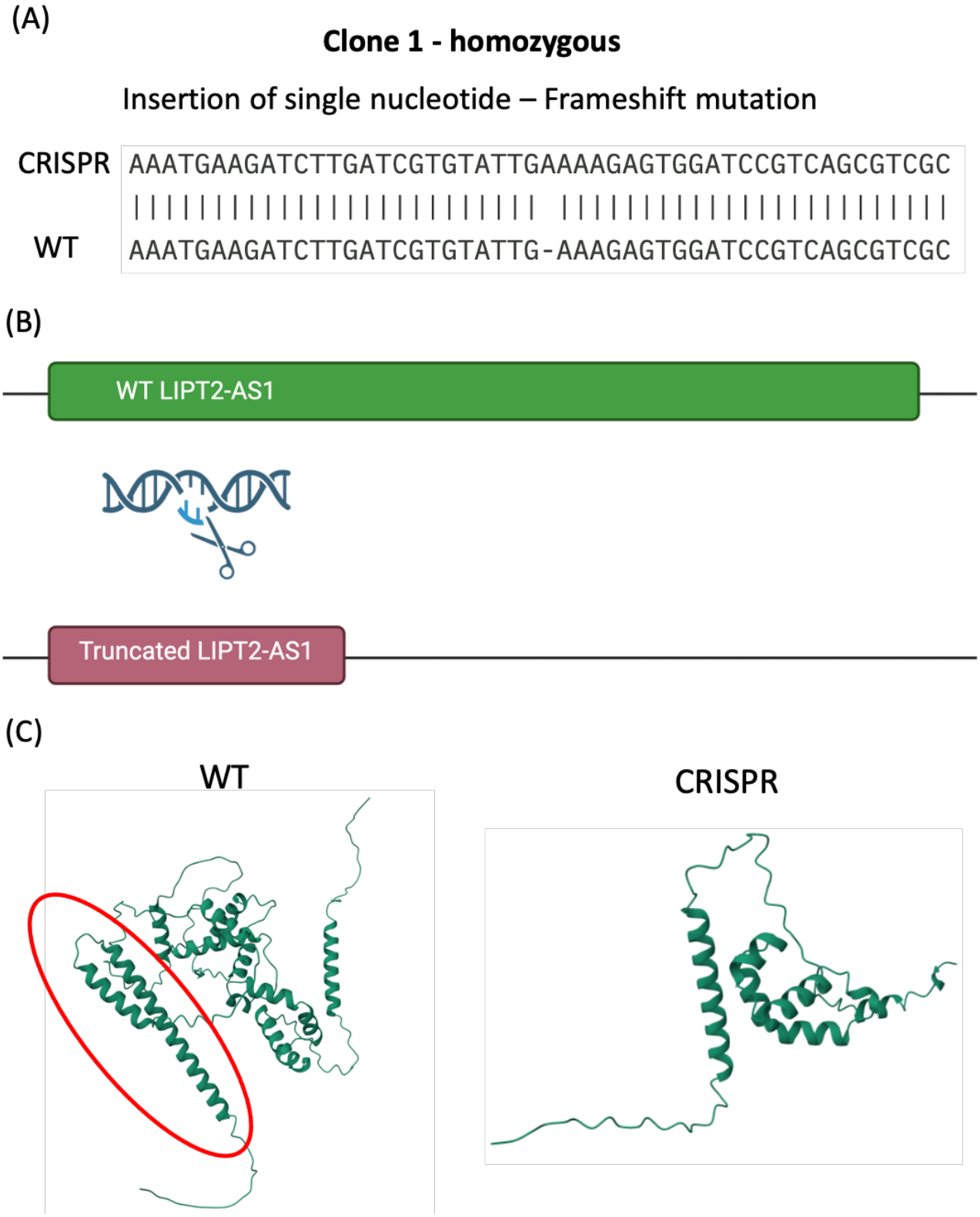


Figure 5.4 Genotype and overview of LIPT2-AS1 CRISPR mutant *(A) One successful knockout mutant was generated for LIPT2-AS1 containing a homozygous single nucleotide insertion resulting in a frameshift mutation. (B) schematic comparing WT LIPT2-AS1 to frameshift truncated LIPT2-AS1. (C) Alphafold models of WT and mutant LIPT2-AS1 protein. The HTH DNA binding domain is highlighted in red on the WT model. It can be seen as absent in the CRISPR proteins model.*

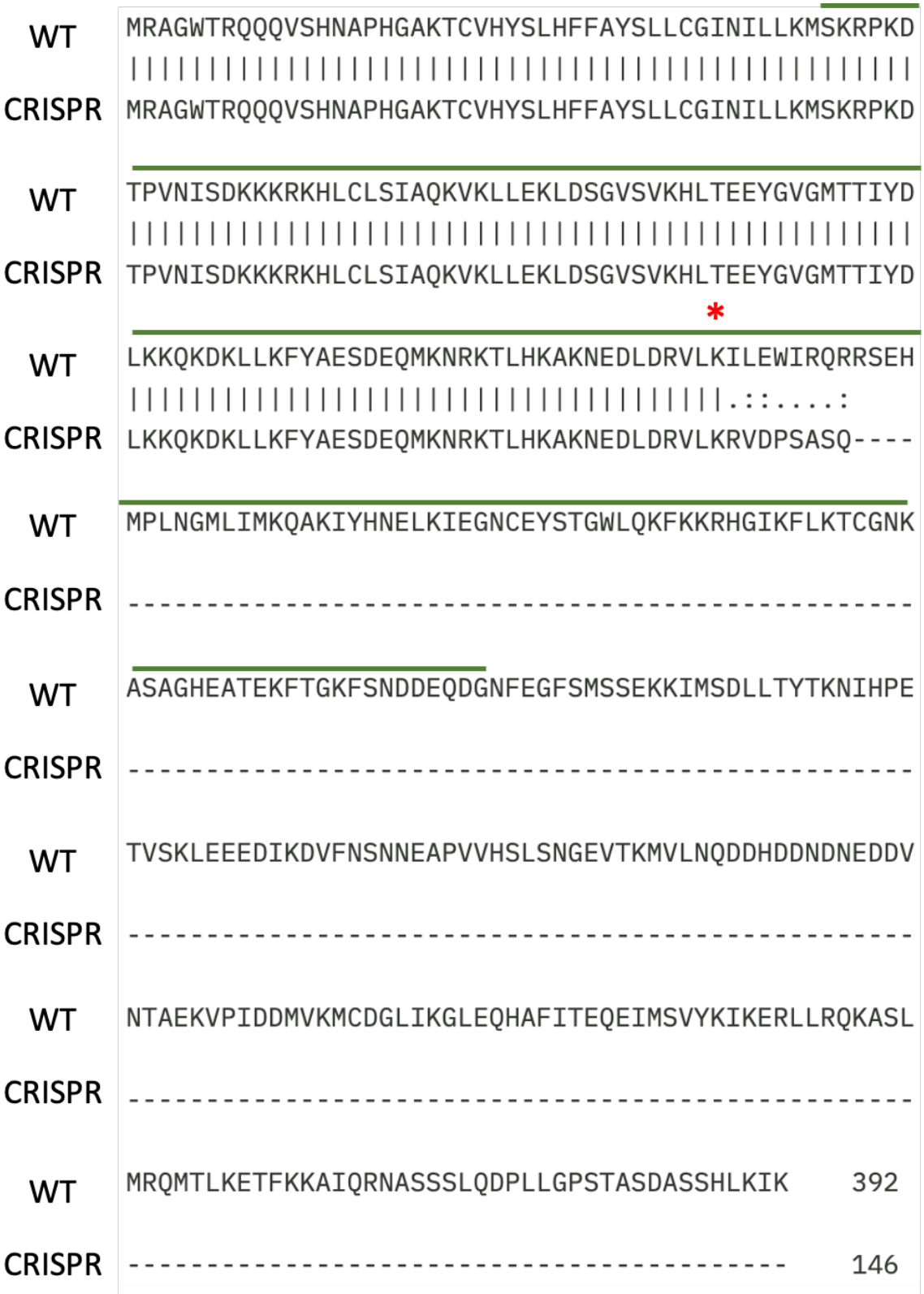


Figure 5.5: Alignment of WT and frameshift LIPT2-AS1 protein. *DNA binding domain is show by green bar above amino acid sequence and runs from 44-223. Mutation at point 137 is marked by red *.*

5.3 CRISPR editing of LIPT2-AS1 does not affect RNA structure or expression

To ensure any phenotype observed is due to the loss of LIPT2-AS1's protein and not a reduction of LIPT2-AS1 RNA levels, LIPT2-AS1 RNA levels were measured by RT-qPCR. There was no difference in RNA abundance between wildtype and LIPT2-AS1$^{-/-}$ mutant cells, both under immature and differentiated culture conditions (Figure 5.6 A). Additionally, to ensure there was no disruption of LIPT2-AS1's sense gene, LIPT2, RNA levels were also measured for the sense transcript. There was no difference in LIPT2 mRNA levels between WT and LIPT2-AS1$^{-/-}$ (Figure 5.6 B). The exact site of LIPT2-AS1's mutation does not overlap LIPT2 sense so will have no effect on the sense genes translation or protein production (Figure 5.6 C).

Additionally, to investigate the potential for loss of RNA activity due to disrupted RNA structure, RNA computational modelling using the RNAFold server (Hofacker, 2003) was used to investigate the structure of WT and LIPT2-AS1$^{-/-}$ RNA. While differences are apparent the structures are highly similar (Figure 5.7). Furthermore, LIPT2-AS1 has no described RNA functions.

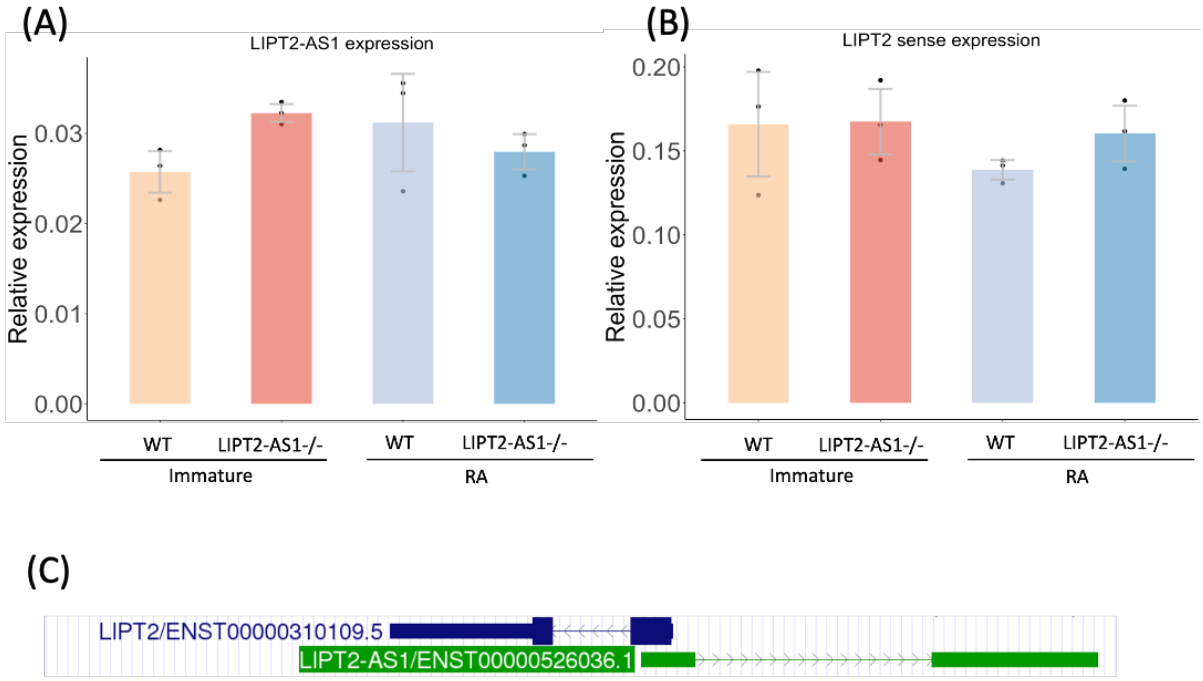


Figure 5.6: Frame shift CRISPR mutation of LIPT2-AS1 does not affect RNA levels of LIPT2-AS1 or its sense gene LIPT2. *(A) LIPT2-AS1 levels and (B) LIPT2 sense RNA abundance is not affected by CRISPR mutation of LIPT2-AS1. Relative quantification was performed using housekeeping genes NDUFV2 and HAUS8 and calculated with the ΔCq method. Two-way paired t-test, N=3. No comparisons passed the threshold for significance (p<0.05). Individual biological repeats denoted with dots on bar charts. (C) LIPT2-AS1 is encoded by a bi-directional promotor. Exon 1 of LIPT2-AS1 overlaps with exon 1 on LIPT2. The ORF (and mutation site) in LIPT2-AS1 is located entirely in the second exon and therefore not overlapping with LIPT2 sense gene.*

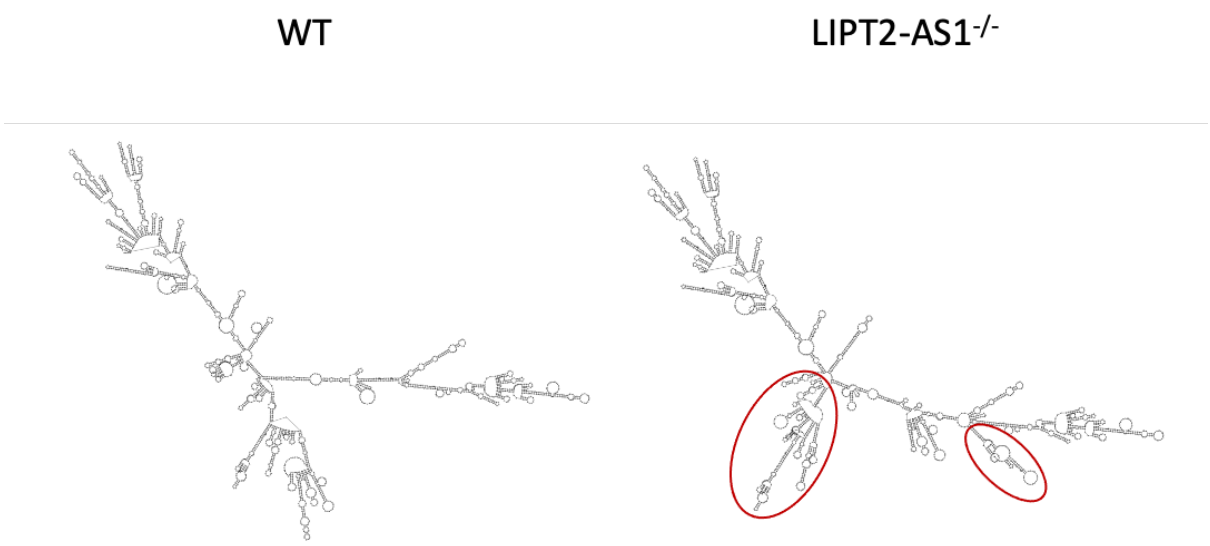


Figure 5.7: RNA structural predictions for WT and CRISPR LIPT2-AS1*. The 2493 nucleotide long WT sequence of transcript ENST00000526036 and 2494 nucleotide mutant were both modelled using the RNAFold server using minimum free energy prediction and default settings* (Hofacker, 2003)*. Main structural differences are highlighted in red.*

5.4 LIPT2-AS1$^{-/-}$ CRISPR cells recapitulate the siRNA knockdown phenotype

To determine if the frameshift mutation introduced into LIPT2-AS1 recapitulated the phenotype characterised for siRNA knockdown of LIPT-AS1, differentiation assays were performed. LIPT2-AS1$^{-/-}$ CRISPR cells compared to wildtype SH-SY5Y cells that underwent a mock transfection and single cell sort in parallel CRISPR cells. Neuronal and pluripotency marker gene levels were assessed by RT-qPCR and neurite outgrowth assay was used. If the phenotypes seen from the CRISPR mutant matched that seen in the siRNA knockdown of LIPT2-AS1 RNA it will indicate that the loss of the protein is underpinning the phenotype, rather than the RNA or combination of both RNA and protein.

5.4.2 LIPT2-AS1$^{-/-}$ cells are unable to grow neurites in response to RA

Neurites of differentiated SH-SY5Y cells following siRNA knockdown of LIPT2-AS1 cells were on average 34μm shorter than differentiated SH-SY5Ys treated with a scrambled control. Similarly, CRISPR LIPT2-AS1$^{-/-}$ cells showed a no increase in neurite length following RA treatment and when quantified the neurites of LIPT2-AS1$^{-/-}$ cells were significantly shorter (36μm) than wild type controls in the differentiated condition (Figure 5.8). This recapitulation of the siRNA phenotype indicates that LIPT2-AS1 protein is involved in neuronal differentiation and attenuation of neurite outgrowth seen with loss of LIPT2-AS1, either through siRNA knockdown or CRISPR editing, is solely due to the loss of LIPT2-AS1 protein product.

5.4.1 LIPT2-AS1$^{-/-}$ shows same disruption of neuronal marker genes as siRNA knockdown

siRNA knockdown of LIPT2-AS1 (chapter 4) resulted in a significant upregulation of neuronal marker genes NTN4 and RET in differentiated SH-SY5Y cells compared to scrambled controls but had no effect on SOX2 levels. Matching the siRNA phenotype, NTN4 and RET RNA levels were significantly higher in differentiated LIPT2-AS1$^{-/-}$ cells than WT SH-SY5Y cells (Figure 5.9 A, B). However, some differences were seen between the siRNA and CRSIPR phenotype with the levels of NTN4 in immature LIPT2-AS1$^{-/-}$ cells being significantly upregulated compared to the WT cells. This effect was not seen in siRNA knockdown of LIPT2-AS1. Additionally, as in the siRNA knockdown, SOX2 levels were not affected by the loss of LIPT2-AS1 (Figure 5.9 C).

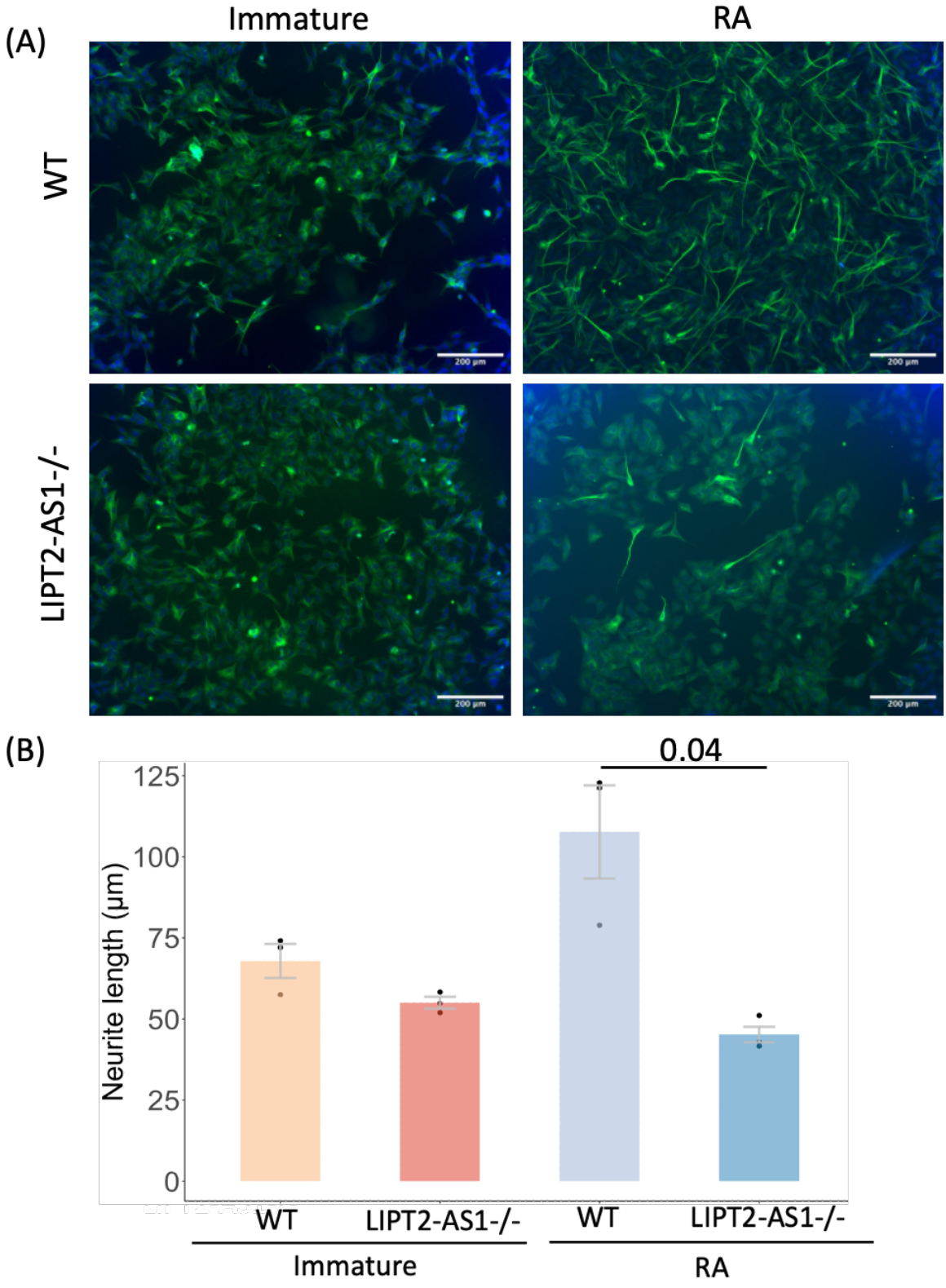


Figure 5.8: CRISPR knockout of LIPT2-AS1 results in significantly reduced neurite outgrowth. *(A) representative images of immature and RA treated WT and CRISPR (LIPT2-AS1$^{-/-}$) SH-SY5Y cells stained for neuronal microtubule protein Tuj1 after 3 days of RA induced differentiation. (B) quantification of neurite length shows LIPT2-AS1$^{-/-}$ cells have significantly shorter neurites compared with WT cells following RA treatment. Two-way paired t-test, N=3, >100 neurites measured per condition.*

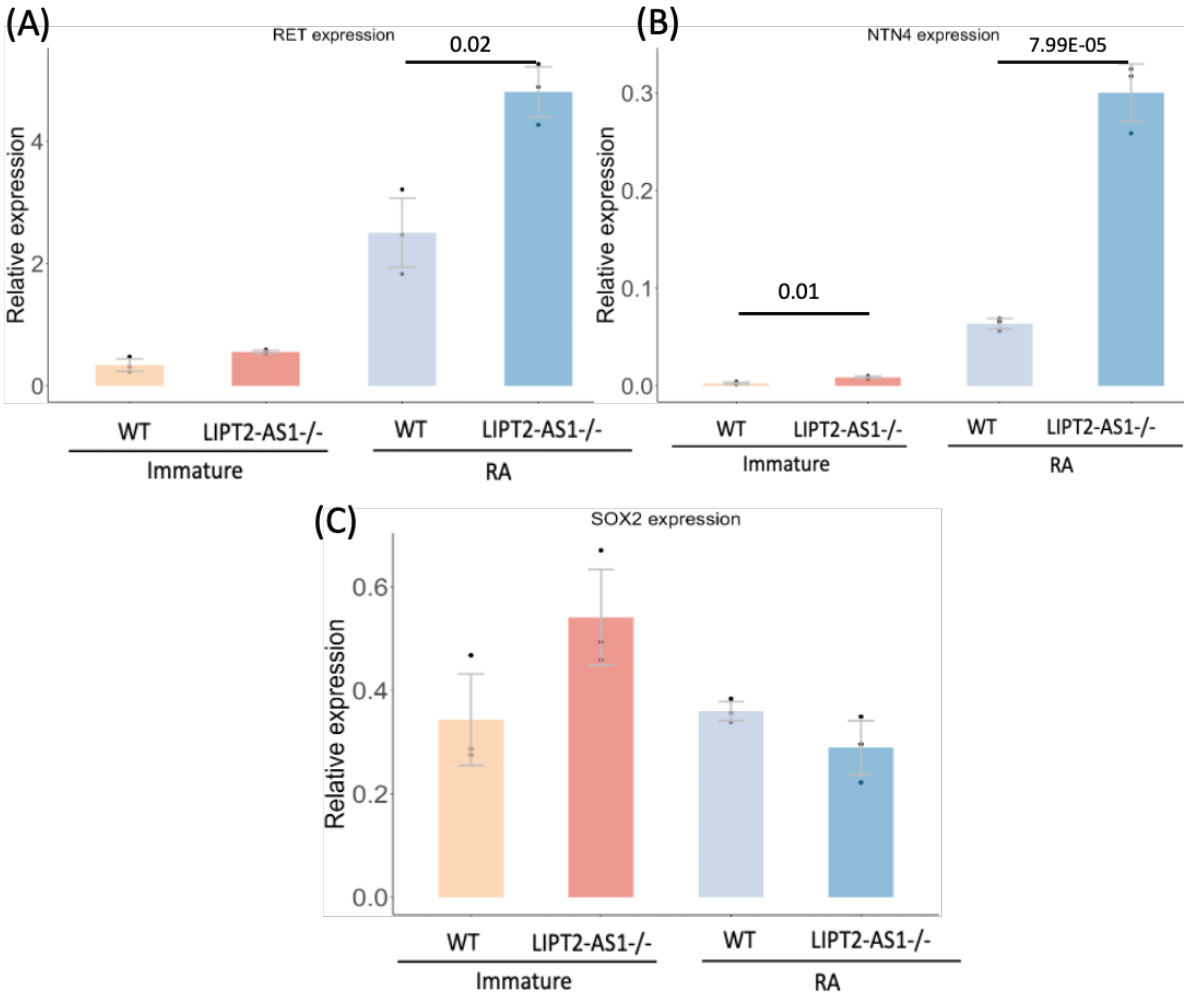


Figure 5.9: CRISPR knockout of LIPT2-AS1 results in significantly increased RNA levels of neuronal marker genes RET and NTN4. *Relative quantification of mature neuronal markers RET and NTN4 and pluripotency marker SOX2 in LIPT2-AS1$^{-/-}$ SH-SY5Y compared to WT SH-SY5Y. LIPT2-AS1 CRISPR resulted in a significant increase of (A) RET and (B) NTN4 in differentiated SH-SY5Y cells. (C) SOX2 levels were not affected by the loss of LIPT2-AS1 protein. Relative quantification was performed housekeeping genes NDUFV2 and HAUS8 and calculated with the ΔCq method. Two-way paired t-test, N=3. Individual biological repeats denoted with dots on bar charts. P value reported on graphs.*

5.4.3 Loss of LIPT2-AS1 protein has no effect on cell growth

The reported doubling time of SH-SY5Y cells varies (Feles et al., 2022; Kovalevich & Langford, 2013). However, after extended periods of culture SH-SY5Y, as is required to generate CRISPR mutants, proliferation tends to slow with doubling time after 63 days of culture being 67.3 h ± 5.8 hour (Feles et al., 2022). To investigate the effect of LIPT2-AS1 loss on SH-SY5Y proliferation, growth curves of LIPT2-AS1$^{-/-}$ and WT cells both without and in the presence of RA were generated. Under both conditions there was no significant difference in growth rate between LIPT2-AS1$^{-/-}$ and WT cells (Figure 5.10).

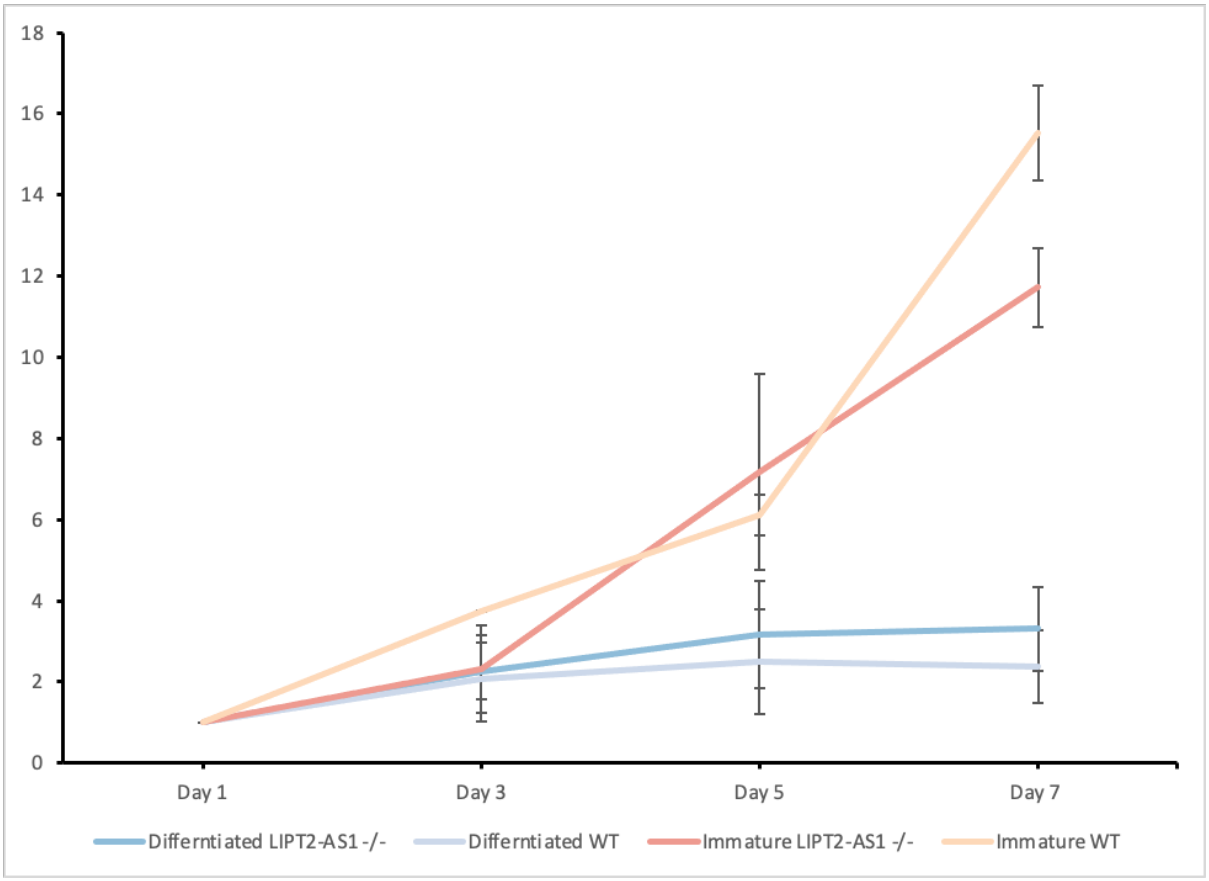


Figure 5.10: Growth curves of LIPT2-AS1$^{-/-}$ and WT SH-SY5Y cells under immature and differentiated culture condition. *Equal numbers of WT and LIPT2-AS1$^{-/-}$ were plated in 24 well plates and allowed to recover for 24 hours before counting to ensure consistent seeding and treatment with RA (Differentiated) or DMSO (Immature). Cells were counted at 48-hour intervals following RA or DMSO treatment. No significant difference was found between LIPT2-AS1$^{-/-}$ and WT cells at any time point. N=3. Error bars represent SEM.*

5.5 Loss of LIPT2-AS1 protein leads to global changes in RNA abundance

siRNA knockdown and frameshift mutation of LIPT2-AS1 protein results in clear phenotypes in SH-SY5Y cells with dysregulated neuronal marker gene expression and attenuated neurite growth. The mechanism by which LIPT2-AS1 protein is functional is currently unknown, however LIPT2-AS1 protein shows co-localisation to the nucleus and is predicted by multiple tools to contain a DNA binding domain (Table 3.2). Therefore, I hypothesise that LIPT2-AS1 may be regulating transcription during neuronal differentiation. To determine the pathways affected by the loss of LIPT2-AS1 protein, RNA-Seq was performed on immature and differentiated LIPT2-AS1$^{-/-}$ and WT SH-SY5Y. All RNA-Seq analysis was completed by Eilidh Ward (University of Leeds, Aspden group).

To visualize variation between our RNA-seq samples a PCA plot was generated. This allowed for the similarities and differences between biological groups to be easily observed. All biological triplicates cluster together separate from other conditions, indicating that all biological replicates are highly similar and show minimal variation. Additionally, all biological conditions cluster separately demonstrating that there are clear RNA abundance differences between all biological conditions (Figure 5.11). One outlier was identified in the immature CRISPR triplicate with one biological repeat localising far to the top left of the PCA. This replicated was included in further analysis as it still clustered closest to the other two immature CRISPR samples than any other group (Figure 5.11).

The largest variation (PC1; 57%) is between immature and differentiated samples. This represents the large change in gene expression that is known to occur throughout SH-SY5Y differentiation with RA (Douka et al., 2021; Korecka et al., 2013). The PCA visualizes this with all immature samples localizing to the left of the PCA plot and all differentiated samples localizing to the right (Figure 5.11). PC2 accounts for 34% of the variation between samples and represents the differential RNA abundance between WT and LIPT2-AS1$^{-/-}$ SH-SY5Y cells. This demonstrates the large variation in RNA abundance between WT and LIPT2-AS1$^{-/-}$ populations, in both immature and RA treated culture condition. A clear distinction can be seen on the PCA plot with all WT samples localising to the bottom of the plot and the CRISPR samples localising at the top (Figure 5.11).

To ensure our RNA-Seq data was accurately reflecting the RNA abundance difference between our biological conditions a correlation was performed between log_2fold change calculated from RNA-Seq or RT-qPCR. Using 7 genes a coefficient of determination of 0.82 was produced showing a strong correlation between RNA-Seq and RT-qPCR (Figure 5.12). Additionally, to ensure the difference in gene expression between WT and LIPT2-AS1$^{-/-}$ cells is not due to different levels of N and S type SH-SY5Y cells multiple marker genes for each cell type were checked. Only one gene out of the 6 investigated showed significant differential expression between WT and LIPT2-AS1$^{-/-}$ cells, and only under differentiated culture condition (Table 5.1). This confirms proportion of S and N-type cells in WT and LIPT2-AS1$^{-/-}$ are similar and unlikely to explain the phenotypically differences observed in the RNA-Seq data.

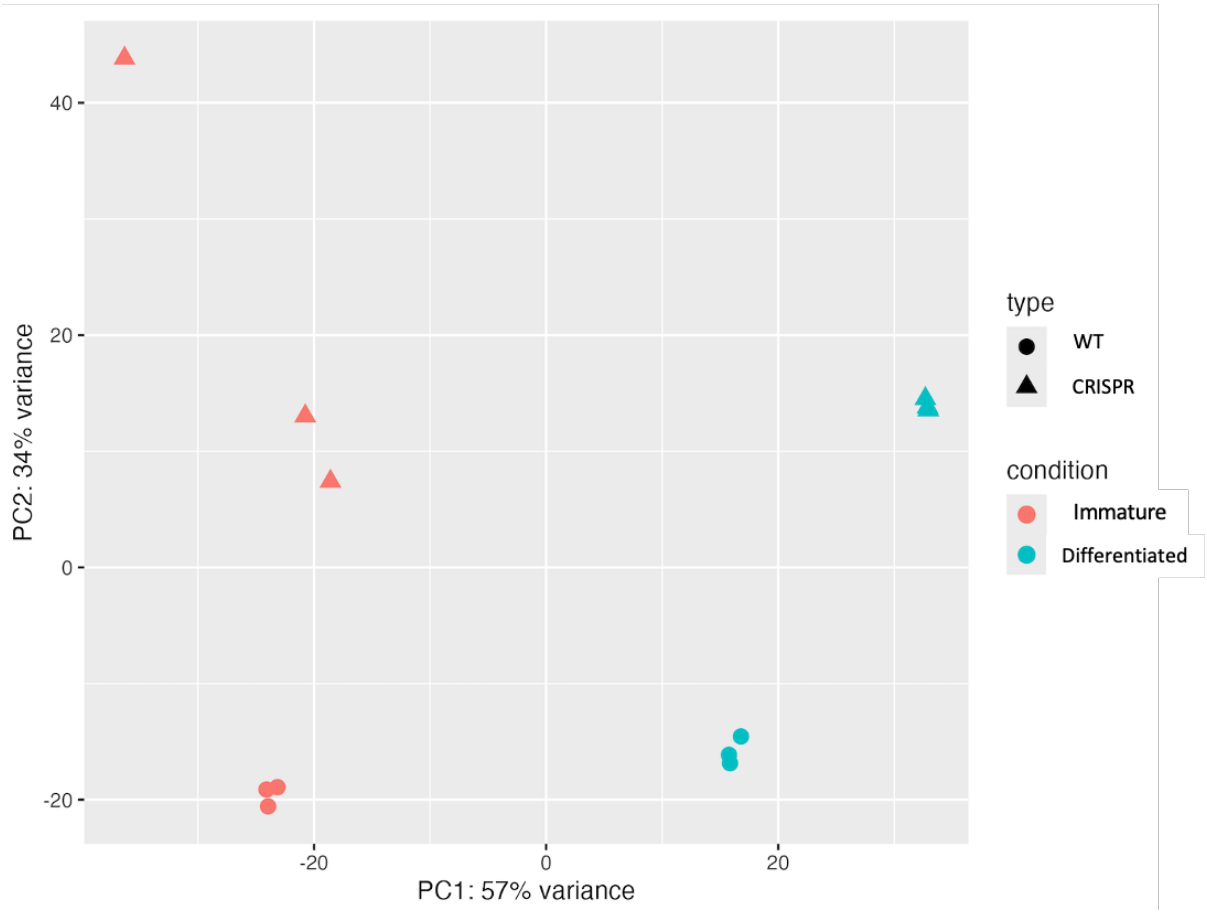


Figure 5.11: PCA plot shows all conditions cluster separately and loss of LIPT2-AS1 protein has similar impact in both immature and differentiated SH-SY5Ys. *PC1 accounts for 57% of the total variance between the data plotted. The immature and differentiated samples can be seen to cluster separately along PC1. PC2 covers 34% of the variance (Y-axis) and shows WT and CRISPR cells clustering apart. All biological repeats (n=3) cluster together and separately from other biological conditions. This indicates variation in gene expression between all samples, but minimal variation between biological repeats. PCA generated by Eilidh Ward.*

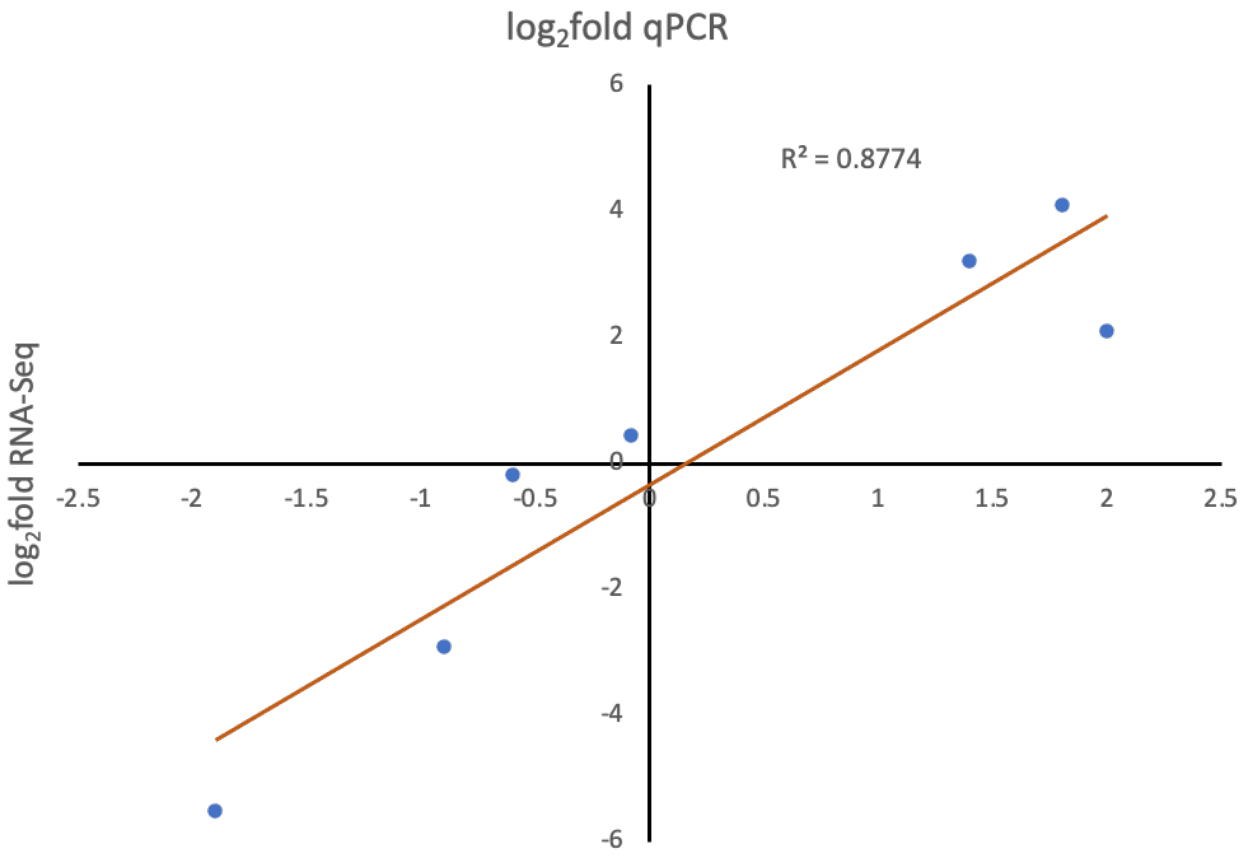


Figure 5.12: Correlation of RT-qPCR data for fold-change for lncRNAs with RNA-Seq analysis. *Genes used were MOXD1-RA, NTN4-RA, NTN4-IM, RET-RA, RET-IM, NEUROG2-RA, NEUROG2-IM with 4 comparing WT RA vs CRIPSR RA and 3 comparing WT IM vs CRISPR IM. R^2=0.8774.*

Gene	N or S type	WT vs CRISPR IM	WT vs CRISPR RA	Reference
LEG1	Marker of S-type	No difference	No difference	(J. Li et al., 2021)
PAI1	Marker of S-type	No difference	No difference	(J. Li et al., 2021)
SPARC	Marker of S-type	No difference	Significantly up	(J. Li et al., 2021)
Tuj1	Marker of N-type	No difference	No difference	(N. Bell et al., 2013)
TRPC1	Marker of N-type	No difference	No difference	(N. Bell et al., 2013)
BCL-2	Marker of N-type	No difference	No difference	(N. Bell et al., 2013)

Table 5.1: Marker genes for N and S-type SH-SY5Y cells and comparison of their RNA levels (determined by RNA-Seq) in WT and LIPT2-AS1$^{-/-}$ cells in both immature and differentiated conditions.

5.5.1 Gene expression changes between LIPT2-AS1$^{-/-}$ and WT SH-SY5Y cells

To determine what genes contribute to the differential RNA abundance between LIPT2-AS1$^{-/-}$ and WT SH-SY5Y differential expression analysis was performed by Eilidh Ward using DESeq2 (Love et al., 2014). Comparisons were made between WT and CRISPR (LIPT2-AS1$^{-/-}$) cells, both when immature and differentiated (Figure 5.13). The loss of LIPT2-AS1 caused global changes in RNA abundance in SH-SY5Y cells with greatest difference seen between differentiated WT and differentiated LIPT2-AS1$^{-/-}$ with 1592 genes showing significant differential expression between the two differentiated conditions (Figure 5.14). Of the differently expressed genes in LIPT2-AS1$^{-/-}$ cells the majority were downregulated (987 genes) compared to 605 genes upregulated.

The loss of LIPT2-AS1 in immature SH-SY5Y cells resulted in a total of 424 genes with significant differential expression between WT and LIPT2-AS1$^{-/-}$ cells (Figure 5.15). Again, more genes were downregulated than upregulated in LIPT2-AS1$^{-/-}$ cells with 275 genes being significantly downregulated compared to WT and 149 genes upregulated. Over half of the differentially expressed genes between WT and LIPT2-AS1$^{-/-}$ cells in immature conditions are also differentially expressed in differentiated conditions (Figure 5.16). This large overlap indicates the effect of losing LIPT2-AS1 protein is similar in both immature and differentiated SH-SY5Y cells.

5.3.2 Loss of LIPT2-AS1 protein results in the downregulation of neuronal marker genes

To identify the processes disrupted by LIPT2-AS1 knockdown, I performed GO term analysis using g:Profiler (Reimand et al., 2007) on genes significantly differentially expressed over +/- 1.5 $\log_2$fold between WT and LIPT2-AS1$^{-/-}$ cells in both immature and differentiated conditions. Additionally, GO terms were calculated for the genes that were consistently differentially expressed in both immature and differentiated conditions.

GO term analysis of genes downregulated in LIPT2-AS1$^{-/-}$ cells compared to WT SH-SY5Y in both immature and differentiated conditions are enriched for terms relating to neuronal differentiation and mature neurons (Figure 5.17 A,C). GO terms for downregulated genes in differentiated LIPT2-AS1$^{-/-}$ cells include terms related to mature neuronal function, for example “Synapse”, “Synaptic signalling” and “Main axon” (Figure 5.17 A). Similarly, GO

terms for downregulated genes in immature LIPT2-AS1$^{-/-}$ cells include the neuronal terms "Nervous system development", "synapse" and "Neuron projection (Figure 5.17 B). Additionally, when repeating GO term analysis on the 212 genes that are consistently downregulated in both immature and differentiated LIPT2-AS1$^{-/-}$ the terms are again enriched for neuronal terms with the top 3 GO terms calculated being "Synapse", "Neuron projection" and "Axon" (Figure 5.18 B). This demonstrates that without LIPT2-AS1 protein, RNA abundance of genes involved in neuronal processes are significantly reduced indicating that the loss of LIPT2-AS1 protein prevents SH-SY5Y cells from differentiating down a neuronal lineage.

The most significantly downregulated genes in LIPT2-AS1$^{-/-}$ cells include synaptic genes (Synapsin I, Synaptotagmin 7), genes involved in neurite and axonal projection (Stathmin 4, ISLR2, Chromogranin A) as well as neuronal specific transcription factors (ACTL6B) and genes required for neurotransmitter production (Dopamine beta-hydroxylase).

5.3.1 Loss of LIPT2-AS1 protein results in the upregulation of connective tissue genes

GO terms calculated from genes upregulated in immature and differentiated LIPT2-AS1$^{-/-}$ show a less clear pattern than GO terms based of downregulated genes. In differentiated LIPT2-AS1$^{-/-}$ cells the top 10 terms included "extracellular region", "collagen-containing extracellular matrix" and "developmental process" (Figure 5.17 C). Similarly in immature LIPT2-AS1$^{-/-}$ GO terms related to development are also present including "developmental process" and "system development" also the term "extracellular region" is also significantly enriched (Figure 5.17 D). Performing GO term analysis on the 90 genes that are consistently upregulated in both immature and differentiated LIPT2-AS1$^{-/-}$ produced the terms including "anatomical structure morphogenesis", "developmental process" and "tube development" (Figure 5.18 A). Consistently upregulated genes in both immature and differentiated LIPT2-AS1$^{-/-}$ cells were also analysed for GO terms. The top 3 terms being "anatomical structure morphogenesis", "multicellular organism development" and "anatomical structure development" (Figure 5.18 B).

Looking at specific genes, TBX18 was the second and third most highly upregulated gene in differentiate and immature LIPT2-AS1$^{-/-}$ cells, respectively. TBX18 is a highly conserved

developmental transcription factor involved in the development of multiple tissues and organs, particularly associated with cardiac development. Cell expressing TBX18 give rise to cardiac fibroblasts and coronary smooth muscle cells (C.-L. Cai et al., 2008). GPX8 is also highly upregulated in LIPT2-AS1$^{-/-}$ cells both in immature and differentiated conditions. GPX8 encodes an oxidoreductase and is known to contribute to glioma pathogenesis by driving epithelial–mesenchymal transformation (S. Li et al., 2022). Additionally, multiple genes involved in extracellular matrix are among the most upregulated genes in LIPT2-AS1$^{-/-}$ cells. This includes collagen genes COL5A1 and COL3A1 and extracellular matrix component Fibrillin 1.

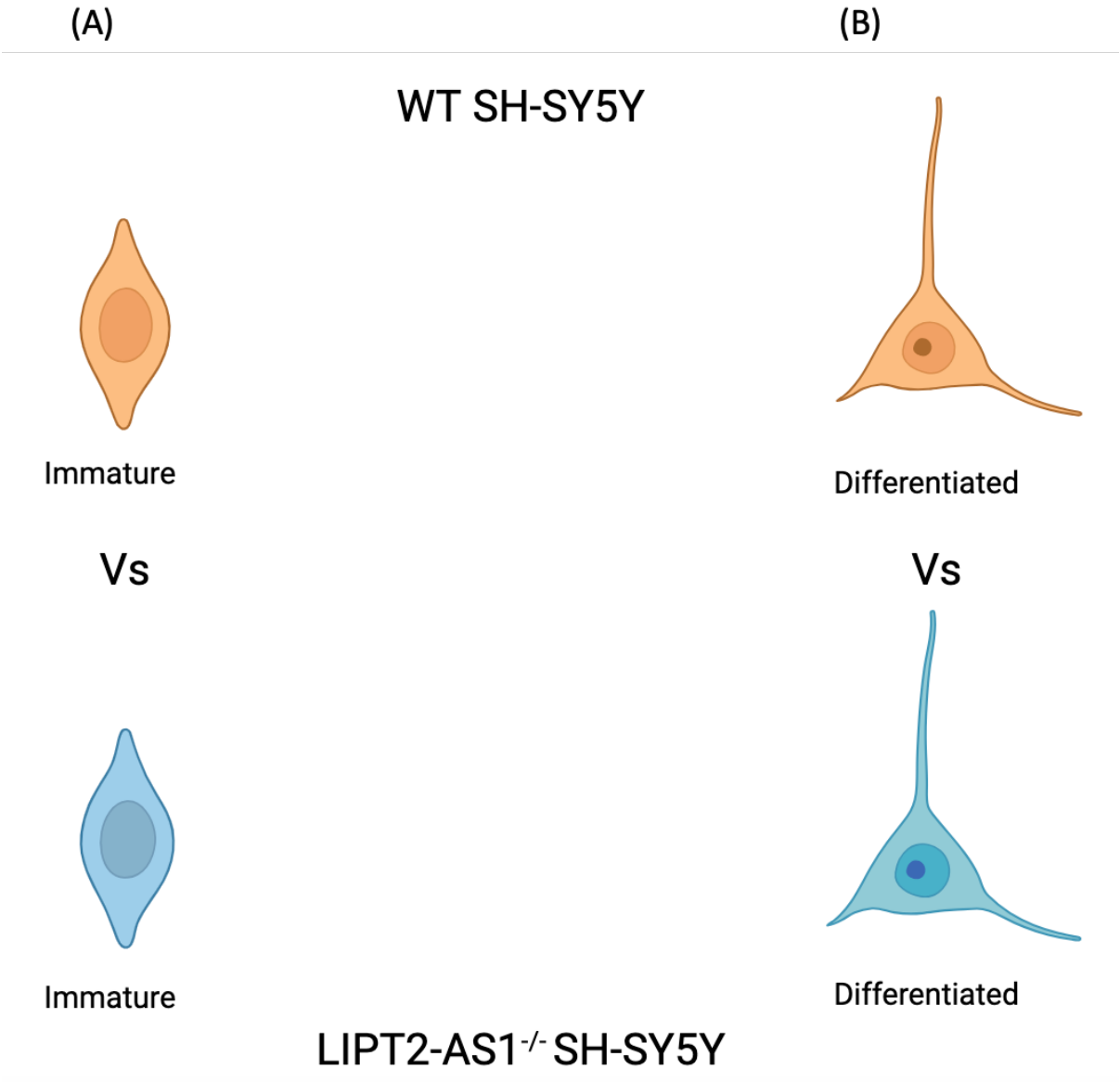


Figure 5.13: Diagram showing the two comparisons investigated through RNA-Seq data analysis. *(A) immature WT SH-SY5Y cells compared to immature LIPT2-AS1$^{-/-}$ SH-SY5Y cells. (B) differentiated WT SH-SY5Y cells compared to differentiated LIPT2-AS1$^{-/-}$ SH-SY5Y cells.*

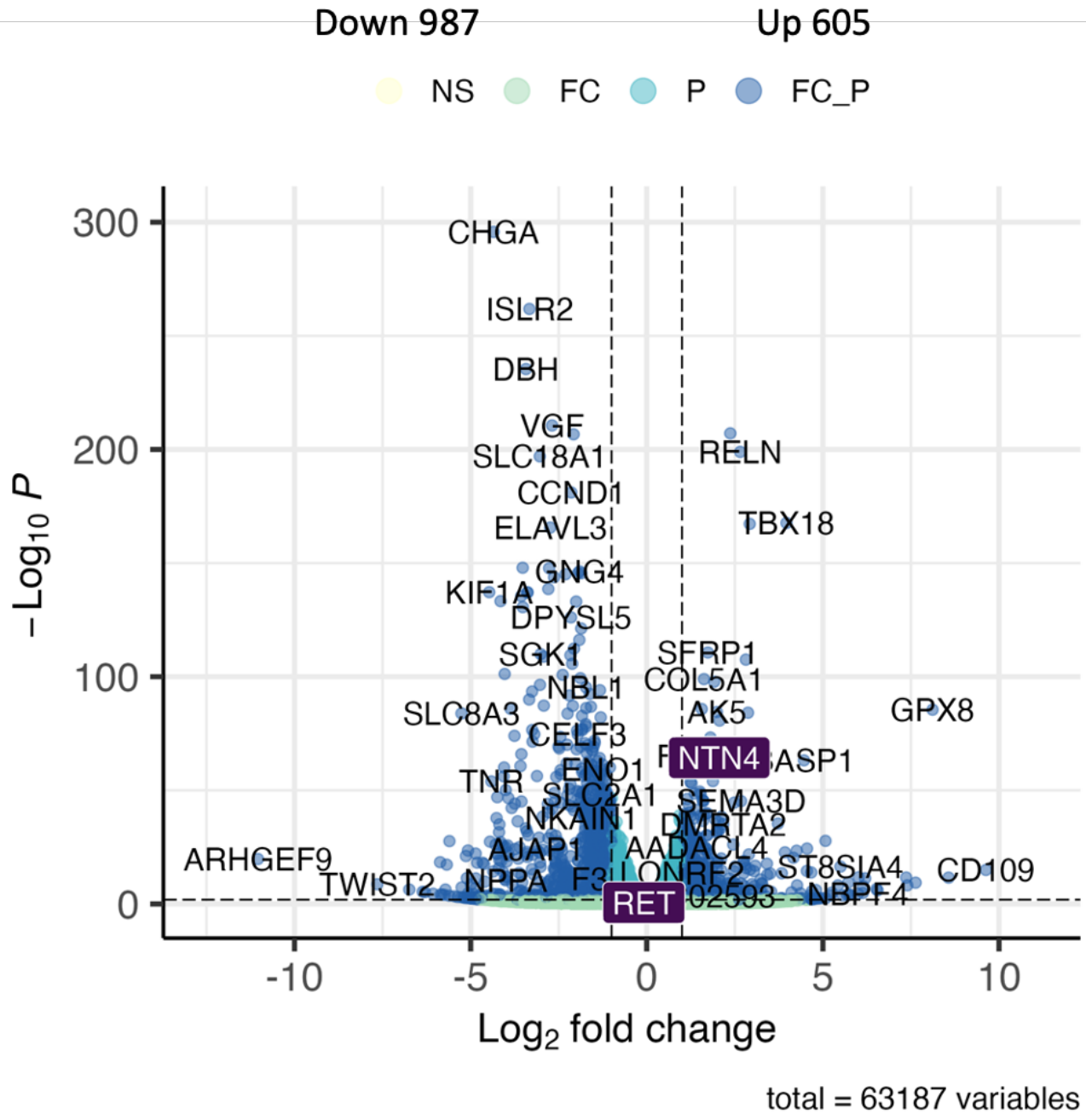


Figure 5.14: Volcano plot showing differently expressed genes between differentiated WT and differentiated LIPT2-AS1$^{-/-}$ SH-SY5Y cells. There are 605 significantly upregulated genes (Log2fold >1) and 987 significantly downregulated (Log_2fold <-1) genes (padj <0.05). NS= genes below both p value and Log_2fold change cut off. FC= genes with above Log_2fold cut off but not reaching p value cut off. P= genes with p value <0.05 but Log_2fold less than $\pm$ 1. FC_P= Log_2fold greater than $\pm$ 1 and p value <0.05. Plot generate by Eilidh Ward.

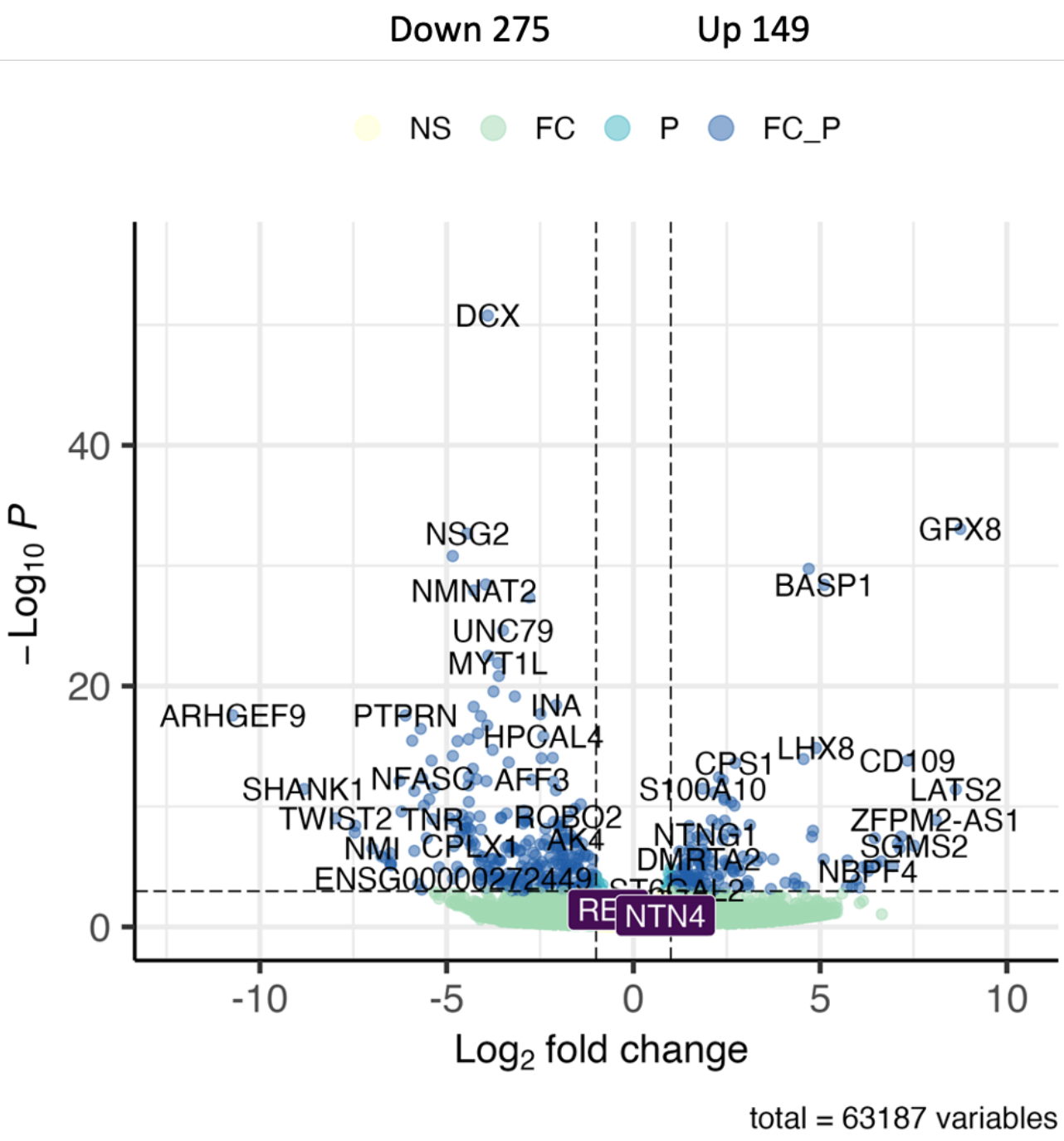


Figure 5.15: Volcano plot showing differently expressed genes between immature WT and immature LIPT2-AS1$^{-/-}$ SH-SY5Y cells. There are 149 significantly upregulated genes (Log_2fold >1) and 275 significantly downregulated (Log_2fold <-1) genes (padj <0.05). NS= genes below both p value and Log_2fold change cut off. FC= genes with above Log_2fold cut off but not reaching p value cut off. P= genes with p value <0.05 but Log_2fold less than ± 1. FC_P= Log_2fold greater than ± 1 and p value <0.05. Plot generate by Eilidh Ward.

(A) Overlap of genes downregulated in differentiated (RA) and immature (IM) LIPT2-AS1-/- SH-SY5Y cells

(B) Overlap of genes upregulated in differentiated (RA) and immature (IM) LIPT2-AS1-/- SH-SY5Y cells

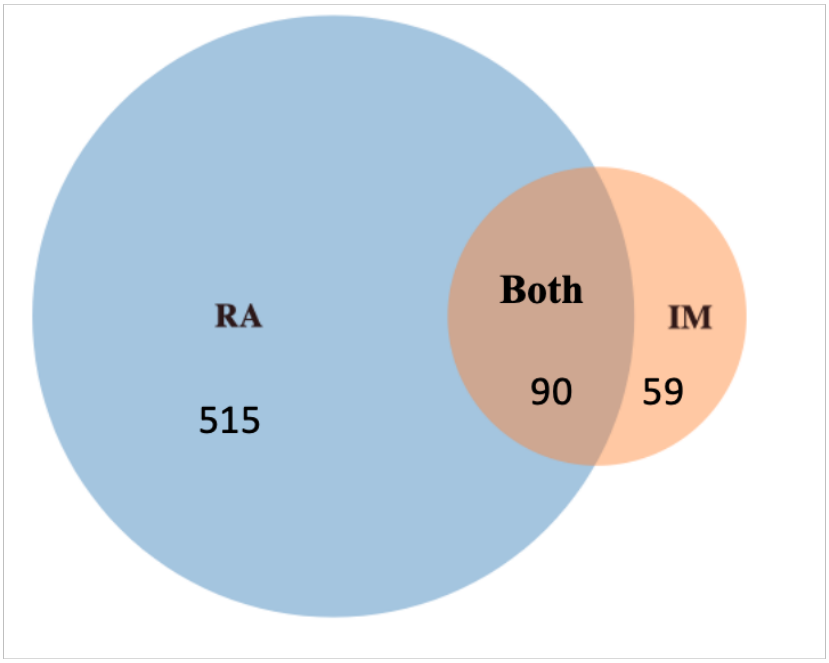


Figure 5.16: Venn diagrams showing overlap between consistently differentially expressed genes between LIPT2-AS1$^{-/-}$ and WT SH-SY5Y in both immature (IM) and differentiated (RA) conditions. *(A) 212 genes are downregulated in LIPT2-AS1$^{-/-}$ SH-SY5Y in both immature and differentiated condition. (B) 90 genes are upregulated in LIPT2-AS1$^{-/-}$ SH-SY5Y cells compared to WT SH-SY5Y in both immature and differentiated conditions.*

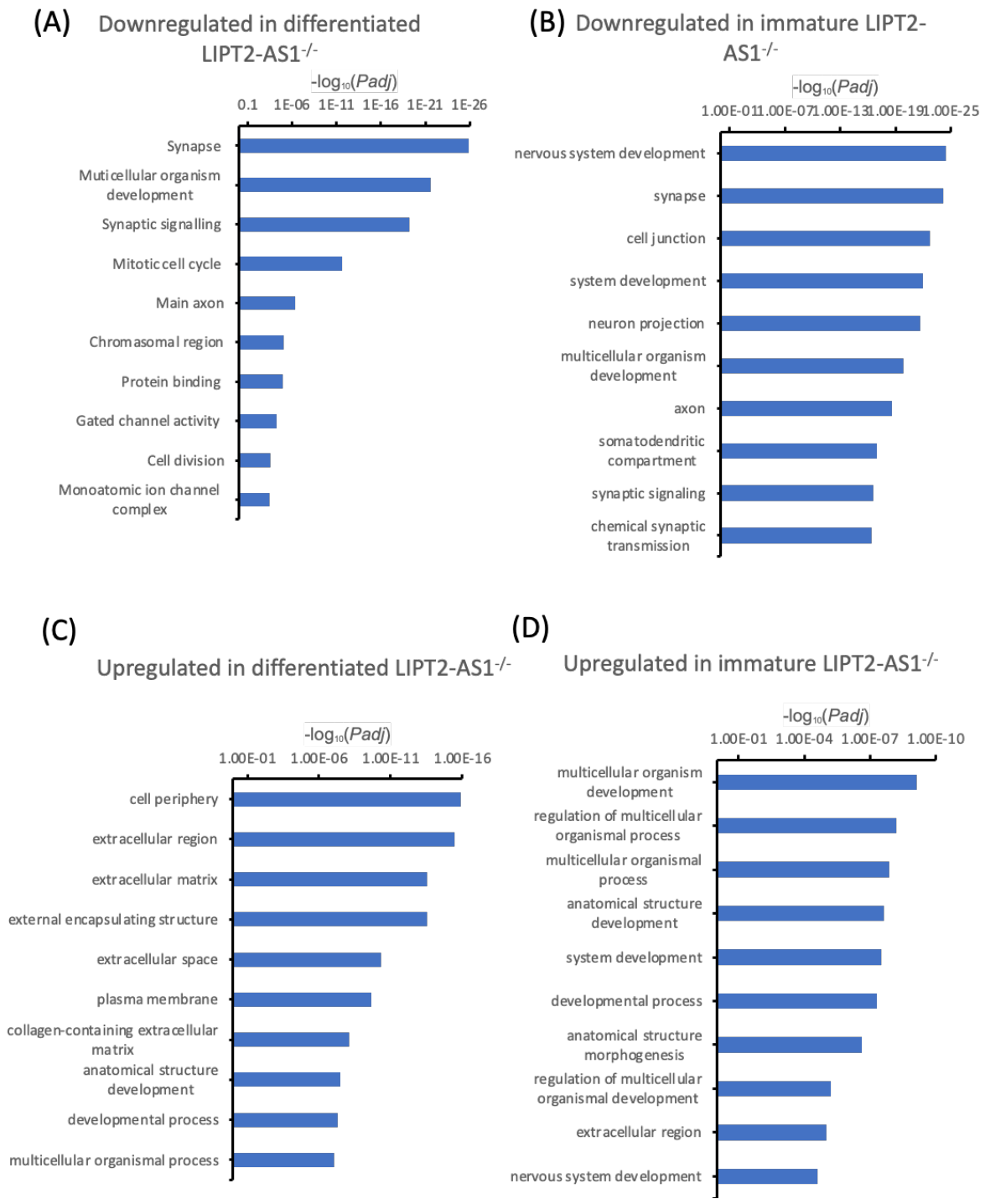


Figure 5.17: Top 10 GO terms for genes significantly differentially expressed in LIPT2-AS1$^{-/-}$ cells compared to WT SH-SY5Y, both in differentiated (RA) and immature (IM) conditions. *GO terms were calculated using g:Profiler for all genes showing significant differential expression over +/-1.5 log2fold between WT and LIPT2-AS1$^{-/-}$. (A) GO terms for downregulated genes in differentiated LIPT2-AS1$^{-/-}$ SH-SY5Y cells. (B) GO terms for downregulated genes in immature LIPT2-AS1$^{-/-}$ SH-SY5Y cells. (C) GO terms for upregulated genes in differentiated LIPT2-AS1$^{-/-}$ SH-SY5Y cells. (D) GO terms for upregulated genes in immature LIPT2-AS1$^{-/-}$ SH-SY5Y cells.*

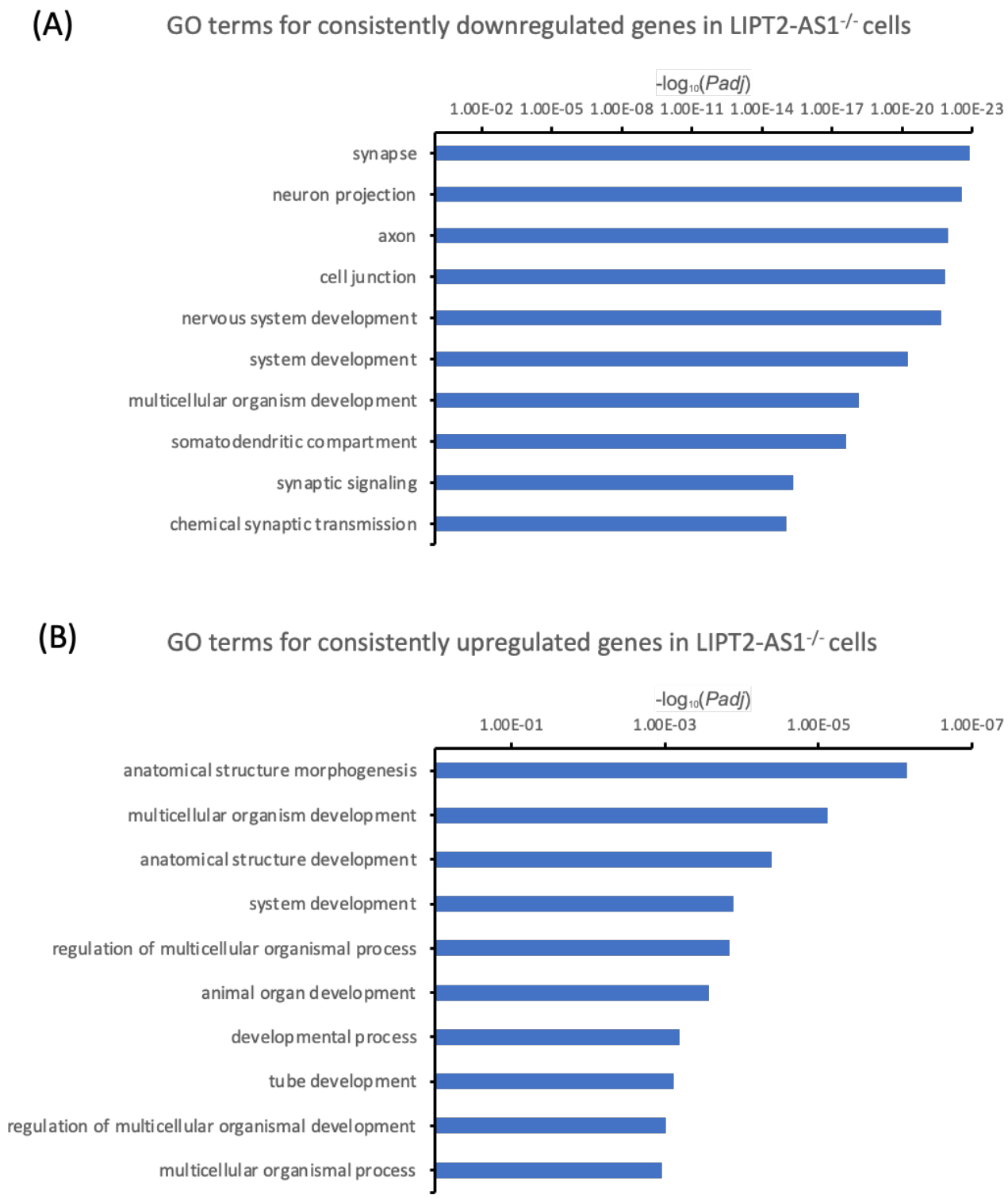


Figure 5.18: Top 10 GO terms for genes consistently significantly differentially expressed in both differentiated (RA) and immature (IM) conditions. *GO terms were calculated using g:Profiler for genes showing consistently (A) downregulated or (B) upregulated over +/-1.5 log_2fold between WT and LIPT2-AS1$^{-/-}$ in both immature and differentiated conditions.*

5.4 Identification of LIPT2-AS1 protein interactions

CRISPR editing experiments and RNA-Seq have demonstrated LIPT2-AS1 protein is essential for normal neuronal differentiation and loss of LIPT2-AS1 protein results in global dysregulation of SH-SY5Y RNA levels. However, the mechanism of action for LIPT2-AS1 protein functions is currently unknown beyond that LIPT2-AS1 protein localises to the nucleus in discreet puncta, as shown by IF experiments and that LIPT2-AS1 is predicted to contain a DNA binding domain. Therefore, it was expected that LIPT2-AS1 would interact with nuclear proteins.

To begin elucidating the molecular function by which LIPT2-AS1 acts a preliminary pull-down experiment was performed to identify the interaction partners of LIPT2-AS1. FLAG tagged LIPT2-AS1 was overexpressed in immature SH-SY5Y cells and pulled down with anti-FLAG beads from whole cell lysate. Label free quantification MS was used to identify proteins interacting with LIPT2-AS1. As a preliminary experiment a n=1 was performed. In parallel a 3X FLAG alone pulldown was performed as a negative control. To identify proteins that specifically bound to LIPT2-AS1 protein-FLAG and not to FLAG alone enrichment analysis was performed. 91 proteins were found to be enriched $\geq$1.5 log_2fold to LIPT-AS1-protein.

5.4.1 String analysis of proteins pulled-down with LIPT2-AS1-FLAG

String network (STRING v12) generated from the 91 proteins enriched 1.5 log_2fold over FLAG alone showed a large cluster containing multiple histone proteins, which are found in the nucleosome, such as Histone H2A type 1-D (Figure 5.19). Additionally, in this cluster is RPABC3, a core subunit of RNA polymerase II, which is required in the transcription of DNA to mRNA. Multiple transcription factors were enriched in LIPT2-AS1-FLAG pulldown over control including E2F7, a transcription factor involved in may processes including spouting angiogenesis (Weijts et al., 2012)and MYEF2, a transcriptional repressor of the myelin basic protein gene (MBP) (Muralidharan et al., 1997).

Also interacting with LIPT2-AS1 are proteins involved in RNA processing including core components of the spliceosome, PRPF4, PPIH, and CDC5L. Additionally, HNRNPH3, a heterogeneous nuclear ribonucleoproteins involved in pre-mRNA processing was also enriched in LIPT2-AS1 pulldowns. Multiple ribosomal proteins are also enriched in the pulldown of LIPT2-AS1 including MRPL21.

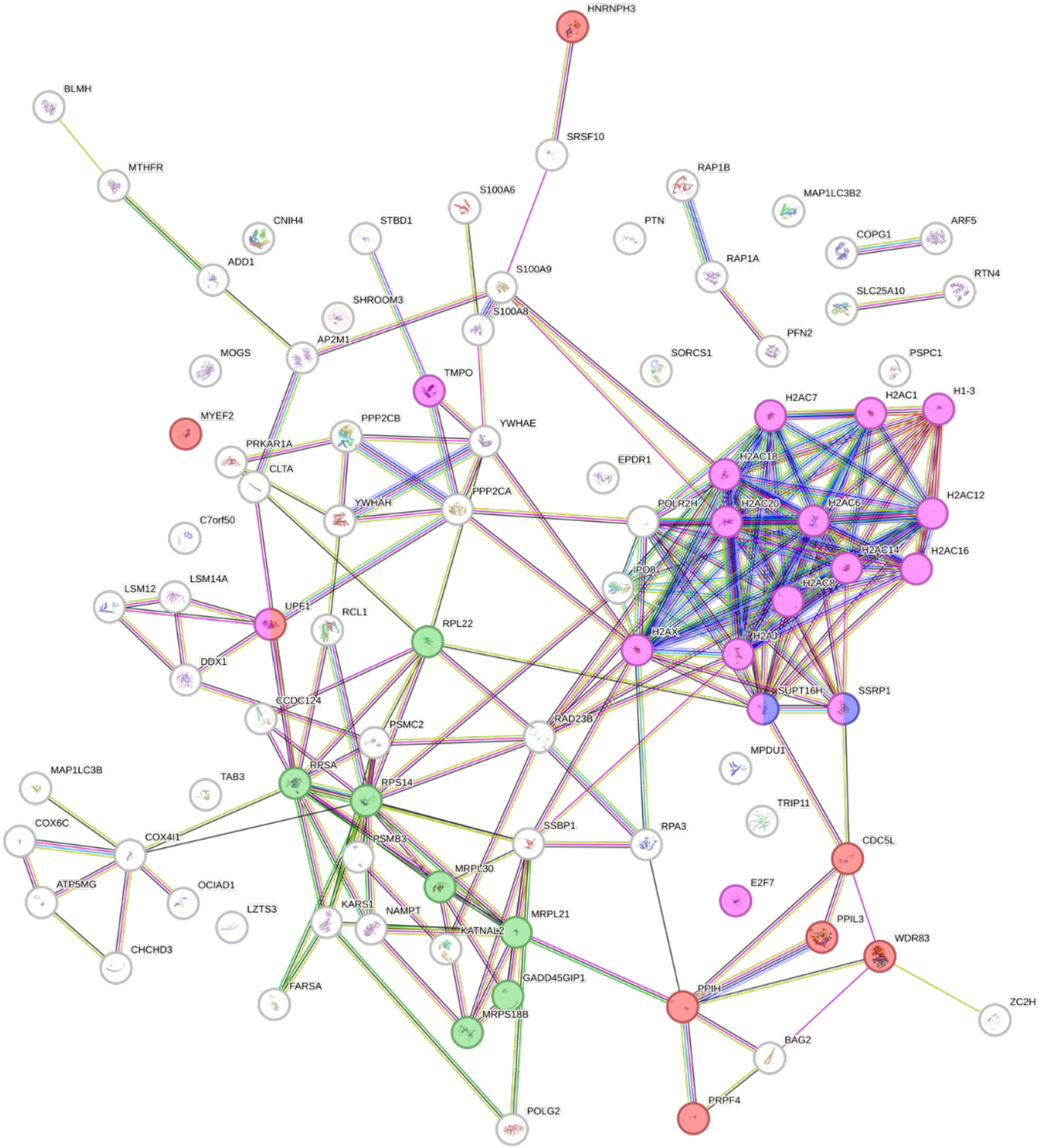


Figure 5.19: String analysis of proteins pulled-down with LIPT2-AS1-FLAG. *The network is made from proteins enriched 1.5 log_2fold in LIPT2-AS1-FLAG pulldowns over FLAG alone. Proteins highlighted in red are components to the spliceosome. Proteins in blue are members of the FACT complex and pink proteins are part of chromatin. Multiple ribosomal proteins can also be seen highlighted in green. Analysis performed in String* (Szklarczyk et al., 2023)

5.4.2 GO term analysis of proteins pulled-down with LIPT2-AS1-FLAG

To understand the potential function of LIPT2-AS1 protein through its interacting partners, GO terms analysis was performed on proteins enriched 1.5 log_2-fold in LIPT2-AS1 pulldown compared to FLAG alone pulldown. These proteins are likely interaction partners of LIPT2-AS1 and therefore could indicate what biological processes LIPT2-AS1 is involved in.

GO molecular function identified the most significant scores as “Structural constituent of chromatin” and “protein heterodimerization activity” suggesting LIPT2-AS1 is part of nuclear protein complexes (Figure 5.20). Also included in the list of significant terms were “DNA binding” and “nucleic acid binding”. Additionally, GO cellular compartment calculated by the String program identified LIPT2-AS1 to have its most strong association with U4/U6 snRNP, FACT complex and nucleosome (Table 5.2). This data indicates that LIPT2-AS1 interacts with multiple nuclear protein complexes involved in pre-RNA splicing and transcription.

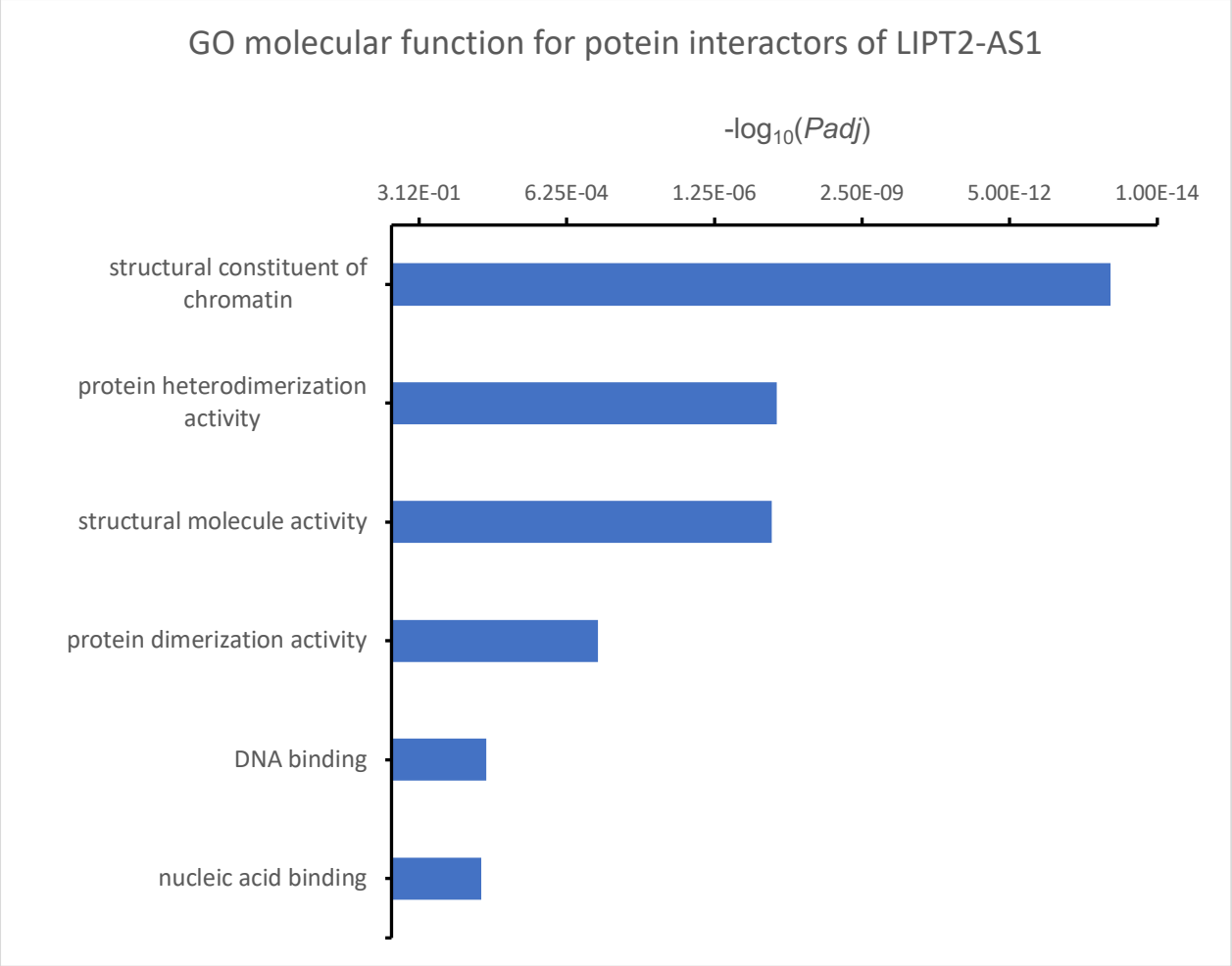


Figure 5.20: GO molecular function for proteins pulled-down with LIPT2-AS1-FLAG. *GO terms were calculated for proteins enriched ≥1.5 log_2fold in LIPT2-AS1-FLAG pulldowns over FLAG alone pulldowns using g:Profiler. All significant (p<0.05) GO terms are shown.*

GO term cellular compartment	Strength
U4/U6 snRNP	2.34
FACT complex	2.34
Nucleosome	1.29
protein-DNA complex	1.13

Table 5.2: Summarising the GO term cellular compartments for protein identified with LIPT2-AS1 pulldown with a strength >1, calculated by String. *Strength is calculated by* Log_{10}*(observed / expected) with observer describing the number of proteins in your network that are annotated with a given term and expected being the number of proteins that would be annotated with the same term in a random sized matched network* (Szklarczyk et al., 2023) .

5.5 Discussion

5.5.1 LIPT2-AS1 protein, not its RNA, functions in neuronal differentiation

Both translated lncRNAs and "canonical" mRNAs have been identified as dual functional genes, which shows how far understanding of gene function has come since the "one gene-one polypeptide hypothesis" (Beadle & Tatum, 1941). This chapter focused on determining if LIPT2-AS1 RNA, protein, or combination of both were responsible for the phenotype seen in the siRNA knockdown screen (Figure 4.10/4.11). CRISPR-Cas9 was used to introduce a homozygous frameshift mutation that prevented LIPT2-AS1 protein expression but left the RNA intact. Therefore, the phenotype of losing LIPT2-AS1 protein alone could be elucidated.

LIPT2-AS1$^{-/-}$ cells exhibited a very similar phenotype to SH-SY5Y cells treated with siRNA against LIPT2-AS1, with both knockdown methods showing consistent effects on both neuronal marker gene expression and neurite outgrowth. This data indicates the phenotype seen with knockdown of LIPT2-AS1 is entirely due to the loss of LIPT2-AS1 protein, and not due to the inhibition of an RNA function.

However, there is the possibility that the single nucleotide mutation introduced into LIPT2-AS1 is disrupting the RNAs ability to function through an unknown mechanism as LIPT2-AS1 currently has no known lncRNA function. RNA structure prediction by RNAFold did show variation between WT and mutant LIPT2-AS1 RNA structure (Figure 5.7). To overcome this issue a rescue experiment could be performed. LIPT2-AS1 protein can be re-introduced to LIPT2-AS1$^{-/-}$ cells, either through transfection of a synthetic protein or plasmid expressing just LIPT2-AS1s ORF. Neurite measuring and RT-qPCR could then be used to establish if the protein can reverse the knockdown phenotype.

5.5.2 Loss of LIPT2-AS1 protein dysregulates SH-SY5Y gene expression

RNA-Seq of LIPT2-AS1$^{-/-}$ cells clearly shows loss of LIPT2-AS1 protein results in a global downregulation of neuronal genes and LIPT2-AS1$^{-/-}$ are unable to grow neurites in response to RA treatment. This indicates that LIPT2-AS1 protein is required for SH-SY5Y maintenance as a neuronal precursor-like cell and for SH-SY5Y differentiation towards a mature neuronal lineage following RA treatment.

However, it is less clear what lineage SH-SY5Y progress down with the loss of LIPT2-AS1. An upregulation of genes and GO terms associated with connective tissue and extracellular

matrix formation is seen in LIPT2-AS1$^{-/-}$ compared to WT cells. Additionally, multiple other genes involved in mesenchymal cell formation are upregulated with LIPT2-AS1 knockdown. TBX18, a cardiac-specific transcription factor expressed in mesenchymal stem cells is highly upregulated in both immature and differentiated LIPT2-AS1$^{-/-}$ cells. TBX18 is and thought to be involved in the epithelial-to-mesenchymal transition of epicardial cells and high expression of TBX18 is associated with the recruitment and differentiation of cells at the cardiac venous pole during embryonic development of the heart (Takeichi et al., 2013). GPX8 is also highly upregulated in both immature and differentiated LIPT2-AS1$^{-/-}$ cells. GPX8 is an oxidoreductase and its expression is positively correlated with mesenchymal markers and negatively correlated with pro-neural markers in primary GBM samples. Additionally, GPX8 is thought to contribute to epithelial–mesenchymal transition in glioma (S. Li et al., 2022). Furthermore, multiple key genes including FAP, FBN1, FN1 and THBS1 known to be involved in glioblastoma epithelial–mesenchymal transition are significantly upregulated in LIPT2-AS1$^{-/-}$ cells compared to WT cells (Cheng et al., 2012; Mikheeva et al., 2010). This suggests that the loss of LIPT2-AS1 may be driving SH-SY5Y from a neuronal lineage into a mesenchymal phenotype.

LIPT2-AS1 knockdown results in a significant upregulation of NTN4, a gene involved in axonal guidance, which I originally interpreted as an indication that SH-SY5Y cells were becoming more neuronal with LIPT2-AS1 knockdown. However, NTN4 is also expressed by multiple other cell types and required for multiple biological processes including angiogenesis and migration (Q. Liu et al., 2019; Villanueva et al., 2019). Similarly, RET, while required for neuronal navigation and differentiation is also involved in development of organs and tissues derived from the neural crest, including sympathetic, parasympathetic, and enteric nervous systems, the kidney, and spermatogenesis(de Groot et al., 2006) . Therefore, this upregulation of NTN4 could reflect LIPT2-AS1$^{-/-}$ cells differentiation towards a non-neuronal phenotype. In future screens of SH-SY5Y differentiation more markers covering a range of neural crest derivative cell types should be used.

5.5.3 LIPT2-AS1 protein interacts with nuclear proteins involved in splicing and transcription

To identify potential interaction partners of LIPT2-AS1, LIPT2-AS1-FLAG was overexpressed in SH-SY5Y cells and pulled down with magnetic beads coated in antibodies against FLAG. Proteins interacting with LIPT2-AS1-FLAG were then identified by label free quantitative MS.

91 proteins were identified as enriched $\geq$1.5 log_2fold in LIPT2-AS1-FLAG pulldowns over FLAG alone pulldowns. Out of the 91, 9 proteins (including PPIH and PRPF4) are key components of the spliceosome, the large RNA-protein complex that catalyses the removal of introns from pre-mRNA. Additionally, SUPT16H and SSRP1 were identified. Together these proteins make up the FACT (FAcilitates Chromatin Transcription) complex. FACT is a highly conserved histone chaperone complex that destabilises nucleosomes allowing for multiple processes that require DNA as a template including mRNA transcription, DNA replication and DNA repair to occur (Formosa & Winston, 2020). This data suggest LIPT2-AS1 protein does function with the nucleus as expected from its nuclear localisation and predicted DNA binding domain and potentially has a role in splicing and or transcriptional control. However, as only one replicate was performed for the pull-down no statistical analysis for significance can be performed. Therefore, it is likely not all the 91 proteins identified are genuine interaction partners. For example, multiple mitochondrial ribosomal proteins were enriched in the LIPT2-AS1 pulldown, which may represent experimental noise. Proteins identified by MS as interacting with LIPT2-AS1 need to be validated by future co-immunoprecipitation experiments.

5.6 Conclusion

In this chapter I have demonstrated that the protein product of LIPT2-AS1, and not the RNA, is the functional element of the LIPT2-AS1 gene in the context of neuronal differentiation. Loss of LIPT2-AS1 protein, achieved through a CRISPR generated frameshift mutation, drives SH-SY5Y cells away from a neuronal phenotype, as characterised by a global downregulation of neuronal genes and attenuated neurite growth. Additionally, loss of LIPT2-AS1 protein may induce SH-SY5Y to undergo a mesenchymal transition. Preliminary evidence suggests LIPT2-AS1 functions in the nucleus and interacts with multiple chromatin, splicing, transcription factors and transcriptional machinery.

Chapter 6: General Discussion

6.1 General discussion

Ribosome profiling has furthered our understanding of the human genome by revealing extensive translation outside of annotated coding sequences. These sequences are termed non-canonical ORFs (Prensner et al., 2023). While estimates about the number of translated non-canonical ORFs in humans varies up to 30-fold, a conservative estimate has been put at ~7000 (Prensner et al., 2023). If accurate, this would expand the known proteome by ~30%. Considering that ~40% of all human lncRNAs are expressed in the brain, non-canonical ORFs within lncRNAs could be the source of hundreds of novel brain specific proteins. Examples of novel proteins functioning in neuronal development (pTUNAR, ENSG00000205704 (N. Lin et al., 2014; Senís et al., 2021)), long-term potentiation (Pants (Kragness et al., 2022)) and Alzheimer's disease (FLJ 33706 and SHMOOSE (C.-Y. Li et al., 2010; Miller et al., 2023)) have already been identified. Still, with the rapid and extensive identification of non-canonical ORFs, characterisation of novel protein products is lagging behind their identification. In this thesis, I have characterised translated non-canonical ORFs from one source, annotated lncRNAs. LncRNAs are a heterogenous population of RNAs, historically thought to be nuclear retained and non-coding in nature. However, it is now accepted that cytoplasmic lncRNAs are involved in many neuronal functions (Grinman et al., 2021; Rani et al., 2016) and a proportion can be translated to produce functional proteins (Brunet et al., 2020; Mudge et al., 2022; Orr et al., 2020).

It is likely that functional proteins identified from lncRNAs will be from lncRNA genes with a range of distinct characteristics including dual function lncRNAs (lncRNAs with both RNA function and (micro)protein coding potential), lncRNAs that encode microproteins but have no RNA function, and misannotated mRNAs that have no RNA function and encode proteins > 100 amino acids long. Additionally, the discovery of canonical genes possessing dual-functional, such as p53, further blurs the lines between mRNA and lncRNA classifications (Candeias et al., 2008). Regardless of nomenclature, the exclusion of non-canonical ORFs identified by Ribo-Seq from genome annotations is detrimental to their study as accurate annotations underpin large-scale genomic projects and human variant interpretation (Mudge et al., 2022). Therefore, the identification and characterisation of novel functional proteins from 'lncRNAs' involved in neuronal differentiation and disease will both be beneficial, both in: (1) furthering our understanding of neuronal development and disease

processes and potentially identifying novel proteins with clinical relevance, and (2) demonstrating that ORFs from genes not annotated as mRNAs can produce functional proteins, hopefully leading to non-canonical ORFs discovered by Ribo-Seq being more readily incorporated into genome annotations (Mudge et al., 2022).

6.2 GIHCG protein may contribute to neuronal mitochondrial dynamics

GIHCG has been extensively characterised as a lncRNA, with proposed roles in a range of cancers (Jiang et al., 2019; X. Zhang, Mao, et al., 2019; S.-Y. Zhu et al., 2022). Poly-Ribo-Seq identified a novel sORF within GIHCG encoding a 64 amino acid peptide translated in differentiated SH-SY5Ys (Douka et al., 2021). Here I have discovered that that GIHCG-peptide is stable enough to be detected by Western blot and exhibits precise localisation to the mitochondria of both immature and differentiated SH-SY5Y cells (Figure 3.12). Overexpressing GIHCG-peptide-FLAG in SH-SY5Y cells resulted in significantly shorter mitochondrial branches than controls (Figure 4.14). Mitochondrial dynamics, the balance between fusion and fission, changes throughout neuronal differentiation (Khacho et al., 2016). In uncommitted NPCs mitochondria exhibit an elongated morphology. However, once committed mitochondria become fragmented before fusing again to form elongated structures in post-mitotic neurons (Khacho et al., 2016). Additionally, following NPC mitosis, in one daughter cell the mitochondria undergo fusion and the cell remains pluripotent NPCs whereas a daughter cell that undergoes more mitochondrial fission differentiates towards a neuronal lineage (Iwata et al., 2020). Therefore, GIHCG-peptide's ability to regulate mitochondrial morphology in SH-SY5Y cells may suggest that it is involved in the fine tuning of mitochondrial dynamics during neuronal maturation and could potentially influence cell fate. The first experiment to characterise GIHCG-peptide function should be to use the 3 CRISPR clones generated here (Figure 5.1) to investigate the impact of GIHCG-peptide knockout on SH-SY5Y differentiation, mitochondrial network structure and cell viability.

Out of the 5 lncRNAs and one mRNA measured, GIHCG showed the highest RNA abundance in cortical organoid polysomes, despite not having the highest total RNA level, indicating it is highly translated in this model (Figure 4.16). Using endogenous tagging or by generating a custom antibody, the translation and subcellular localisation of GIHCG-peptide in cortical organoids could be investigated in the future. Frame-shift CRISPR mutants could be

generated in cortical organoids to determine the contribution to mitochondria dynamics in the range of neuronal cell types that make up cortical organoids.

6.3 LINC00839 produces a cytoplasmic protein with a potential role in neuronal differentiation

Poly-Ribo-Seq determined that LINC00839 ORF showed high translation efficiency in differentiated SH-SY5Y cells (Douka et al., 2021). In this thesis I have shown that LINC00839-protein exhibited robust expression in the FLAG-tagged expression assay producing a protein that localises to the periphery of SH-SY5Y cells (Figure 3.15). However, LINC00839 RNA was abundant at low levels in total and polysomal RNA from cortical organoids (Figure 4.15/4.16). However, organoids were only grown to day 26 to assess total RNA levels and day 25 to investigate polysome association. Therefore, organoids will only just be beginning to undergo neuronal differentiation, so it is possible LINC00839 is required at a late stage of development and therefore is not robustly expressed at early stages of organoid development. Investigation of translated lncRNA expression at later time points of organoid growth may reveal what stage of brain development individual lncRNA function at.

siRNA knockdowns of LINC00839 resulted in a significant downregulation of neuronal marker genes NTN4 and RET and significantly shorter neurite in differentiated SH-SY5Y (Figure 4.6). This phenotype suggests LINC00839 is required for differentiation and neurite outgrowth of SH-SY5Y cells in response to RA. Although CRISPR mutants were generated none were complete knockouts for LINC00839. This could indicate that complete knockouts were not viable enough to recover from a single cell sort. The CRISPR process involves transfection and single cell sorting of SH-SY5Y. Both processes are highly stressful to the cell and therefore if loss of LINC00839 has an impact on cell viability it is unlikely complete knockout clones will survive to generate clonal colonies. SH-SY5Y cell viability was not assessed following siRNA knockdown of LINC00839 but should be done to identify if LINC00839 knockdown impacts neuronal cell viability.

Furthermore, work by Eilidh Ward in the Aspden Group, using data from the 100,000 Genomes Project (Turnbull et al., 2018) identified multiple possibly pathogenic mutations resulting in amino acid changes in LINC00839s ORF associated with neuronal conditions

including hereditary ataxia, congenital myopathy, and intellectual disability. This suggests LINC00839 protein could potentially have a role in neuronal development and its mutation contributes to neurodevelopmental disorders, although no current literature has described LINC00839 in this context. However, LINC00839 has been identified as a modulator of key developmental signalling pathways including Wnt/β-catenin (Rosso & Inestrosa, 2013; Yin et al., 2023).

6.4 OLMALINC produces a nuclear protein with a potential role in neuronal differentiation

The FLAG-tagging expression assay validated that OLMALINC lncRNA encodes a 123aa protein that localises to the nucleus of SH-SY5Y cells. ddPCR showed that OLMALINC RNA is highly enriched in the polysomes of day 25 cortical organoids (Figure 4.16). Evidence from CRISPR screens using tiling sgRNAs have discovered that the ORF, but not UTRs, of OLMALINC are required for cancer cell viability in multiple cancer cell lines including 2 medulloblastoma cell lines (Hofman et al., 2024; Prensner et al., 2021). This indicates OLMALINC protein rather than RNA is contributing to the disease. Additionally, analysis of the 100k genome project by Eilidh Ward identified 2 potentially pathogenic mutations that alter OLMALINCs amino acid sequence associated with intellectual disability, indicating OLMALINC may be required for human neuronal development.

My siRNA knockdown of OLMALINC resulted in a significant upregulation of NTN4 and RET but had no effect on neurite outgrowth (Figures 4.9). Additionally, OLMALINC expression has previously been shown to be essential in oligodendrocyte differentiation (Mills et al., 2015). Both the data presented in this thesis and other publications have demonstrated that OLMALINC has a role in physiological brain development (Mills et al., 2015). Additionally, multiple publications have identified OLMALINC to have a role in a range of cancers including Brest cancer (R. Zhang et al., 2023), osteosarcoma (H. Liu et al., 2022) and neuroblastoma (W. J. Lee et al., 2021) indicating that dysregulation of OLMALINC contributes to multiple pathologies.

Distinguishing if the lncRNA or protein is the functional component of OLMALINC is essential to understanding its role in neuronal biology and potentially has clinical relevance in the

treatment of cancers, specifically medulloblastoma. The strong enrichment of OLMALINC RNA in polysomes of cortical organoids suggesting it is being actively translated in both SH-SY5Y and organoid cells. Multiple OLMALINC knockout clones have been generated by CRISPR-Cas9 and mutants are currently being investigated for their viability and ability to differentiate compared to WT SH-SY5Y cells.

6.5 LIPT2-AS1 encodes a novel nuclear protein

6.5.1 LIPT2-AS1 protein contains a DNA binding domain and interacts with nuclear proteins

Current genome annotations list LIPT2-AS1 as a lncRNA and state it lacks protein-coding ability (Martin et al., 2023; O'Leary et al., 2016). However, Poly-Ribo-Seq performed by Dr Katerina Douka and Dr Isabel Birds (Douka et al RNA 2021) revealed that a 1179 nt unannotated ORF in LIPT2-AS1 is translated. Interestingly, LIPT2-AS1 was previously annotated as a protein-coding gene "hypothetical protein LOC374408", NCBI Reference Sequence: NP_940921.1. However, this sequence was permanently suppressed due to its repetitive nucleotide sequence (O'Leary et al., 2016). Despite LIPT2-AS1's repetitive DNA sequence its ORF produces a protein detected in both FLAG-tagging assays and western blot (Figure 3.13/ 3.19).

Phyre2 (Kelley et al., 2015) and EMBI-EBI SSS (Madeira et al., 2022) both predict LIPT2-AS1-protein to share structural homology to Jerky protein homolog (JKR), Jerky protein homolog-like (JRKL) and Major centromere autoantigen B (CENPB). The homology between these proteins and LIPT2-AS1-protein arises from their shared inclusion of a HTH CenpB-type DNA-binding domain. This domain is derived from the pogo transposase superfamily, which has been repeatedly domesticated in mammals (Gao et al., 2020). Therefore, it seems likely that LIPT2-AS1 has evolved in a similar way. Work by Dr Isabel Birds identified that LIPT2-AS1 ORF is conserved back to *Oryctolagus cuniculus* (European rabbit) (Douka et al., 2021) indicating that the incorporation of pogo transposase DNA encoding a HTH CenpB-type DNA-binding domain at the LIPT2-AS1 locus occurred before the divergence of *Oryctolagus cuniculus* and *Homo sapiens* from their last common ancestor.

JKR is the most similar protein to LIPT2-AS1 in both predicted structure and sequence and therefore may give insights into the potential function of LIPT2-AS1. JKR has been extensively characterisation and in humans is only translated in the central nervous system (W. Liu et al., 2002). Deletion of the jerky gene in mice results in a phenotype similar to that of familiar temporal lobe epilepsy in humans (Toth et al., 1995). Additionally, JRK is known to bind a specific subset of mRNAs through its N-terminal HTH-DNA binding domain, although the exact sequence that binds RNA has not been elucidated (W. Liu et al., 2003). JRK targets mRNAs encoding proteins involved in ribosome assembly, cell stress and cytoskeletal structure, suggesting the epilepsy phenotype observed with JRK deletion is potentially associated with defects in ribosome biogenesis and neuronal survival (W. Liu et al., 2003). The region of JKR responsible for binding DNA/RNA shares high homology with the N-terminus of LIPT2-AS1 suggesting LIPT2-AS1 protein may possess similar properties. Beyond LIPT2-AS1-proteins's N-terminal HTH-DNA binding domain, which makes up the first ~220 amino acids of LIPT2-AS1, no high confidence structural predictions (E value < 0.05) were identified. Therefore, the functional role LIPT2-AS1 proteins C-terminus is unclear. Work performed by Eilidh Ward using data from the 100,000 Genomes Project (Turnbull et al., 2018) suggests LIPT2-AS1-proteins's DNA binding domain could be important to its function. Analysis revealed a SNP in LIPT2-AS1 that is associated with intellectual disability (unpublished work). This mutation results in an amino acid change (arginine to cysteine) in LIPT2-AS1's protein and AlphaFold modelling predicts this mutation to disrupt the structure, and therefore the function, of LIPT2-AS1's DNA binding domain.

To identify which cellular processes LIPT2-AS1 is potentially involved in LIPT2-AS1-proteins's interaction partners were identified with pull down experiments followed by MS (Chapter 5.4). This allowed us to identify protein interactors with known functions so we can infer potential roles for LIPT2-AS1-protein that can then be tested experimentally in the future. From the preliminary data, LIPT2-AS1 protein primarily interacts with nuclear proteins including multiple chromatin proteins (H1-3, H2A, TMPO) and splicing factors (PPIH, PRPF4, CDC5L) (Figure 5.19). Further work is required validate the interaction partners identified by FLAG-tagged pulldown and MS of LIPT2-AS1-protein. Validation could be achieved by Co-IP experiments to ensure detected proteins are genuine interactions. Targets for validation include components of the nucleosome and members of the spliceosome.

6.5.3 LIPT2-AS1 protein is essential for neuronal differentiation

Knockdown of LIPT2-AS1 siRNA and LIPT2-AS1 protein knockout by CRISPR is SH-SY5Y cells resulted in dysregulated neuronal differentiation. Specifically, SH-SY5Y cells without LIPT2-AS1 were not able to grow neurites (Figure 5.8) and showed disrupted expression of neuronal marker genes (Figure 5.9). To understand the global effects of LIPT2-AS1 protein knockdown RNA-Seq of LIPT2-AS1$^{-/-}$ cells and WT cells was performed (Chapter 5.5). There were substantial changes in RNA levels, with the most significantly downregulated genes being involved in neuronal maturation (e.g. ISLR2, INA, Doublecortin), neurotransmitter production (e.g. Dopamine beta-hydroxylase, PTPRN), neurotransmitter release (e.g. Chromogranin A, NSG2), neuronal excitability (e.g. KCNQ2, UNC79) amongst other functions. Throughout neuronal development multiple specific transcription factors are required to control all processes required for neuronal development (Santiago & Bashaw, 2014) and function including axonal growth (Polleux et al., 2007), synaptic partner selection and dendritic arborization (Jan & Jan, 2010). The clear downregulated of neuronal genes with LIPT2-AS1-protein knockdown suggests LIPT2-AS1 may act as a pro-neuronal transcription factor.

Concurrently with the global downregulation of neuronal marker genes following LIPT2-AS1-protein knockout, genes associated with epithelial–mesenchymal transformation were upregulated. These include the oxidoreductase GPX8 (S. Li et al., 2022) and extracellular matrix components COL5A1, COL3A1 and Fibrillin 1. Additionally, FAP, FN1 and THBS1 were all upregulated in LIPT2-AS1 knockout cells and are known markers of epithelial–mesenchymal transition in glioblastoma (Cheng et al., 2012; Mikheeva et al., 2010). This upregulation of extracellular matrix proteins suggests that SH-SY5Y cells without LIPT2-AS1 differentiate down a non-neuronal lineage. SH-SY5Y cells are derived from a neuroblastoma cancer biopsied from a metastatic bone tumour (Kovalevich & Langford, 2013). Neuroblastoma originates from neural crest cells which give rise to a range of tissues beyond neurons including glia cells of the peripheral nervous system (PNS), medulla cells of the adrenal gland, pigment producing cells of the epidermis and skeletal and connective tissue components of the head and endothelial cells of aortic arch arteries (Gilbert, 2000). Supporting the hypothesis that LIPT2-AS1 protein loss contributes to SH-SY5Y differentiation

down a non-neuronal lineage is the upregulation of TBX18 in knockout cells. TBX18 is a transcription factor expressed in cardiac progenitor cells that give rise to cardiac fibroblasts, coronary smooth muscle cells and myocytes in the ventricular septum and the atrial and ventricular walls (C.-L. Cai et al., 2008; Takeichi et al., 2013). Functionally, knockout of LIPT2-AS1 resulted in a greater number of genes being downregulated than upregulated. This suggests that LIPT2-AS1 may act as a transcriptional activator mediated through its DNA binding domain. This data demonstrates LIPT2-AS1 protein is essential for SH-SY5Y cells to maintain their NPC-like status and for them to differentiate further down the neuronal lineage following treatment with RA.

Further analysis of the RNA-Seq data could be used to elucidate a more specific role for LIPT2-AS1-protein in the nucleus. Firstly, genomic DNA binding sites of LIPT2-AS1 could be predicted through analysis of the promotor regions of differentially expressed genes following LIPT2-AS1 knockdown, to identify if there are any enriched promotor or transcription factor binding sites controlling transcription of effected genes. Alternatively, ChIP-Seq could be used to identify the precise locations and DNA sequence motifs that LIPT2-AS1 binds, although this would first require an antibody against LIPT2-AS1-protein (Guo et al., 2023). Additionally, as LIPT2-AS1-protein was found to interact with splicing factors, splice-sensitive RNA-Seq analysis could be performed on the RNA-Seq. This will allow us to determine if LIPT2-AS1 deletion has any effect on RNA splicing. Alternative splicing is prominent in the brain with temporal control over transcript generation being essential to many neuronal processes including neurogenesis, synaptogenesis, and axon guidance (Wright et al., 2022). However, the depth of the sequencing in our RNA-Seq may not be sufficient to detect changes in splicing. Additionally, short read sequencing is inherently limited in identifying transcript isoforms as the reads cannot cover full length transcripts at once and therefore rely on bioinformatic tools to assemble full-length transcripts from the fragments detected. This often leads to inconsistent results depending on the algorithm used (Byrne et al., 2017). Long read sequencing avoids these issues as can sequence entire transcripts and is being widely adopted to quantify isoforms in a range of biological contexts (Joglekar et al., 2024).

The reduced number of differentially expressed genes in comparisons of immature LIPT2-AS1$^{-/-}$ and WT cells, compared to differentiated LIPT2-AS1$^{-/-}$ and WT cells, is likely due to sequencing issues rather than a genuine biological observation. The PCA (Figure 5.11) shows an outlier sample in immature LIPT2-AS1$^{-/-}$. This outlier sample showed poor coverage and high levels of multimapping reads compared to the other 11 samples. Therefore, its inclusion will likely weaken the ability of DESeq2 to identify differentially expressed genes.

To confirm the phenotype seen in siRNA and CRISPR knockdowns of LIPT2-AS1 is due to the loss of the protein alone and not from the mild disruption of LIPT2-AS1 RNA structure predicted by RNAFold (Figure 5.7), a rescue experiment could be performed. LIPT2-AS1$^{-/-}$ cells would be transfected with either a synthetic LIPT2-AS1-protein or a plasmid encoding LIPT2-AS1 protein but using different codons so the RNA sequence produced will be different from the WT RNA, therefore avoiding the possibility that the RNA encoding the ORF may itself have an RNA function. Neurite length and marker gene expression of transfected LIPT2-AS1$^{-/-}$ cells could then be analysed to determine the phenotype is rescued with reintroduction of the protein alone.

6.6 Future perspectives

Poly-Ribo-Seq identifies ORFs undergoing active translation, by isolation of RNAs bound by multiple ribosomes. However, it potentially excludes very small ORFs, or ORFs primarily translated by monosomes. This is of note in neurons where it has been proposed that in space limited parts of specialised cells, such as synapses, translation if primarily performed by monosomes (Biever et al., 2020). Further novel ORF discovery in neuronal contexts therefore should arguably use monosome Ribo-Seq in conjunction with Poly-Ribo-Seq to avoid losing transcripts containing very small ORF and ones preferentially translated by monosomes, while also identifying novel ORFs undergoing active translation in polysomes.

Additionally, while SH-SY5Y cells are a widely used model for neuronal differentiation they are limited by short differentiation timescales, cell type they can model and their cancerous origin. Therefore, performing Rib-Seq ORF discovery on a more biologically relevant model of neuronal differentiation could identify the synthesis of additional novel proteins. iPSCs can be differentiated into neuronal cells that develop into 3D cortical organoids that show

transcriptomic profiles consistent with the developing brain (Paşca et al., 2015). Advantages of cortical organoids over immortalised cell models of neuronal differentiation include their ability to differentiate over long time periods (over 100 days) and therefore will allow us to identify novel translated ORFs involved in later stages of neuronal development. Furthermore, as cortical organoids are composed of a range of cell types including neurons and glia, the identification of novel proteins involved in cell-to-cell communication will be possible.

In Chapter 4, the expression and enrichment to the polysome of 4 translated lncRNAs, including LIPT2-AS1, were identified in day 25 human cortical organoids. Using cortical organoids for further characterisation of translated-lncRNA candidates will provide a greater understanding of their role in brain development, beyond a what can be assessed using 2D single cell type models such as SH-SY5Ys. Additional parameters such as organoid structure, cellular composition and viability can be assessed and will inform us on how novel protein knockdown effects neuronal development on a whole brain level.

As previously indicated, work performed by Eilidh Ward using 100k genome data, identified multiple possibly pathogenic mutations in LIPT2-AS1, LINC00839 and OLMALINC ORFs. To investigate the impact of these mutations on neuronal differentiation and brain development CRISPR-Cas9 with Homology Directed Repair (Jasin & Rothstein, 2013) could be used to re-create these mutations in SH-SY5Y or human iPSCs. Additionally, potentially pathogenic mutations could be introduced into our FLAG-tagged reporter assay using site-directed mutagenesis. If, for example, the DNA binding domain of LIPT2-AS1 is disrupted, then the mutated LIPT2-AS1-protein should no longer localise to discreet foci in the nucleus. This will allow us to screen if the mutations likely disrupt protein function before performing CRISPR.

6.7 Conclusion

Overall, this work highlights the diversity of lncRNAs and ORFs they may contain. Of the top 4 candidates investigated throughout this thesis, all produced a stable protein in FLAG-tagged overexpression assays, were detected in the polysomes of cortical organoids and resulted in dysregulated SH-SY5Y neuronal differentiation upon siRNA knockdown (Figure 6.1 A). From these 4, GIHCG, LINC00839 and OLMALINC have characterised RNA functions in addition to the protein coding ability, indicating they are potentially dual-functional genes. LIPT2-AS1 has no known RNA function and encodes a large protein of 392 amino acids that had escaped annotation due to LIPT2-AS1s repetitive DNA sequence. LIPT2-AS1-protein contains a DNA binding domain derived from Pogo transposase, localises to the nucleus and interacts with multiple nuclear proteins including transcription and splicing factors. Knockout of LIPT2-AS1 induces a global downregulation on neuronal genes in SH-SY5Y cells indicating LIPT2-AS1 is essential for maintenance of neuronal lineage and normal neuronal differentiation (Figure 6.1 B). Further work is required to continue elucidating the role of LIPT2-AS1 and other novel ORFs described in this thesis in neuronal development and disease.

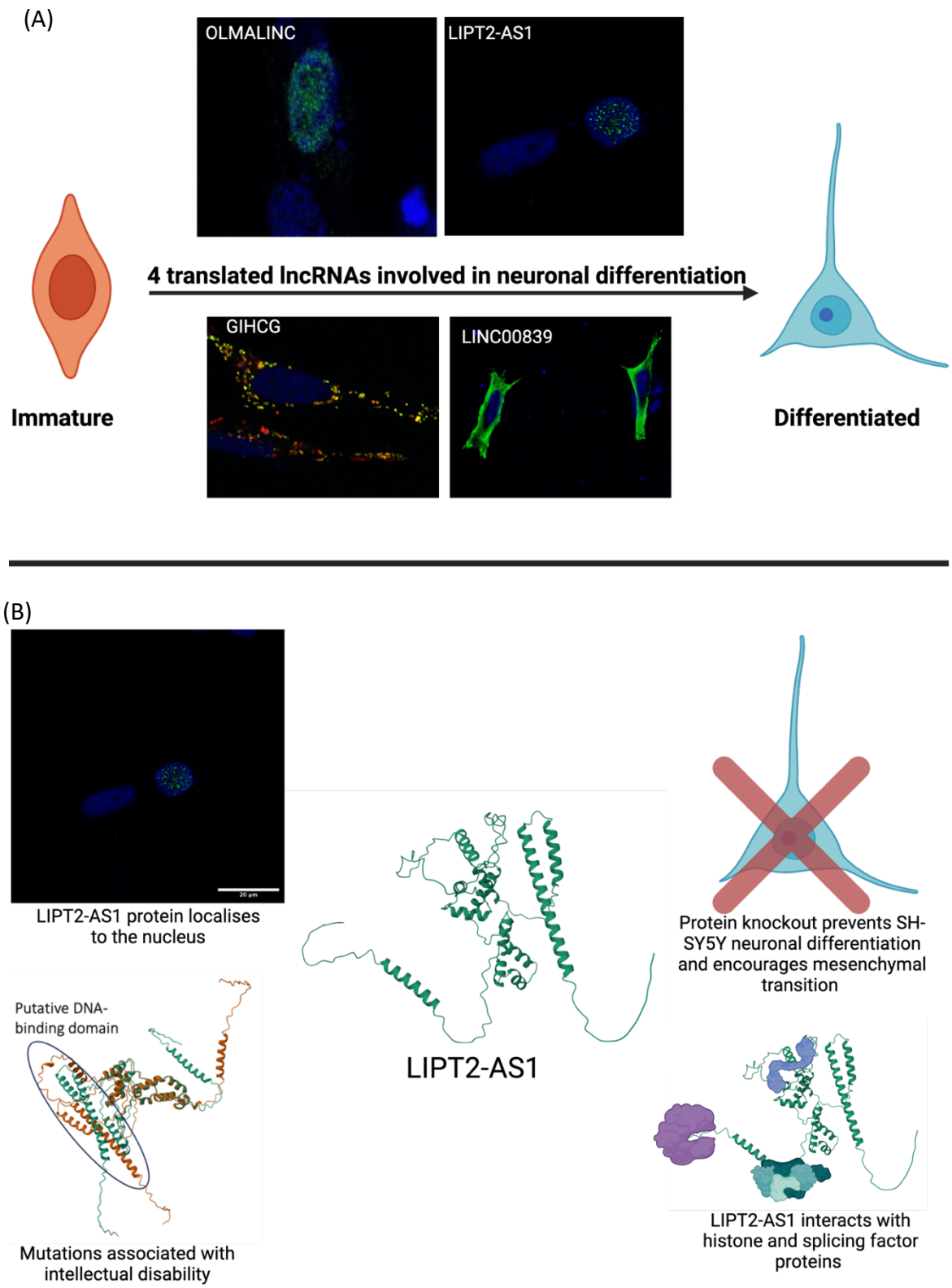


Figure 6.1: Overview of results from this thesis. *(A) multiple translated lncRNAs produce stable proteins with diverse subcellular localisation and are required for normal neuronal differentiation. (B) LIPT2-AS1 encodes a protein that contains a conserved DNA binding domain, localises to the nucleus, interacts with chromatin and splicing factors, is essential for SH-SY5Y maintenance as an NPC and differentiation following RA treatment and mutations potentially contribute to intellectual disability*.

Appendix

Significantly upregulated genes differentiated LIPT2-AS1$^{-/-}$ vs WT

gene_id	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj
ENSG00000156535.15	75.75755	9.627611917	1.201819	8.010864	1.14E-15	2.68E-14
ENSG00000251003.9	36.54549	8.567775692	1.225142	6.993294	2.69E-12	4.95E-11
ENSG00000164294.14	624.9862	8.112965658	0.412296	19.67754	3.36E-86	1.23E-83
ENSG00000135678.12	37.34849	7.632251346	1.222766	6.241791	4.33E-10	6.43E-09
ENSG00000258053.1	16.14692	7.395651146	1.274162	5.804326	6.46E-09	8.44E-08
ENSG00000168899.5	57.63434	7.363752305	1.050766	7.007983	2.42E-12	4.48E-11
ENSG00000186086.19	8.990764	6.553221268	1.34154	4.884849	1.04E-06	1.01E-05
ENSG00000146233.8	17.40284	6.520862313	1.268571	5.140322	2.74E-07	2.92E-06
ENSG00000112280.18	38.73873	6.198905741	0.89917	6.894029	5.42E-12	9.77E-11
ENSG00000218233.1	25.11134	6.1447442	1.088524	5.645027	1.65E-08	2.04E-07
ENSG00000170160.19	6.684867	6.13440799	1.417666	4.327119	1.51E-05	0.00012
ENSG00000150457.9	245.5352	6.015784256	0.853482	7.048523	1.81E-12	3.38E-11
ENSG00000196427.14	6.189734	5.99757586	1.43448	4.18101	2.9E-05	0.000218
ENSG00000279675.1	5.189384	5.756532306	1.450995	3.9673	7.27E-05	0.000501
ENSG00000176532.4	5.147492	5.746427283	1.461168	3.932764	8.4E-05	0.000569
ENSG00000169064.13	28.15213	5.720358107	0.910557	6.282265	3.34E-10	5E-09
ENSG00000131620.18	4.953666	5.685573552	1.465802	3.878813	0.000105	0.000694
ENSG00000231652.2	4.526801	5.561856167	1.485718	3.743547	0.000181	0.001139
ENSG00000113532.13	48.10403	5.499073327	0.653418	8.415857	3.9E-17	1.02E-15
ENSG00000169946.14	22.5356	5.427606106	0.934597	5.807429	6.34E-09	8.3E-08
ENSG00000250305.9	13.46667	5.251572612	1.187485	4.422434	9.76E-06	8.01E-05
ENSG00000119121.22	7.201489	5.232670635	1.380601	3.790141	0.000151	0.000964
ENSG00000198523.6	3.22258	5.076197806	1.704321	2.978427	0.002897	0.013078
ENSG00000251493.5	78.18886	5.073576223	0.45831	11.0702	1.75E-28	8.99E-27
ENSG00000283148.1	2.980364	4.968000174	1.659419	2.99382	0.002755	0.012494
ENSG00000234147.4	2.841524	4.886441691	1.649126	2.963049	0.003046	0.013629
ENSG00000225180.8	2.837562	4.872221084	1.663909	2.928177	0.00341	0.015064
ENSG00000182492.16	15.15239	4.799673747	0.968513	4.955716	7.21E-07	7.25E-06
ENSG00000258734.2	2.694832	4.795899684	1.715697	2.795306	0.005185	0.02155
ENSG00000057593.14	2.549538	4.743077609	1.710299	2.773244	0.00555	0.022828
ENSG00000290976.1	2.486701	4.711605721	1.735383	2.715023	0.006627	0.026605
ENSG00000260911.2	2.539007	4.709605328	1.773244	2.655926	0.007909	0.03079
ENSG00000261646.1	2.503685	4.706790886	1.69487	2.777081	0.005485	0.022601
ENSG00000289067.1	2.482739	4.696361377	1.698905	2.764346	0.005704	0.023385
ENSG00000273218.1	13.87188	4.688405646	0.997286	4.701164	2.59E-06	2.36E-05
ENSG00000132259.13	2.448699	4.680110534	1.710112	2.736727	0.006205	0.025109
ENSG00000102445.20	2.448699	4.680110534	1.710112	2.736727	0.006205	0.025109
ENSG00000293309.1	2.478778	4.679915942	1.719416	2.721805	0.006493	0.026142
ENSG00000122824.11	36.55701	4.668505932	0.60652	7.697206	1.39E-14	3.03E-13
ENSG00000257869.1	9.112126	4.652286214	1.206594	3.855718	0.000115	0.000758
ENSG00000272030.1	2.432925	4.644154236	1.791553	2.59225	0.009535	0.036041
ENSG00000287294.1	4.78951	4.637525886	1.475537	3.142942	0.001673	0.008084

ENSG00000111728.11	67.36508	4.548686672	0.439653	10.34607	4.36E-25	1.87E-23
ENSG00000101000.6	37.86328	4.54551939	0.57607	7.890572	3.01E-15	6.85E-14
ENSG00000176788.9	188.4781	4.462448436	0.26441	16.877	6.64E-64	1.41E-61
ENSG00000179088.16	4.128209	4.408546613	1.515827	2.908344	0.003633	0.015865
ENSG00000222033.1	4.010385	4.372147563	1.511256	2.893056	0.003815	0.01658
ENSG00000283511.2	10.9974	4.363335274	1.024602	4.258568	2.06E-05	0.000159
ENSG00000235855.3	3.922641	4.333661159	1.513799	2.862772	0.0042	0.017994
ENSG00000143341.12	20.70669	4.267154374	0.733916	5.81423	6.09E-09	7.98E-08
ENSG00000162624.16	62.47093	4.254425066	0.419622	10.1387	3.72E-24	1.51E-22
ENSG00000153291.16	55.49426	4.163729619	0.442138	9.417258	4.63E-21	1.59E-19
ENSG00000261616.1	3.446033	4.150507357	1.571625	2.640903	0.008269	0.031927
ENSG00000236078.2	3.45101	4.11278797	1.580982	2.601413	0.009284	0.035249
ENSG00000143816.8	12.35091	4.110019634	0.916478	4.484582	7.31E-06	6.15E-05
ENSG00000143119.14	3.354327	4.092359803	1.585943	2.580395	0.009869	0.037083
ENSG00000112837.17	576.8498	3.967905095	0.143458	27.65904	2.2E-168	4.4E-165
ENSG00000186340.17	73.72308	3.923506584	0.393466	9.971654	2.03E-23	7.86E-22
ENSG00000286288.2	5.420718	3.848259888	1.319411	2.91665	0.003538	0.015511
ENSG00000197935.7	36.52712	3.833035424	0.505154	7.58785	3.25E-14	6.88E-13
ENSG00000272588.1	5.26936	3.824072449	1.304194	2.932135	0.003366	0.014889
ENSG00000267737.1	5.218264	3.804035507	1.318676	2.884737	0.003917	0.016952
ENSG00000214106.9	104.9493	3.723362874	0.296887	12.54134	4.43E-36	3.12E-34
ENSG00000141574.8	4.669419	3.5776576	1.35249	2.645237	0.008163	0.031619
ENSG00000087494.16	4.257392	3.565558655	1.397586	2.551227	0.010734	0.039758
ENSG00000241362.2	4.427765	3.563283502	1.398866	2.547266	0.010857	0.040111
ENSG00000164023.15	14.86594	3.531381648	0.76943	4.589604	4.44E-06	3.89E-05
ENSG00000189409.14	38.34443	3.504210466	0.463522	7.559959	4.03E-14	8.46E-13
ENSG00000178343.5	41.59305	3.421018942	0.462337	7.399399	1.37E-13	2.77E-12
ENSG00000164342.14	52.17334	3.410834119	0.38913	8.765279	1.86E-18	5.38E-17
ENSG00000167311.14	9.644999	3.34199824	0.891641	3.748141	0.000178	0.00112
ENSG00000255794.11	7.416963	3.291779981	1.010109	3.258835	0.001119	0.005699
ENSG00000189292.17	9.32375	3.281239201	0.91424	3.589034	0.000332	0.001947
ENSG00000072952.21	137.5109	3.27318353	0.534395	6.125026	9.07E-10	1.3E-08
ENSG00000156395.14	25.53872	3.258008692	0.535983	6.078568	1.21E-09	1.71E-08
ENSG00000164442.11	34.43731	3.234209704	0.481124	6.722199	1.79E-11	3.06E-10
ENSG00000119917.15	18.52573	3.232219937	0.660361	4.894626	9.85E-07	9.63E-06
ENSG00000108846.16	58.91787	3.079295687	0.361183	8.525579	1.52E-17	4.1E-16
ENSG00000204335.4	24.30652	3.055440299	0.540384	5.654204	1.57E-08	1.94E-07
ENSG00000125872.9	47.89941	3.043106605	0.387096	7.861383	3.8E-15	8.61E-14
ENSG00000162267.14	4.621645	3.035086641	1.214051	2.499967	0.01242	0.045008
ENSG00000175984.16	19.21201	3.013392126	0.607714	4.958571	7.1E-07	7.17E-06
ENSG00000115594.12	14.44119	2.99374384	0.728547	4.109199	3.97E-05	0.00029
ENSG00000105278.12	31.65813	2.989081429	0.497165	6.012256	1.83E-09	2.54E-08
ENSG00000158156.9	6.152408	2.975843426	1.036629	2.870692	0.004096	0.017616
ENSG00000169752.17	55.54046	2.931117485	0.363428	8.065204	7.31E-16	1.74E-14
ENSG00000128606.13	1467.542	2.923401933	0.105825	27.62478	5.6E-168	1E-164
ENSG00000093072.19	7.136915	2.920889588	1.007991	2.897734	0.003759	0.016355

ENSG00000249684.6	5.711895	2.912422947	1.081238	2.693601	0.007068	0.028073
ENSG00000156466.11	82.20116	2.89838216	0.296837	9.76423	1.6E-22	5.9E-21
ENSG00000177469.13	418.9817	2.874903083	0.147264	19.52216	7.12E-85	2.51E-82
ENSG00000205832.8	5.74095	2.874194886	1.109396	2.590773	0.009576	0.036176
ENSG00000186838.15	7.234994	2.837896683	0.952989	2.977891	0.002902	0.013095
ENSG00000011201.12	6.848308	2.829449639	1.060398	2.668291	0.007624	0.029901
ENSG00000143167.12	8.400543	2.826153856	0.897098	3.150329	0.001631	0.007918
ENSG00000099250.18	61.54344	2.817391293	0.341423	8.251909	1.56E-16	3.92E-15
ENSG00000074410.14	702.2267	2.812741949	0.127263	22.10185	3E-108	1.7E-105
ENSG00000189171.15	64.14049	2.779911272	0.321103	8.657371	4.83E-18	1.36E-16
ENSG00000150938.10	9.461969	2.778385111	0.89319	3.110633	0.001867	0.008913
ENSG00000169247.14	13.55343	2.777176051	0.679578	4.086618	4.38E-05	0.000317
ENSG00000150556.17	17.67547	2.777101138	0.603956	4.598183	4.26E-06	3.75E-05
ENSG00000149596.7	6.660546	2.773361567	0.988682	2.80511	0.00503	0.020997
ENSG00000186197.15	73.43028	2.758598088	0.304354	9.063793	1.26E-19	3.95E-18
ENSG00000171759.10	27.99347	2.757696361	0.482629	5.7139	1.1E-08	1.4E-07
ENSG00000205795.5	73.79048	2.74542487	0.306097	8.96914	2.99E-19	9.13E-18
ENSG00000116711.10	49.62349	2.745110511	0.36195	7.584224	3.34E-14	7.07E-13
ENSG00000286015.1	19.62862	2.731312697	0.569945	4.79224	1.65E-06	1.56E-05
ENSG00000286214.1	36.55485	2.678364698	0.437233	6.125708	9.03E-10	1.3E-08
ENSG00000153993.14	220.3307	2.675923515	0.187981	14.23507	5.55E-46	6.16E-44
ENSG00000112333.12	11.47413	2.664153473	0.740158	3.599439	0.000319	0.001881
ENSG00000189056.15	7741.306	2.651658793	0.087954	30.14839	1.1E-199	3.3E-196
ENSG00000185761.11	23.14349	2.64982284	0.552953	4.792127	1.65E-06	1.56E-05
ENSG00000250608.2	6.335988	2.638263423	1.007846	2.617724	0.008852	0.033891
ENSG00000157601.15	8.902016	2.629275237	0.85662	3.069359	0.002145	0.01007
ENSG00000288871.2	6.434852	2.623983402	1.008198	2.602646	0.009251	0.035142
ENSG00000188869.13	27.27153	2.551303431	0.506979	5.032363	4.84E-07	5E-06
ENSG00000263326.1	6.957112	2.543924317	0.950913	2.675244	0.007467	0.029412
ENSG00000169903.7	7.314917	2.525349549	0.934639	2.70195	0.006893	0.027502
ENSG00000205403.15	73.16793	2.511324505	0.297772	8.433719	3.35E-17	8.86E-16
ENSG00000123243.15	39.98918	2.50272599	0.401084	6.239903	4.38E-10	6.5E-09
ENSG00000172987.14	6.799567	2.489945414	0.9353	2.662189	0.007763	0.030353
ENSG00000073792.16	217.467	2.485000045	0.175889	14.12826	2.54E-45	2.71E-43
ENSG00000063127.16	37.94776	2.481965182	0.402921	6.159932	7.28E-10	1.06E-08
ENSG00000092529.26	7.941085	2.480527991	0.860575	2.882406	0.003947	0.01706
ENSG00000180739.15	62.99215	2.478016593	0.315558	7.852799	4.07E-15	9.2E-14
ENSG00000137462.9	33.6687	2.463485339	0.417183	5.905049	3.53E-09	4.74E-08
ENSG00000124134.9	12.60426	2.459130319	0.706916	3.478674	0.000504	0.002818
ENSG00000170961.7	121.1485	2.455389931	0.228873	10.72816	7.51E-27	3.58E-25
ENSG00000120337.10	15.98338	2.450875292	0.620038	3.952779	7.72E-05	0.000529
ENSG00000157306.14	8.593378	2.428010768	0.860504	2.821614	0.004778	0.020128
ENSG00000287855.1	8.299556	2.418282515	0.933873	2.589519	0.009611	0.036288
ENSG00000183114.8	26.80245	2.414110472	0.479849	5.030981	4.88E-07	5.03E-06
ENSG00000292998.1	31.56692	2.403931935	0.427422	5.624264	1.86E-08	2.29E-07
ENSG00000224468.3	11.51249	2.403702463	0.74028	3.24702	0.001166	0.005913

ENSG00000007174.18	9.002459	2.400317987	0.826756	2.903298	0.003693	0.016099
ENSG00000262903.1	14.18005	2.399647741	0.634823	3.780027	0.000157	0.000999
ENSG00000181072.12	15.24106	2.394208819	0.66322	3.609975	0.000306	0.001816
ENSG00000243836.7	41.27101	2.390908057	0.388651	6.15182	7.66E-10	1.11E-08
ENSG00000174640.15	22.65218	2.38590663	0.53341	4.472934	7.72E-06	6.46E-05
ENSG00000099953.10	4734.121	2.371671778	0.077093	30.76379	8E-208	3.3E-204
ENSG00000293432.1	26.21122	2.360632698	0.484663	4.870665	1.11E-06	1.08E-05
ENSG00000270959.2	50.66952	2.350447212	0.343091	6.850806	7.34E-12	1.3E-10
ENSG00000099860.9	28.88731	2.334061575	0.505588	4.616527	3.9E-06	3.46E-05
ENSG00000138772.13	20.96424	2.300915686	0.532165	4.323689	1.53E-05	0.000121
ENSG00000103528.17	41.8704	2.300070874	0.374831	6.136296	8.45E-10	1.22E-08
ENSG00000275793.1	19.71453	2.288056402	0.560036	4.085555	4.4E-05	0.000318
ENSG00000276399.1	20.04739	2.28231921	0.530034	4.305986	1.66E-05	0.000131
ENSG00000182177.15	8.351183	2.271269758	0.865987	2.622754	0.008722	0.033445
ENSG00000106823.13	7.796167	2.244180119	0.870986	2.576597	0.009978	0.037397
ENSG00000166960.17	9.02824	2.234116084	0.783633	2.850973	0.004359	0.01859
ENSG00000233175.2	10.58839	2.214649711	0.737768	3.001824	0.002684	0.012235
ENSG00000112562.20	10.68666	2.211582898	0.761649	2.903676	0.003688	0.016086
ENSG00000182240.16	10.06279	2.207312256	0.791575	2.788506	0.005295	0.021934
ENSG00000275557.1	10.7682	2.199803983	0.716674	3.069463	0.002144	0.010069
ENSG00000103534.17	11.03196	2.196977771	0.74574	2.946039	0.003219	0.014314
ENSG00000115267.10	17.97601	2.181711236	0.55835	3.907428	9.33E-05	0.000625
ENSG00000174348.14	8.289178	2.180180963	0.845782	2.57771	0.009946	0.037331
ENSG00000142700.12	207.4512	2.171611102	0.173773	12.49682	7.77E-36	5.42E-34
ENSG00000168356.13	22.4477	2.147954402	0.556056	3.862841	0.000112	0.000737
ENSG00000068079.8	9.458909	2.144329498	0.765591	2.800881	0.005096	0.021244
ENSG00000149488.14	7.459092	2.140926536	0.840459	2.547329	0.010855	0.040111
ENSG00000111057.11	37.97281	2.137048501	0.407831	5.240031	1.61E-07	1.77E-06
ENSG00000138472.11	35.26836	2.132226471	0.394572	5.403899	6.52E-08	7.51E-07
ENSG00000119943.13	83.47997	2.125107546	0.268024	7.928803	2.21E-15	5.1E-14
ENSG00000206384.11	333.9441	2.118085787	0.161839	13.08757	3.88E-39	3.07E-37
ENSG00000187838.17	8.190176	2.116996444	0.817411	2.589881	0.009601	0.036263
ENSG00000224149.1	10.95799	2.115398909	0.737661	2.86771	0.004135	0.017757
ENSG00000122861.16	12.96173	2.114968197	0.659761	3.205659	0.001348	0.006703
ENSG00000167281.20	10.14441	2.113560095	0.77625	2.722781	0.006473	0.026091
ENSG00000184304.17	125.9247	2.113220111	0.226161	9.343858	9.29E-21	3.15E-19
ENSG00000168016.15	50.26266	2.099148682	0.368361	5.69862	1.21E-08	1.52E-07
ENSG00000143869.7	214.7631	2.084853039	0.172353	12.0964	1.1E-33	7.05E-32
ENSG00000274370.1	18.71162	2.073524839	0.542909	3.819288	0.000134	0.000866
ENSG00000166033.13	63.14402	2.072582157	0.31297	6.622294	3.54E-11	5.89E-10
ENSG00000151468.11	713.5314	2.0666813	0.108135	19.11209	2E-81	6.5E-79
ENSG00000074527.13	475.9261	2.057161142	0.120778	17.03253	4.71E-65	1.04E-62
ENSG00000261761.5	11.48377	2.052052609	0.709697	2.89145	0.003835	0.016651
ENSG00000036530.9	18.05777	2.051868889	0.581382	3.529293	0.000417	0.002376
ENSG00000234362.7	10.59392	2.046191049	0.751282	2.7236	0.006457	0.026033
ENSG00000151388.11	54.56477	2.045673829	0.32975	6.203705	5.51E-10	8.1E-09

ENSG00000228661.1	11.27948	2.043365087	0.741254	2.756634	0.00584	0.023852
ENSG00000114251.15	300.7217	2.042241424	0.161463	12.64837	1.14E-36	8.28E-35
ENSG00000144227.5	14.3321	2.041302964	0.612282	3.333925	0.000856	0.004506
ENSG00000162989.5	27.65955	2.040394584	0.461732	4.419005	9.92E-06	8.13E-05
ENSG00000168671.10	64.17685	2.035137904	0.303931	6.696057	2.14E-11	3.64E-10
ENSG00000205413.8	218.1388	2.028447936	0.169353	11.97764	4.65E-33	2.91E-31
ENSG00000151025.11	131.8125	2.025815484	0.22381	9.051477	1.41E-19	4.4E-18
ENSG00000007062.12	39.96633	2.006549907	0.36736	5.462084	4.71E-08	5.52E-07
ENSG00000253361.3	9.629545	2.003786743	0.75951	2.638263	0.008333	0.03214
ENSG00000021826.18	128.7578	2.002485336	0.225286	8.888625	6.19E-19	1.85E-17
ENSG00000146197.9	102.1682	1.995065361	0.243337	8.19876	2.43E-16	5.98E-15
ENSG00000127418.15	215.6396	1.994654592	0.167533	11.90604	1.1E-32	6.8E-31
ENSG00000162595.8	9.428883	1.99135232	0.74235	2.682499	0.007307	0.028893
ENSG00000154027.19	890.2598	1.990875618	0.101785	19.55953	3.42E-85	1.23E-82
ENSG00000108001.16	17.50102	1.989720082	0.566045	3.515127	0.00044	0.002491
ENSG00000177706.9	18.43149	1.985750764	0.558931	3.552767	0.000381	0.002199
ENSG00000260577.2	38.35766	1.980763645	0.372073	5.323583	1.02E-07	1.15E-06
ENSG00000196337.14	15.12992	1.97682784	0.591389	3.342687	0.00083	0.004384
ENSG00000142089.17	241.7658	1.975497047	0.167504	11.79375	4.2E-32	2.54E-30
ENSG00000279205.1	10.57346	1.972838806	0.725029	2.721049	0.006508	0.026197
ENSG00000286389.2	10.42212	1.964685076	0.738336	2.660964	0.007792	0.03042
ENSG00000156042.19	24.50221	1.957442754	0.470111	4.163786	3.13E-05	0.000233
ENSG00000066230.12	32.3047	1.951795689	0.490724	3.977378	6.97E-05	0.000482
ENSG00000137198.10	64.72211	1.948106005	0.308134	6.322268	2.58E-10	3.96E-09
ENSG00000290007.1	20.95873	1.939915966	0.502142	3.863282	0.000112	0.000736
ENSG00000165124.19	292.8881	1.934901077	0.150221	12.88038	5.8E-38	4.39E-36
ENSG00000189129.14	15.88575	1.921162021	0.593272	3.238248	0.001203	0.006074
ENSG00000168394.13	27.62334	1.91892598	0.439398	4.36717	1.26E-05	0.000101
ENSG00000272323.1	14.28111	1.915191355	0.630873	3.035778	0.002399	0.011107
ENSG00000103316.12	2884.935	1.914469621	0.090888	21.064	1.7E-98	8.28E-96
ENSG00000168679.18	510.3968	1.877652717	0.120476	15.58532	9.16E-55	1.41E-52
ENSG00000062038.15	59.36082	1.868527505	0.318776	5.861574	4.58E-09	6.08E-08
ENSG00000148488.17	954.2387	1.865911868	0.112664	16.56177	1.32E-61	2.59E-59
ENSG00000183853.18	1066.577	1.861466176	0.112016	16.61792	5.17E-62	1.03E-59
ENSG00000131634.14	65.38569	1.850522414	0.295816	6.255644	3.96E-10	5.9E-09
ENSG00000095303.17	113.5626	1.849049128	0.23067	8.016002	1.09E-15	2.57E-14
ENSG00000174740.8	28.01771	1.84394169	0.447298	4.122405	3.75E-05	0.000275
ENSG00000124507.11	14.21443	1.843675571	0.63171	2.918546	0.003517	0.015444
ENSG00000108231.14	89.3916	1.841729557	0.247916	7.428836	1.1E-13	2.24E-12
ENSG00000172752.16	12.87904	1.840461813	0.670375	2.745422	0.006043	0.024555
ENSG00000230615.8	13.26295	1.836940066	0.710566	2.585179	0.009733	0.03666
ENSG00000164920.10	11.29048	1.833048499	0.688919	2.660759	0.007796	0.030427
ENSG00000215097.3	13.93445	1.828618375	0.692339	2.641217	0.008261	0.031911
ENSG00000110002.16	77.58174	1.827912939	0.266272	6.864846	6.66E-12	1.19E-10
ENSG00000246777.1	16.25889	1.807875238	0.575306	3.142461	0.001675	0.008095
ENSG00000147251.16	2948.642	1.806972585	0.099332	18.19121	6.06E-74	1.67E-71

ENSG00000244586.1	29.13018	1.799549131	0.427756	4.20695	2.59E-05	0.000196
ENSG00000102554.14	31.38939	1.798815898	0.415267	4.331714	1.48E-05	0.000117
ENSG00000290867.1	66.11254	1.798274816	0.285678	6.294763	3.08E-10	4.65E-09
ENSG00000151572.18	56.05773	1.793895759	0.363135	4.940024	7.81E-07	7.8E-06
ENSG00000261485.1	15.31509	1.793797295	0.622316	2.882456	0.003946	0.01706
ENSG00000168994.14	60.1781	1.791817608	0.320006	5.599326	2.15E-08	2.64E-07
ENSG00000182022.18	79.43786	1.783415211	0.266663	6.68789	2.26E-11	3.84E-10
ENSG00000163661.4	185.204	1.782978169	0.181128	9.843771	7.29E-23	2.73E-21
ENSG00000184809.13	87.75363	1.778683133	0.268239	6.630975	3.33E-11	5.57E-10
ENSG00000089847.13	17.40117	1.776686809	0.578876	3.069201	0.002146	0.010071
ENSG00000119408.17	78.80987	1.768162907	0.270706	6.531684	6.5E-11	1.07E-09
ENSG00000164236.12	36.43443	1.767729235	0.380543	4.645286	3.4E-06	3.03E-05
ENSG00000163235.16	13.45811	1.767511017	0.629869	2.806156	0.005014	0.020944
ENSG00000225138.8	163.0572	1.766905799	0.195454	9.040029	1.57E-19	4.86E-18
ENSG00000151892.16	63.2355	1.766361647	0.322507	5.476975	4.33E-08	5.1E-07
ENSG00000148680.16	20.89416	1.763966447	0.506356	3.483647	0.000495	0.002772
ENSG00000103742.12	21.37811	1.761493236	0.485343	3.629377	0.000284	0.001705
ENSG00000182667.15	134.041	1.760675147	0.213331	8.253268	1.54E-16	3.89E-15
ENSG00000123496.8	170.7147	1.753810958	0.196986	8.903215	5.43E-19	1.63E-17
ENSG00000167779.9	3334.886	1.751821488	0.106635	16.42815	1.2E-60	2.25E-58
ENSG00000293404.1	16.51514	1.739308028	0.555558	3.130741	0.001744	0.008389
ENSG00000104332.12	11301.8	1.73549067	0.0774	22.42246	2.4E-111	1.5E-108
ENSG00000289207.1	18.63289	1.735327713	0.546342	3.176266	0.001492	0.007358
ENSG00000140795.13	194.5369	1.733481025	0.176352	9.829674	8.39E-23	3.13E-21
ENSG00000109072.14	230.5329	1.731521076	0.169692	10.20393	1.9E-24	7.84E-23
ENSG00000183508.5	105.3119	1.72941558	0.231381	7.47431	7.76E-14	1.6E-12
ENSG00000255248.11	130.9409	1.727785871	0.207322	8.333832	7.83E-17	2.02E-15
ENSG00000153976.3	19.34126	1.726110822	0.517123	3.337912	0.000844	0.004451
ENSG00000187244.12	38.15596	1.720454991	0.423383	4.063587	4.83E-05	0.000346
ENSG00000150687.12	53.88715	1.71988695	0.326332	5.270355	1.36E-07	1.51E-06
ENSG00000196932.12	26.6688	1.718820197	0.439677	3.90928	9.26E-05	0.000621
ENSG00000169760.18	327.0158	1.714918997	0.148317	11.5625	6.38E-31	3.65E-29
ENSG00000169891.18	392.0742	1.706408376	0.127795	13.35271	1.14E-40	9.89E-39
ENSG00000204936.10	136.7749	1.698019731	0.257478	6.594816	4.26E-11	7.04E-10
ENSG00000182326.17	61.64387	1.696386561	0.297399	5.704077	1.17E-08	1.48E-07
ENSG00000113805.9	286.1317	1.696368981	0.146629	11.5691	5.91E-31	3.4E-29
ENSG00000277287.1	35.75659	1.694726538	0.398402	4.253813	2.1E-05	0.000162
ENSG00000106483.12	389.1842	1.691297596	0.127063	13.31073	2.01E-40	1.72E-38
ENSG00000120594.17	78.02906	1.668985454	0.265234	6.292503	3.12E-10	4.71E-09
ENSG00000163359.17	244.9281	1.660985238	0.159495	10.41399	2.14E-25	9.35E-24
ENSG00000196549.13	283.2877	1.659805232	0.158848	10.44904	1.48E-25	6.52E-24
ENSG00000107611.16	14.11891	1.649137024	0.608615	2.709654	0.006735	0.026981
ENSG00000132561.14	192.8381	1.646721289	0.170236	9.673164	3.92E-22	1.42E-20
ENSG00000284773.1	29.00373	1.644551214	0.418045	3.933911	8.36E-05	0.000567
ENSG00000188818.13	37.39098	1.644126078	0.383042	4.29229	1.77E-05	0.000138
ENSG00000179846.11	22.22984	1.635702532	0.475915	3.436963	0.000588	0.003237

ENSG00000173641.18	37.65576	1.634144906	0.369367	4.424181	9.68E-06	7.95E-05
ENSG00000130635.17	2684.228	1.624814205	0.076671	21.19197	1.1E-99	5.64E-97
ENSG00000095587.9	21.5164	1.623899259	0.490593	3.310072	0.000933	0.004846
ENSG00000183778.19	126.0739	1.617286509	0.224618	7.200177	6.01E-13	1.16E-11
ENSG00000162614.19	353.2998	1.61668921	0.156693	10.31756	5.87E-25	2.48E-23
ENSG00000105963.15	21.75887	1.614806343	0.477889	3.37904	0.000727	0.003901
ENSG00000142279.13	507.065	1.612713458	0.121904	13.22933	5.94E-40	4.96E-38
ENSG00000137507.11	92.44638	1.610123549	0.276781	5.817319	5.98E-09	7.85E-08
ENSG00000160161.9	560.1694	1.606444575	0.112902	14.22862	6.09E-46	6.72E-44
ENSG00000173705.9	177.8179	1.600708445	0.186168	8.598177	8.1E-18	2.23E-16
ENSG00000107738.20	23.68157	1.599637287	0.475566	3.363652	0.000769	0.00409
ENSG00000273312.2	14.53503	1.598669493	0.592403	2.698616	0.006963	0.027729
ENSG00000076344.16	124.2559	1.598437567	0.227774	7.017634	2.26E-12	4.2E-11
ENSG00000106484.16	896.2062	1.598249444	0.112083	14.2595	3.91E-46	4.37E-44
ENSG00000187260.16	1368.448	1.59431928	0.08916	17.88151	1.64E-71	4.3E-69
ENSG00000139352.4	691.0237	1.592353423	0.111359	14.29927	2.21E-46	2.51E-44
ENSG00000163520.14	413.9767	1.592180808	0.124821	12.75569	2.9E-37	2.15E-35
ENSG00000076716.9	48.62892	1.588440589	0.374901	4.236964	2.27E-05	0.000174
ENSG00000182179.13	126.8806	1.584258788	0.205521	7.708516	1.27E-14	2.79E-13
ENSG00000163827.14	50.84708	1.58305355	0.33957	4.661936	3.13E-06	2.83E-05
ENSG00000203772.8	181.1699	1.57811224	0.195084	8.089415	6E-16	1.43E-14
ENSG00000161638.11	15.02963	1.571894	0.580307	2.708729	0.006754	0.027051
ENSG00000147041.12	27.51752	1.56906267	0.450925	3.479651	0.000502	0.002809
ENSG00000163749.18	35.73472	1.56651726	0.389596	4.020875	5.8E-05	0.000408
ENSG00000261888.1	78.4919	1.564616752	0.284509	5.499357	3.81E-08	4.52E-07
ENSG00000021300.14	200.6572	1.563856964	0.176638	8.853458	8.48E-19	2.51E-17
ENSG00000120885.22	3045.895	1.559802116	0.079124	19.71347	1.65E-86	6.14E-84
ENSG00000141527.19	22.34649	1.558937337	0.479846	3.248831	0.001159	0.005884
ENSG00000164855.16	14.5264	1.55436951	0.625935	2.483278	0.013018	0.046824
ENSG00000150471.17	16.62243	1.543726901	0.590281	2.615242	0.008916	0.0341
ENSG00000185774.17	34.6583	1.542328927	0.38614	3.994225	6.49E-05	0.000453
ENSG00000184226.15	381.3357	1.535697154	0.140898	10.89932	1.16E-27	5.7E-26
ENSG00000110841.14	651.0017	1.532702881	0.116217	13.18823	1.03E-39	8.38E-38
ENSG00000185201.18	896.9768	1.530790418	0.103881	14.736	3.79E-49	4.78E-47
ENSG00000287750.3	14.83369	1.529772222	0.571703	2.675816	0.007455	0.029379
ENSG00000073146.17	98.77364	1.522051131	0.233187	6.527165	6.7E-11	1.1E-09
ENSG00000262155.2	28.11307	1.521658393	0.446692	3.406505	0.000658	0.003569
ENSG00000006016.12	31.80005	1.519450512	0.403335	3.767218	0.000165	0.001046
ENSG00000047457.14	17.56782	1.519161975	0.535512	2.836841	0.004556	0.019324
ENSG00000245812.5	21.98445	1.518044751	0.487732	3.112456	0.001855	0.008866
ENSG00000073756.13	38.41559	1.50773091	0.371088	4.063005	4.84E-05	0.000347
ENSG00000129422.15	713.0164	1.506986102	0.101183	14.89372	3.62E-50	4.8E-48
ENSG00000156113.25	804.4068	1.506876703	0.137395	10.96747	5.48E-28	2.72E-26
ENSG00000101955.15	201.649	1.501737653	0.16904	8.883904	6.46E-19	1.93E-17
ENSG00000170425.4	126.9876	1.501116637	0.202666	7.406851	1.29E-13	2.63E-12
ENSG00000116396.15	68.47213	1.498838064	0.27205	5.50942	3.6E-08	4.29E-07

ENSG00000267296.2	32.39072	1.496553676	0.392616	3.811747	0.000138	0.00089
ENSG00000234899.11	35.0749	1.495212898	0.376403	3.972369	7.12E-05	0.000491
ENSG00000158270.12	2989.041	1.491227233	0.084203	17.7099	3.52E-70	8.76E-68
ENSG00000091490.11	401.1926	1.486979277	0.130974	11.35321	7.15E-30	3.89E-28
ENSG00000184363.10	35.2357	1.485561599	0.392747	3.782486	0.000155	0.000991
ENSG00000130768.15	28.61269	1.483768435	0.427281	3.472585	0.000515	0.002873
ENSG00000198626.18	170.5689	1.482567818	0.185631	7.986657	1.39E-15	3.25E-14
ENSG00000138744.16	79.08833	1.482331765	0.259514	5.711947	1.12E-08	1.41E-07
ENSG00000118898.16	34.56236	1.476239646	0.40004	3.690232	0.000224	0.001374
ENSG00000291123.1	16.1374	1.475086785	0.547562	2.693917	0.007062	0.028063
ENSG00000215769.8	37.13194	1.474709483	0.37367	3.946557	7.93E-05	0.000542
ENSG00000115363.14	60.17221	1.473270304	0.297195	4.957243	7.15E-07	7.21E-06
ENSG00000103196.12	53.23544	1.469251969	0.337986	4.347082	1.38E-05	0.00011
ENSG00000169418.10	96.34182	1.462307764	0.247519	5.907861	3.47E-09	4.67E-08
ENSG00000107249.24	24.43906	1.461999525	0.452359	3.231945	0.00123	0.00619
ENSG00000151948.12	20.19032	1.460927716	0.504684	2.894737	0.003795	0.016498
ENSG00000108691.10	115.4927	1.45407072	0.23725	6.12885	8.85E-10	1.27E-08
ENSG00000100504.17	215.8327	1.452447846	0.16364	8.875866	6.94E-19	2.07E-17
ENSG00000163817.17	17.14601	1.445580713	0.538558	2.684171	0.007271	0.028766
ENSG00000254602.2	34.48756	1.442538805	0.389729	3.701386	0.000214	0.001322
ENSG00000167617.3	25.82945	1.441291538	0.485568	2.96826	0.002995	0.013435
ENSG00000136068.16	9382.944	1.426649308	0.083209	17.14546	6.8E-66	1.51E-63
ENSG00000136869.16	20.11877	1.421582334	0.511968	2.7767	0.005491	0.022619
ENSG00000029153.15	83.05069	1.421375463	0.26521	5.359431	8.35E-08	9.48E-07
ENSG00000159403.18	546.5429	1.4200056	0.126917	11.18843	4.65E-29	2.43E-27
ENSG00000188338.15	81.4773	1.414627714	0.248239	5.698662	1.21E-08	1.52E-07
ENSG00000134115.13	41.01994	1.411074399	0.352542	4.002571	6.27E-05	0.000438
ENSG00000204518.3	342.8945	1.40995067	0.134917	10.45052	1.46E-25	6.43E-24
ENSG00000107736.22	68.79367	1.407997455	0.313401	4.492638	7.03E-06	5.94E-05
ENSG00000223764.2	36.12639	1.402897635	0.39215	3.577454	0.000347	0.00202
ENSG00000173918.15	825.1828	1.400132954	0.102921	13.604	3.79E-42	3.46E-40
ENSG00000148935.12	172.013	1.398848423	0.20174	6.93393	4.09E-12	7.43E-11
ENSG00000074590.14	78.67545	1.39462761	0.256906	5.428552	5.68E-08	6.6E-07
ENSG00000078098.15	790.4297	1.38864233	0.121141	11.46305	2.02E-30	1.12E-28
ENSG00000138193.17	433.9248	1.388530552	0.152966	9.077381	1.11E-19	3.5E-18
ENSG00000289511.1	60.31333	1.384316314	0.288047	4.805862	1.54E-06	1.46E-05
ENSG00000139410.15	108.22	1.380232945	0.231363	5.965664	2.44E-09	3.33E-08
ENSG00000152672.8	19.1152	1.379972554	0.520032	2.653628	0.007963	0.030976
ENSG00000005379.17	328.5984	1.379301035	0.147843	9.329491	1.06E-20	3.59E-19
ENSG00000166394.15	129.4542	1.378822657	0.222856	6.187047	6.13E-10	8.97E-09
ENSG00000172771.13	19.83064	1.376269664	0.527385	2.609611	0.009065	0.034557
ENSG00000061455.11	28.82338	1.37202969	0.419186	3.273083	0.001064	0.005452
ENSG00000286162.4	28.19242	1.368471151	0.441998	3.096104	0.001961	0.009308
ENSG00000137965.11	54.84536	1.368340538	0.30468	4.491067	7.09E-06	5.98E-05
ENSG00000293505.1	30.30821	1.367206925	0.412765	3.312312	0.000925	0.004822
ENSG00000079931.15	2534.826	1.361839242	0.092278	14.75802	2.73E-49	3.47E-47

ENSG00000175356.14	27.45887	1.360173743	0.495525	2.744914	0.006053	0.024589
ENSG00000260874.5	20.45321	1.36008109	0.514222	2.64493	0.008171	0.03164
ENSG00000260778.5	19.77009	1.358886021	0.503548	2.698623	0.006963	0.027729
ENSG00000171462.15	214.8863	1.358251535	0.16419	8.272431	1.31E-16	3.32E-15
ENSG00000138131.4	142.9025	1.35688053	0.210761	6.438001	1.21E-10	1.93E-09
ENSG00000107807.13	28.003	1.356524884	0.421446	3.218741	0.001288	0.006444
ENSG00000176720.6	794.6921	1.353723531	0.116145	11.65546	2.15E-31	1.26E-29
ENSG00000186301.8	39.11595	1.352438528	0.352849	3.832907	0.000127	0.000824
ENSG00000166106.4	20.20643	1.349565927	0.488919	2.760305	0.005775	0.02362
ENSG00000132470.14	120.1001	1.343826132	0.213482	6.294801	3.08E-10	4.65E-09
ENSG00000197093.11	409.3468	1.342462061	0.126759	10.59069	3.29E-26	1.51E-24
ENSG00000251615.3	45.16048	1.342203025	0.349519	3.840146	0.000123	0.000804
ENSG00000166801.17	266.8883	1.341983108	0.153346	8.751345	2.11E-18	6.08E-17
ENSG00000178031.18	407.5319	1.339552193	0.125898	10.63996	1.94E-26	9E-25
ENSG00000078596.11	506.1839	1.337420538	0.320093	4.178227	2.94E-05	0.00022
ENSG00000131724.11	944.3733	1.337345268	0.103251	12.95239	2.28E-38	1.77E-36
ENSG00000166825.15	616.9087	1.335199989	0.125169	10.66721	1.45E-26	6.79E-25
ENSG00000225950.9	35.20456	1.334367134	0.380244	3.509243	0.000449	0.002542
ENSG00000266714.11	518.4618	1.33129955	0.140638	9.466176	2.9E-21	1.01E-19
ENSG00000123689.6	37.95203	1.327341816	0.377315	3.517865	0.000435	0.002469
ENSG00000244731.10	21.27729	1.325891534	0.491143	2.699604	0.006942	0.027663
ENSG00000182218.10	106.9728	1.325844551	0.21735	6.100048	1.06E-09	1.51E-08
ENSG00000205456.12	19.14014	1.324522451	0.513571	2.579043	0.009907	0.037215
ENSG00000120820.13	89.01504	1.32418435	0.254401	5.2051	1.94E-07	2.11E-06
ENSG00000168306.13	18.78565	1.315285175	0.518711	2.535682	0.011223	0.041246
ENSG00000162591.17	239.4514	1.308953112	0.175486	7.459016	8.72E-14	1.79E-12
ENSG00000143416.21	40.3681	1.308086466	0.354708	3.687779	0.000226	0.001385
ENSG00000134321.13	48.98279	1.307183757	0.334993	3.90212	9.54E-05	0.000638
ENSG00000150907.10	222.8272	1.303738992	0.160862	8.10468	5.29E-16	1.27E-14
ENSG00000284968.1	26.4218	1.302762929	0.453562	2.872291	0.004075	0.017542
ENSG00000146674.16	936.6183	1.298632581	0.113775	11.414	3.56E-30	1.94E-28
ENSG00000131370.16	195.556	1.298300183	0.169894	7.641842	2.14E-14	4.61E-13
ENSG00000143954.13	54.18972	1.295919101	0.308716	4.197775	2.7E-05	0.000203
ENSG00000053328.9	122.2298	1.295642328	0.232309	5.577239	2.44E-08	2.98E-07
ENSG00000175928.6	431.2154	1.295554024	0.129143	10.0319	1.1E-23	4.36E-22
ENSG00000008513.17	489.1	1.295454868	0.136143	9.515375	1.81E-21	6.34E-20
ENSG00000174804.4	105.7542	1.294426508	0.239906	5.395551	6.83E-08	7.85E-07
ENSG00000278917.1	30.52285	1.287206506	0.411393	3.128899	0.001755	0.008432
ENSG00000145703.17	148.8366	1.284955175	0.196411	6.542161	6.06E-11	9.96E-10
ENSG00000213316.10	22.97217	1.282648217	0.47979	2.673351	0.00751	0.029539
ENSG00000182771.19	65.73283	1.280456868	0.271147	4.722365	2.33E-06	2.14E-05
ENSG00000163694.15	24.84096	1.274132329	0.451052	2.824801	0.004731	0.019978
ENSG00000137474.23	120.2335	1.273758424	0.211275	6.028913	1.65E-09	2.3E-08
ENSG00000124942.14	11016.1	1.273128127	0.082421	15.44663	7.95E-54	1.17E-51
ENSG00000104760.17	102.7148	1.271768266	0.235061	5.41038	6.29E-08	7.28E-07
ENSG00000124212.6	306.2961	1.269499754	0.142138	8.931481	4.2E-19	1.27E-17

ENSG00000186088.16	646.216	1.269418382	0.108331	11.71795	1.03E-31	6.14E-30
ENSG00000169851.16	129.1731	1.26924343	0.23029	5.511493	3.56E-08	4.24E-07
ENSG00000011638.11	286.717	1.26812517	0.148454	8.542213	1.32E-17	3.57E-16
ENSG00000109743.11	134.8982	1.266231811	0.201664	6.278905	3.41E-10	5.11E-09
ENSG00000177283.8	77.8642	1.26566646	0.257383	4.917439	8.77E-07	8.69E-06
ENSG00000149633.12	96.12853	1.265183553	0.231342	5.468886	4.53E-08	5.32E-07
ENSG00000177685.17	90.36606	1.264956172	0.243154	5.202287	1.97E-07	2.14E-06
ENSG00000094755.17	50.0297	1.263902772	0.328786	3.844156	0.000121	0.000792
ENSG00000099840.14	50.91452	1.262764996	0.325948	3.874133	0.000107	0.000706
ENSG00000249846.7	36.38886	1.261978915	0.38983	3.237258	0.001207	0.006091
ENSG00000198053.12	995.927	1.256202764	0.09551	13.15258	1.64E-39	1.33E-37
ENSG00000113296.14	73.64357	1.254698623	0.267724	4.686542	2.78E-06	2.53E-05
ENSG00000154102.11	23.04461	1.254373231	0.46236	2.71298	0.006668	0.026733
ENSG00000258728.2	19.36748	1.251140561	0.504458	2.480167	0.013132	0.04721
ENSG00000090539.16	20.05415	1.250739119	0.491649	2.543968	0.01096	0.040404
ENSG00000117480.16	190.1921	1.25031341	0.176017	7.103363	1.22E-12	2.3E-11
ENSG00000183775.11	47.21061	1.25028328	0.355418	3.517788	0.000435	0.002469
ENSG00000197580.13	104.2254	1.248102549	0.220614	5.657395	1.54E-08	1.91E-07
ENSG00000008311.16	117.6052	1.247481882	0.235742	5.29173	1.21E-07	1.35E-06
ENSG00000203392.3	20.15124	1.246862368	0.490046	2.54438	0.010947	0.040364
ENSG00000165171.11	25.06812	1.244783753	0.443074	2.809423	0.004963	0.020769
ENSG00000137501.18	51.80402	1.243417629	0.330929	3.757354	0.000172	0.001084
ENSG00000166147.15	6805.503	1.240017362	0.079876	15.52424	2.38E-54	3.55E-52
ENSG00000119684.16	278.8384	1.239879052	0.145429	8.525651	1.52E-17	4.1E-16
ENSG00000156869.14	164.6748	1.237877533	0.184659	6.703587	2.03E-11	3.47E-10
ENSG00000165757.9	165.6406	1.236163651	0.183195	6.747793	1.5E-11	2.6E-10
ENSG00000113140.11	4770.673	1.235016839	0.076654	16.11162	2.11E-58	3.63E-56
ENSG00000159674.12	22.36502	1.232585436	0.49471	2.491534	0.012719	0.045936
ENSG00000140398.14	267.5006	1.231620492	0.15073	8.171014	3.06E-16	7.46E-15
ENSG00000179921.15	23.09661	1.231298607	0.448612	2.744686	0.006057	0.024601
ENSG00000050628.21	82.32548	1.230490569	0.255563	4.814818	1.47E-06	1.41E-05
ENSG00000162645.13	89.06285	1.229938597	0.244745	5.025396	5.02E-07	5.17E-06
ENSG00000267666.2	30.45259	1.229131976	0.395475	3.107985	0.001884	0.00898
ENSG00000140675.13	25.089	1.227827191	0.454389	2.702151	0.006889	0.0275
ENSG00000183914.15	32.85534	1.22649507	0.433864	2.82691	0.0047	0.019867
ENSG00000265298.1	71.2891	1.225623824	0.267355	4.584259	4.56E-06	3.98E-05
ENSG00000148175.13	54.19149	1.224120478	0.312463	3.917644	8.94E-05	0.000602
ENSG00000156804.7	131.3171	1.221858507	0.1965	6.21812	5.03E-10	7.41E-09
ENSG00000122735.16	29.2498	1.220710533	0.460683	2.649785	0.008054	0.031289
ENSG00000099204.22	3266.374	1.220621391	0.102601	11.89681	1.23E-32	7.57E-31
ENSG00000070404.10	588.9196	1.220477447	0.118108	10.3336	4.97E-25	2.12E-23
ENSG00000251364.9	112.8421	1.219106386	0.225341	5.410055	6.3E-08	7.29E-07
ENSG00000113594.10	1888.032	1.21737055	0.098762	12.32625	6.54E-35	4.44E-33
ENSG00000280623.1	80.30164	1.217333543	0.279841	4.350094	1.36E-05	0.000109
ENSG00000181513.15	119.7541	1.216597933	0.213757	5.691504	1.26E-08	1.58E-07
ENSG00000167984.18	31.0659	1.215465846	0.427108	2.845803	0.00443	0.018851

ENSG00000221817.12	68.57909	1.215262665	0.267727	4.539178	5.65E-06	4.86E-05
ENSG00000173175.16	344.3155	1.213820709	0.132213	9.180795	4.28E-20	1.39E-18
ENSG00000162630.6	58.50216	1.21380319	0.328685	3.692908	0.000222	0.001362
ENSG00000171357.6	24.35081	1.210777463	0.463302	2.613367	0.008966	0.034253
ENSG00000204176.14	23.04671	1.207038632	0.466556	2.587125	0.009678	0.036501
ENSG00000152784.16	1166.698	1.205630824	0.089945	13.40414	5.72E-41	4.97E-39
ENSG00000169026.13	24.20553	1.205407846	0.445232	2.707368	0.006782	0.027152
ENSG00000041982.17	37.61418	1.204829533	0.379759	3.172618	0.001511	0.007437
ENSG00000176381.6	79.80293	1.200546418	0.306791	3.913233	9.11E-05	0.000612
ENSG00000134250.21	9318.166	1.198458397	0.086128	13.9148	5.15E-44	5.11E-42
ENSG00000197496.6	556.7758	1.198172837	0.111982	10.69972	1.02E-26	4.84E-25
ENSG00000215014.5	22.41346	1.196516058	0.460021	2.601004	0.009295	0.035285
ENSG00000231187.3	23.23381	1.196223656	0.456412	2.62093	0.008769	0.0336
ENSG00000011465.18	507.0811	1.195495966	0.123496	9.680428	3.65E-22	1.33E-20
ENSG00000062524.16	287.2165	1.190195554	0.140996	8.441337	3.14E-17	8.31E-16
ENSG00000198121.15	580.5338	1.188297441	0.114783	10.35255	4.08E-25	1.76E-23
ENSG00000170412.18	336.8209	1.187708087	0.157561	7.538107	4.77E-14	9.93E-13
ENSG00000145431.11	72.63738	1.184317171	0.259318	4.567038	4.95E-06	4.29E-05
ENSG00000142627.13	774.3095	1.184270803	0.10145	11.67339	1.74E-31	1.03E-29
ENSG00000165272.16	37.32989	1.183729248	0.353846	3.345324	0.000822	0.004345
ENSG00000248719.2	91.35628	1.180909037	0.259378	4.552852	5.29E-06	4.57E-05
ENSG00000148948.8	78.95984	1.178126591	0.265465	4.437976	9.08E-06	7.5E-05
ENSG00000013364.19	785.0559	1.17746442	0.097333	12.09731	1.09E-33	6.99E-32
ENSG00000143994.14	52.06356	1.176780563	0.32569	3.613193	0.000302	0.001797
ENSG00000273151.1	84.50206	1.176290659	0.251651	4.6743	2.95E-06	2.67E-05
ENSG00000128805.15	65.03801	1.175480479	0.279742	4.20202	2.65E-05	0.0002
ENSG00000196338.15	901.2431	1.171937084	0.096629	12.12821	7.49E-34	4.83E-32
ENSG00000260941.1	20.73923	1.17021748	0.473009	2.473986	0.013361	0.047891
ENSG00000273066.5	78.29992	1.168526156	0.271856	4.298333	1.72E-05	0.000135
ENSG00000251661.3	46.73002	1.168189531	0.363096	3.217299	0.001294	0.006471
ENSG00000056998.21	238.7941	1.167731037	0.15242	7.661284	1.84E-14	3.98E-13
ENSG00000259343.9	29.83506	1.165570826	0.419472	2.77866	0.005458	0.022501
ENSG00000166816.15	32.31699	1.164322881	0.418722	2.780657	0.005425	0.02239
ENSG00000250986.1	75.31597	1.164294559	0.260363	4.471809	7.76E-06	6.48E-05
ENSG00000152104.12	1361.615	1.161545912	0.097157	11.95529	6.09E-33	3.8E-31
ENSG00000001617.12	430.617	1.160371151	0.12775	9.083173	1.05E-19	3.33E-18
ENSG00000211955.2	31.02594	1.160078534	0.39646	2.926094	0.003432	0.015136
ENSG00000120658.14	547.5968	1.159690648	0.118735	9.76708	1.56E-22	5.75E-21
ENSG00000182118.8	107.1702	1.159455627	0.233741	4.960437	7.03E-07	7.1E-06
ENSG00000182600.10	24.31312	1.155046758	0.441982	2.613334	0.008966	0.034253
ENSG00000158113.13	51.65501	1.153346169	0.310309	3.716768	0.000202	0.001252
ENSG00000224316.1	23.85364	1.15285634	0.443298	2.600637	0.009305	0.035309
ENSG00000007384.15	745.197	1.150672516	0.104314	11.03088	2.71E-28	1.38E-26
ENSG00000246731.3	22.27965	1.150230375	0.459445	2.503522	0.012296	0.044622
ENSG00000142102.16	162.1825	1.14817316	0.207254	5.539934	3.03E-08	3.65E-07
ENSG00000149177.15	392.194	1.146789527	0.122737	9.343448	9.32E-21	3.15E-19

ENSG00000275131.4	95.1177	1.146574593	0.237374	4.83024	1.36E-06	1.31E-05
ENSG00000149150.9	52.1523	1.146495915	0.333928	3.433363	0.000596	0.003277
ENSG00000111325.16	27.16998	1.144056516	0.438297	2.610232	0.009048	0.034501
ENSG00000075223.14	1759.777	1.142426251	0.090336	12.64639	1.17E-36	8.46E-35
ENSG00000272668.2	35.57199	1.141605775	0.378495	3.016175	0.00256	0.011752
ENSG00000197321.16	4652.635	1.141442452	0.079401	14.37565	7.36E-47	8.54E-45
ENSG00000148411.8	497.518	1.138949083	0.118328	9.625365	6.25E-22	2.24E-20
ENSG00000183798.5	48.42937	1.138126616	0.326317	3.487792	0.000487	0.002737
ENSG00000104518.11	52.58765	1.136108331	0.305008	3.724846	0.000195	0.001218
ENSG00000119227.8	85.11786	1.134126814	0.263796	4.299258	1.71E-05	0.000134
ENSG00000158008.10	95.97662	1.132795443	0.232755	4.866902	1.13E-06	1.1E-05
ENSG00000113494.17	24.75235	1.129036824	0.442148	2.553528	0.010664	0.03954
ENSG00000162520.15	34.79633	1.127496704	0.430738	2.617595	0.008855	0.033898
ENSG00000173376.14	430.0273	1.127128531	0.128991	8.738063	2.37E-18	6.82E-17
ENSG00000075651.17	75.79478	1.126853152	0.262784	4.288128	1.8E-05	0.00014
ENSG00000120156.22	28.89771	1.125004568	0.423972	2.653488	0.007966	0.030983
ENSG00000185561.10	223.118	1.123365806	0.16477	6.817762	9.25E-12	1.63E-10
ENSG00000154252.12	29.35697	1.122593237	0.415395	2.702471	0.006883	0.027479
ENSG00000124593.16	73.68116	1.119560943	0.277238	4.038265	5.38E-05	0.000382
ENSG00000230438.9	82.77443	1.118263883	0.24576	4.550232	5.36E-06	4.62E-05
ENSG00000164142.16	148.576	1.117533076	0.192551	5.803842	6.48E-09	8.46E-08
ENSG00000080503.25	1521.122	1.115936253	0.094399	11.82145	3.02E-32	1.84E-30
ENSG00000125398.8	425.9708	1.11322117	0.133995	8.30796	9.74E-17	2.49E-15
ENSG00000197822.12	90.66027	1.113015928	0.242598	4.587904	4.48E-06	3.92E-05
ENSG00000169129.16	39.72263	1.111506379	0.369824	3.005503	0.002651	0.01211
ENSG00000168542.17	61174.16	1.109001244	0.065156	17.02069	5.77E-65	1.25E-62
ENSG00000118292.9	32.73652	1.108935641	0.393706	2.816658	0.004853	0.020395
ENSG00000197816.16	25.10299	1.107627835	0.431034	2.569699	0.010179	0.037989
ENSG00000145012.15	490.3712	1.104396549	0.135753	8.135359	4.11E-16	9.94E-15
ENSG00000049283.19	24.29305	1.102683251	0.444221	2.482284	0.013054	0.046947
ENSG00000130226.18	151.6742	1.100817835	0.187008	5.886487	3.94E-09	5.29E-08
ENSG00000130303.14	48.50011	1.093606549	0.316208	3.458503	0.000543	0.00301
ENSG00000290758.1	49.33066	1.093055512	0.330212	3.310166	0.000932	0.004846
ENSG00000176845.13	416.9678	1.092479234	0.132388	8.252108	1.56E-16	3.92E-15
ENSG00000106538.10	32.62893	1.090995295	0.385338	2.831268	0.004636	0.019627
ENSG00000152804.11	74.24641	1.090935679	0.258281	4.223834	2.4E-05	0.000183
ENSG00000104826.15	29.95453	1.090372216	0.406667	2.681244	0.007335	0.028996
ENSG00000167968.13	111.5153	1.088950617	0.222772	4.888195	1.02E-06	9.93E-06
ENSG00000168546.11	427.2878	1.088918377	0.120752	9.017835	1.92E-19	5.91E-18
ENSG00000077522.15	39.98105	1.087656994	0.354653	3.066817	0.002164	0.010147
ENSG00000260855.1	53.47611	1.085421634	0.302738	3.585348	0.000337	0.00197
ENSG00000167608.12	43.28859	1.085131224	0.334128	3.24765	0.001164	0.005903
ENSG00000104368.19	1977.555	1.084847417	0.086324	12.56709	3.2E-36	2.27E-34
ENSG00000168702.18	110.2256	1.084229568	0.232998	4.65339	3.27E-06	2.93E-05
ENSG00000111341.10	412.9685	1.084211882	0.169018	6.414756	1.41E-10	2.22E-09
ENSG00000075213.11	3403.869	1.082784001	0.081403	13.30153	2.27E-40	1.92E-38

ENSG00000127415.13	227.9329	1.081927125	0.171459	6.310106	2.79E-10	4.26E-09
ENSG00000164946.20	871.3432	1.080601552	0.124118	8.706279	3.14E-18	8.91E-17
ENSG00000114019.15	2256.713	1.080089524	0.082612	13.07417	4.63E-39	3.64E-37
ENSG00000105639.21	135.2143	1.07692688	0.195669	5.503833	3.72E-08	4.42E-07
ENSG00000176945.18	33.1907	1.076632892	0.373121	2.885476	0.003908	0.01692
ENSG00000182747.5	916.3079	1.075417114	0.091648	11.73427	8.51E-32	5.08E-30
ENSG00000130822.16	31.94426	1.073716304	0.389983	2.753242	0.005901	0.024043
ENSG00000230555.2	37.08624	1.073131231	0.358906	2.99001	0.00279	0.012627
ENSG00000169439.12	2112.526	1.073073852	0.098952	10.84444	2.12E-27	1.03E-25
ENSG00000128242.13	312.0318	1.073072841	0.142158	7.548441	4.4E-14	9.2E-13
ENSG00000288905.2	38.28535	1.072464052	0.369662	2.901201	0.003717	0.016186
ENSG00000197558.14	147.4102	1.070931493	0.217524	4.923273	8.51E-07	8.45E-06
ENSG00000157617.17	332.5036	1.070466755	0.134337	7.968544	1.61E-15	3.74E-14
ENSG00000118849.10	92.78758	1.067575634	0.240377	4.441263	8.94E-06	7.4E-05
ENSG00000180447.7	3177.338	1.066831702	0.094963	11.23414	2.77E-29	1.46E-27
ENSG00000092421.17	547.6515	1.064915801	0.149105	7.142045	9.2E-13	1.75E-11
ENSG00000138119.17	35.33394	1.058261646	0.398018	2.65883	0.007841	0.03056
ENSG00000135926.15	47.59785	1.05521644	0.330492	3.192865	0.001409	0.006986
ENSG00000177990.12	70.95329	1.051117591	0.288986	3.637259	0.000276	0.001656
ENSG00000153237.19	28.97502	1.051035917	0.414521	2.535542	0.011227	0.041255
ENSG00000112319.21	941.0344	1.049185072	0.099375	10.55781	4.67E-26	2.13E-24
ENSG00000138650.9	82.51137	1.047815874	0.280357	3.73743	0.000186	0.001163
ENSG00000275367.1	53.92058	1.04682117	0.309287	3.384629	0.000713	0.003839
ENSG00000176170.14	73.01772	1.044811796	0.280409	3.726022	0.000195	0.001213
ENSG00000107282.9	89.6655	1.039825937	0.231421	4.493213	7.02E-06	5.92E-05
ENSG00000239407.5	84.48765	1.039283056	0.245762	4.228825	2.35E-05	0.000179
ENSG00000141524.19	525.4902	1.037768114	0.112094	9.258004	2.08E-20	6.88E-19
ENSG00000102024.19	7254.776	1.037555312	0.090161	11.5078	1.21E-30	6.78E-29
ENSG00000128604.21	45.48964	1.037490602	0.346451	2.994624	0.002748	0.012473
ENSG00000173706.15	776.3874	1.037241333	0.103499	10.02175	1.22E-23	4.81E-22
ENSG00000109861.17	1826.602	1.036657193	0.083764	12.37592	3.53E-35	2.41E-33
ENSG00000174607.11	53.11761	1.034068592	0.315419	3.278395	0.001044	0.005363
ENSG00000116774.12	1300.297	1.033279461	0.09831	10.51044	7.73E-26	3.47E-24
ENSG00000160951.4	143.2129	1.033169772	0.198495	5.205013	1.94E-07	2.11E-06
ENSG00000142227.11	40.08943	1.031195719	0.393222	2.622426	0.008731	0.033465
ENSG00000124313.19	410.7802	1.027631144	0.126925	8.096363	5.66E-16	1.36E-14
ENSG00000290762.1	40.68579	1.026973628	0.357958	2.868982	0.004118	0.017693
ENSG00000115896.17	112.8579	1.024776767	0.215367	4.758273	1.95E-06	1.82E-05
ENSG00000135116.9	47.36009	1.022701407	0.32308	3.165478	0.001548	0.007594
ENSG00000196954.14	101.9108	1.022057893	0.246952	4.138698	3.49E-05	0.000257
ENSG00000177694.16	36.71141	1.021951433	0.357219	2.860852	0.004225	0.018092
ENSG00000279375.1	67.39481	1.021022184	0.276152	3.697321	0.000218	0.001341
ENSG00000185499.17	72.52828	1.020005791	0.272988	3.736445	0.000187	0.001167
ENSG00000118946.12	57.97784	1.01976678	0.338025	3.016839	0.002554	0.011735
ENSG00000169432.19	739.5238	1.018446245	0.110341	9.229982	2.71E-20	8.89E-19
ENSG00000168970.22	93.23585	1.017820396	0.242723	4.193347	2.75E-05	0.000207

ENSG00000183801.8	70.53553	1.015812052	0.273078	3.719865	0.000199	0.001239
ENSG00000138311.18	36.37124	1.01393478	0.360678	2.81119	0.004936	0.020681
ENSG00000049089.15	111.903	1.01266205	0.220213	4.598564	4.25E-06	3.75E-05
ENSG00000133460.20	258.0807	1.008701964	0.155452	6.488824	8.65E-11	1.39E-09
ENSG00000165475.15	170.3914	1.007543597	0.186118	5.413478	6.18E-08	7.16E-07
ENSG00000116663.11	45.76251	1.007508625	0.336449	2.994532	0.002749	0.012473
ENSG00000290600.1	45.27415	1.005990291	0.369543	2.722254	0.006484	0.026123
ENSG00000170500.13	447.802	1.005553916	0.131147	7.667392	1.76E-14	3.8E-13
ENSG00000115232.14	39.59628	1.0037333	0.358687	2.798352	0.005136	0.021394
ENSG00000001461.17	548.5237	1.002621866	0.112476	8.914098	4.92E-19	1.49E-17
ENSG00000243244.7	556.3691	1.000714529	0.116721	8.573563	1E-17	2.74E-16

Significantly downregulated genes differentiated LIPT2-AS1$^{-/-}$ vs WT

gene_id	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj
ENSG00000131089.17	183.7934	-11.0375708	1.188862	-9.28415	1.63E-20	5.42E-19
ENSG00000233608.5	17.5279	-7.651441946	1.277849	-5.98775	2.13E-09	2.93E-08
ENSG00000258279.4	9.370211	-6.741680864	1.337648	-5.03995	4.66E-07	4.82E-06
ENSG00000101282.9	7.226343	-6.362700225	1.38468	-4.59507	4.33E-06	3.8E-05
ENSG00000274330.1	6.860965	-6.295055551	1.399539	-4.49795	6.86E-06	5.81E-05
ENSG00000110436.13	5.822241	-6.06490969	1.454633	-4.16937	3.05E-05	0.000228
ENSG00000104833.12	55.27246	-5.866814837	0.65685	-8.93174	4.19E-19	1.27E-17
ENSG00000270372.1	5.064913	-5.852710574	1.465607	-3.99337	6.51E-05	0.000454
ENSG00000164326.5	4.912813	-5.807138876	1.467807	-3.95634	7.61E-05	0.000522
ENSG00000254202.2	4.695769	-5.744322169	1.47838	-3.88555	0.000102	0.000678
ENSG00000129993.15	16.56492	-5.715206412	1.133116	-5.0438	4.56E-07	4.73E-06
ENSG00000101144.13	49.36466	-5.678400533	0.656304	-8.65209	5.06E-18	1.42E-16
ENSG00000230805.8	4.295778	-5.609510467	1.529859	-3.66668	0.000246	0.001491
ENSG00000063015.21	85.82864	-5.601981482	0.507147	-11.0461	2.29E-28	1.16E-26
ENSG00000123610.5	3.862651	-5.460678944	1.552079	-3.5183	0.000434	0.002466
ENSG00000253477.5	3.573652	-5.365414207	1.618635	-3.31478	0.000917	0.004793
ENSG00000100678.20	271.4895	-5.258784371	0.269887	-19.4851	1.47E-84	5.08E-82
ENSG00000289192.1	3.328639	-5.258118069	1.610221	-3.26546	0.001093	0.005584
ENSG00000168594.15	18.6529	-5.254325963	0.96502	-5.44478	5.19E-08	6.06E-07
ENSG00000248837.7	3.205838	-5.188480184	1.626955	-3.18907	0.001427	0.007069
ENSG00000287726.1	3.154251	-5.180784707	1.639433	-3.16011	0.001577	0.00771
ENSG00000100095.19	56.07799	-5.133213546	0.534832	-9.59781	8.17E-22	2.89E-20
ENSG00000171722.13	3.042888	-5.123902546	1.693195	-3.02617	0.002477	0.011434
ENSG00000111199.12	28.25493	-5.107678521	0.738446	-6.91679	4.62E-12	8.37E-11
ENSG00000279692.1	2.991672	-5.090606984	1.634026	-3.11538	0.001837	0.008785
ENSG00000171303.8	65.05758	-5.07564604	0.499748	-10.1564	3.1E-24	1.27E-22

ENSG00000266074.11	10.83062	-5.068740121	1.183918	-4.28133	1.86E-05	0.000145
ENSG00000125851.10	2.867541	-5.039356641	1.680211	-2.99924	0.002707	0.012323
ENSG00000187416.13	5.735286	-5.023950312	1.446391	-3.47344	0.000514	0.002865
ENSG00000176293.20	2.823927	-5.021785093	1.67678	-2.9949	0.002745	0.012469
ENSG00000183439.10	2.783189	-5.003684145	1.696108	-2.9501	0.003177	0.014161
ENSG00000197406.8	20.81452	-4.991536844	0.839359	-5.94684	2.73E-09	3.71E-08
ENSG00000113262.16	2.775587	-4.985881058	1.66877	-2.98776	0.00281	0.012713
ENSG00000293490.1	2.753301	-4.976840143	1.659844	-2.99838	0.002714	0.012347
ENSG00000105976.16	2.752342	-4.976743573	1.67448	-2.97211	0.002958	0.013297
ENSG00000278012.1	2.731973	-4.967729698	1.671671	-2.97171	0.002961	0.013308
ENSG00000226650.6	2.72533	-4.948883358	1.693294	-2.92264	0.003471	0.015282
ENSG00000145451.13	2.62917	-4.920891489	1.714132	-2.87078	0.004095	0.017615
ENSG00000271830.1	2.628211	-4.920269543	1.736903	-2.83278	0.004614	0.019542
ENSG00000215397.4	9.804471	-4.905914179	1.203988	-4.07472	4.61E-05	0.000332
ENSG00000251297.1	5.220684	-4.891178459	1.476	-3.31381	0.00092	0.004804
ENSG00000146469.13	18.56813	-4.86465313	0.904678	-5.37722	7.56E-08	8.63E-07
ENSG00000013293.6	56.80019	-4.857392269	0.492977	-9.85318	6.64E-23	2.49E-21
ENSG00000230922.1	2.517807	-4.853023234	1.730733	-2.80403	0.005047	0.021059
ENSG00000162105.21	28.31484	-4.85233879	0.705439	-6.87847	6.05E-12	1.09E-10
ENSG00000237194.1	4.986586	-4.830110635	1.460389	-3.30741	0.000942	0.004888
ENSG00000164176.13	46.80263	-4.816021007	0.54277	-8.87304	7.12E-19	2.12E-17
ENSG00000074317.11	2.432496	-4.809475106	1.865846	-2.57764	0.009948	0.037332
ENSG00000120251.22	4.773312	-4.743949021	1.468935	-3.22952	0.00124	0.006238
ENSG00000123609.11	8.521306	-4.735159709	1.236181	-3.83048	0.000128	0.000831
ENSG00000205037.2	4.671468	-4.716949176	1.480779	-3.18545	0.001445	0.007149
ENSG00000101180.17	4.502242	-4.676609227	1.479161	-3.16166	0.001569	0.007674
ENSG00000135346.9	4.479955	-4.671019928	1.489914	-3.13509	0.001718	0.008278
ENSG00000011347.10	45.44896	-4.659710456	0.525306	-8.87047	7.28E-19	2.17E-17
ENSG00000139767.10	44.22454	-4.617809	0.529344	-8.72365	2.69E-18	7.69E-17
ENSG00000130054.5	23.32341	-4.589827564	0.716773	-6.40346	1.52E-10	2.38E-09
ENSG00000136944.19	7.77861	-4.560671814	1.251974	-3.64279	0.00027	0.001625
ENSG00000254035.1	4.151167	-4.550231756	1.51738	-2.99874	0.002711	0.012337
ENSG00000135480.17	7.60044	-4.549636289	1.245648	-3.65243	0.00026	0.001571
ENSG00000130294.18	587.5006	-4.473307699	0.178948	-24.9979	6.4E-138	6.3E-135
ENSG00000196092.14	72.21382	-4.449524086	0.404612	-10.997	3.95E-28	1.98E-26
ENSG00000116147.18	162.6348	-4.409976433	0.283237	-15.5699	1.17E-54	1.76E-52
ENSG00000165606.10	10.28793	-4.384514068	1.04513	-4.19519	2.73E-05	0.000205
ENSG00000187957.8	36.37247	-4.358805891	0.557135	-7.82361	5.13E-15	1.15E-13
ENSG00000100604.13	1526.98	-4.348572991	0.118164	-36.8012	1.8E-296	3.6E-292
ENSG00000109956.13	67.62252	-4.291601915	0.400384	-10.7187	8.32E-27	3.95E-25
ENSG00000180818.5	104.4664	-4.264193864	0.328655	-12.9747	1.7E-38	1.33E-36
ENSG00000155052.15	58.18107	-4.263640354	0.434357	-9.81598	9.61E-23	3.58E-21
ENSG00000104327.7	9.801046	-4.261100084	1.032891	-4.12541	3.7E-05	0.000272
ENSG00000053438.11	127.4984	-4.244010174	0.292454	-14.5117	1.02E-47	1.24E-45
ENSG00000116544.12	56.08014	-4.230207476	0.430071	-9.83608	7.87E-23	2.94E-21
ENSG00000070886.12	6.208723	-4.227927301	1.285461	-3.28904	0.001005	0.005185

ENSG00000089169.15	42.82252	-4.201460765	0.510912	-8.22346	1.98E-16	4.91E-15
ENSG00000166510.14	90.1722	-4.194130317	0.354462	-11.8324	2.65E-32	1.62E-30
ENSG00000157445.16	9.53465	-4.187412511	1.066903	-3.92483	8.68E-05	0.000586
ENSG00000166111.10	48.83135	-4.185774018	0.457137	-9.15649	5.36E-20	1.73E-18
ENSG00000136960.13	96.85955	-4.184062303	0.33544	-12.4734	1.04E-35	7.23E-34
ENSG00000198028.4	9.222324	-4.175179552	1.052419	-3.96722	7.27E-05	0.000501
ENSG00000287260.2	9.145574	-4.171412158	1.041363	-4.00572	6.18E-05	0.000432
ENSG00000152954.12	74.49641	-4.156593534	0.37905	-10.9658	5.58E-28	2.76E-26
ENSG00000144550.14	3.153562	-4.154815877	1.613893	-2.57441	0.010041	0.037586
ENSG00000054356.14	467.9951	-4.151778377	0.168551	-24.6322	5.7E-134	4.7E-131
ENSG00000274641.2	3.16877	-4.138719911	1.610882	-2.56923	0.010193	0.038027
ENSG00000178171.12	48.79877	-4.135027555	0.457335	-9.04157	1.54E-19	4.8E-18
ENSG00000255883.1	3.105745	-4.112303626	1.659864	-2.47749	0.013231	0.047473
ENSG00000186487.21	71.37001	-4.099142136	0.383801	-10.6804	1.26E-26	5.92E-25
ENSG00000029534.22	38.78038	-4.072103294	0.528177	-7.70974	1.26E-14	2.76E-13
ENSG00000122254.7	177.3939	-4.061243341	0.255541	-15.8927	7.12E-57	1.13E-54
ENSG00000101489.20	66.13914	-4.059670839	0.389477	-10.4234	1.94E-25	8.48E-24
ENSG00000110076.21	172.9294	-4.050964124	0.246337	-16.4448	9.13E-61	1.73E-58
ENSG00000130477.16	350.3138	-4.034042836	0.188182	-21.437	6E-102	3.2E-99
ENSG00000186642.16	27.58922	-4.032480196	0.596851	-6.75626	1.42E-11	2.46E-10
ENSG00000135902.10	81.0553	-4.012008477	0.381568	-10.5145	7.4E-26	3.33E-24
ENSG00000175206.11	26.95166	-4.005881416	0.606078	-6.60952	3.86E-11	6.4E-10
ENSG00000158089.15	163.8528	-4.002630204	0.266848	-14.9997	7.38E-51	1.03E-48
ENSG00000161681.17	29.25782	-3.992665403	0.560424	-7.12437	1.05E-12	1.98E-11
ENSG00000131055.5	32.23893	-3.975414855	0.560496	-7.09268	1.32E-12	2.47E-11
ENSG00000105409.19	50.62484	-3.973798111	0.440882	-9.01329	2E-19	6.14E-18
ENSG00000253230.11	5.49586	-3.963015093	1.303828	-3.03952	0.00237	0.010992
ENSG00000112038.18	73.8426	-3.949738508	0.371044	-10.6449	1.84E-26	8.55E-25
ENSG00000124194.17	56.94556	-3.940657452	0.408696	-9.64202	5.31E-22	1.91E-20
ENSG00000075035.10	7.816565	-3.939382571	1.107758	-3.55618	0.000376	0.002173
ENSG00000077080.10	130.096	-3.928054246	0.271817	-14.4511	2.47E-47	2.95E-45
ENSG00000187714.7	59.16618	-3.920027257	0.395572	-9.90977	3.78E-23	1.43E-21
ENSG00000177511.6	288.6194	-3.865881198	0.196003	-19.7236	1.35E-86	5.12E-84
ENSG00000236437.2	4.854359	-3.849014131	1.347106	-2.85725	0.004273	0.018268
ENSG00000145920.15	118.9355	-3.834809464	0.279169	-13.7365	6.13E-43	5.8E-41
ENSG00000166206.16	21.17368	-3.810352991	0.63894	-5.96355	2.47E-09	3.36E-08
ENSG00000101187.16	217.4766	-3.755500872	0.205377	-18.2859	1.07E-74	3.04E-72
ENSG00000262703.1	4.44486	-3.752406977	1.407269	-2.66645	0.007666	0.030037
ENSG00000260903.3	122.2288	-3.744884421	0.266489	-14.0527	7.42E-45	7.77E-43
ENSG00000144583.5	67.62494	-3.737290085	0.362439	-10.3115	6.25E-25	2.63E-23
ENSG00000143196.5	4.266636	-3.687649664	1.368912	-2.69385	0.007063	0.028063
ENSG00000106004.5	4.612484	-3.668259181	1.369532	-2.67848	0.007396	0.029186
ENSG00000278195.2	4.089754	-3.666031485	1.444376	-2.53814	0.011144	0.041002
ENSG00000148408.14	111.2044	-3.664552837	0.287945	-12.7266	4.21E-37	3.09E-35
ENSG00000237412.7	76.69404	-3.647128206	0.391416	-9.31779	1.19E-20	3.98E-19
ENSG00000259803.8	12.4848	-3.642376699	0.81899	-4.4474	8.69E-06	7.21E-05

ENSG00000149295.14	37.80912	-3.627873908	0.474914	-7.63902	2.19E-14	4.71E-13
ENSG00000154118.13	74.14917	-3.610954347	0.344322	-10.4871	9.9E-26	4.42E-24
ENSG00000083454.23	4.189313	-3.606541256	1.368842	-2.63474	0.00842	0.032421
ENSG00000010671.18	15.9044	-3.599676994	0.741348	-4.85558	1.2E-06	1.16E-05
ENSG00000226777.9	209.7268	-3.585249594	0.216845	-16.5337	2.1E-61	4.04E-59
ENSG00000254656.3	1370.053	-3.564488157	0.143502	-24.8393	3.4E-136	2.9E-133
ENSG00000099869.8	16.1019	-3.564085984	0.746162	-4.77656	1.78E-06	1.67E-05
ENSG00000170091.11	199.7489	-3.559474411	0.206384	-17.2469	1.18E-66	2.65E-64
ENSG00000129990.15	134.6042	-3.558124215	0.250394	-14.2101	7.93E-46	8.71E-44
ENSG00000174672.16	35.27091	-3.554822948	0.481761	-7.37881	1.6E-13	3.22E-12
ENSG00000178187.8	25.53504	-3.537312536	0.561741	-6.29705	3.03E-10	4.61E-09
ENSG00000163377.16	10.13211	-3.529501779	0.899974	-3.92178	8.79E-05	0.000593
ENSG00000257829.1	3.97227	-3.526277225	1.383001	-2.54973	0.010781	0.039908
ENSG00000198910.14	990.7459	-3.522864615	0.135657	-25.969	1.1E-148	1.7E-145
ENSG00000132437.18	619.7036	-3.52075596	0.144381	-24.3852	2.5E-131	1.9E-128
ENSG00000125878.7	11.68892	-3.500220993	0.824412	-4.24572	2.18E-05	0.000168
ENSG00000262904.1	6.054869	-3.454637127	1.160372	-2.97718	0.002909	0.013117
ENSG00000152818.20	12.7242	-3.44246658	0.776704	-4.43215	9.33E-06	7.68E-05
ENSG00000124191.18	116.802	-3.431373284	0.282213	-12.1588	5.15E-34	3.34E-32
ENSG00000123454.12	17781.4	-3.430057607	0.104541	-32.8107	4.1E-236	2.8E-232
ENSG00000258422.6	5.637054	-3.419448994	1.160069	-2.94763	0.003202	0.014262
ENSG00000157064.11	106.1031	-3.407742981	0.286057	-11.9128	1.01E-32	6.28E-31
ENSG00000160013.9	70.69241	-3.399733423	0.339516	-10.0135	1.33E-23	5.2E-22
ENSG00000015592.17	66.4773	-3.39187362	0.345652	-9.81298	9.9E-23	3.68E-21
ENSG00000155265.12	61.75464	-3.391306467	0.356871	-9.50288	2.04E-21	7.13E-20
ENSG00000196990.10	789.0121	-3.390618062	0.135707	-24.9848	8.9E-138	7.9E-135
ENSG00000077279.21	579.0309	-3.383266516	0.135372	-24.9923	7.4E-138	6.9E-135
ENSG00000261678.3	97.07758	-3.363785163	0.290513	-11.5788	5.28E-31	3.06E-29
ENSG00000197859.11	420.3791	-3.339131681	0.165352	-20.1941	1.1E-90	4.6E-88
ENSG00000167178.17	1663.144	-3.332208705	0.096262	-34.6159	1.5E-262	1.5E-258
ENSG00000102174.10	5.176339	-3.307297175	1.159927	-2.8513	0.004354	0.018574
ENSG00000180767.11	5.123204	-3.28380815	1.242557	-2.64278	0.008223	0.031799
ENSG00000133020.4	5.210379	-3.278404326	1.172928	-2.79506	0.005189	0.02156
ENSG00000134107.5	46.87419	-3.262429877	0.402154	-8.1124	4.96E-16	1.2E-14
ENSG00000248540.3	329.0389	-3.25612226	0.158129	-20.5915	3.27E-94	1.45E-91
ENSG00000229056.2	4.86319	-3.254700271	1.205249	-2.70044	0.006925	0.02761
ENSG00000156959.9	307.4001	-3.25170113	0.174745	-18.6083	2.75E-77	8.4E-75
ENSG00000183150.8	68.79805	-3.249233681	0.333955	-9.72956	2.26E-22	8.26E-21
ENSG00000158258.16	341.2182	-3.242622347	0.180508	-17.9638	3.74E-72	1.01E-69
ENSG00000162188.6	8.209744	-3.239946178	0.982826	-3.29656	0.000979	0.005067
ENSG00000280989.2	9.468214	-3.231525993	0.890691	-3.62811	0.000286	0.001711
ENSG00000134343.14	12.22106	-3.218817096	0.821069	-3.92027	8.84E-05	0.000596
ENSG00000082397.18	38.89197	-3.217972727	0.430782	-7.47007	8.02E-14	1.65E-12
ENSG00000105889.16	4.783372	-3.213190074	1.22178	-2.62993	0.00854	0.032822
ENSG00000172020.13	261.9841	-3.192957213	0.173737	-18.3781	1.97E-75	5.74E-73
ENSG00000188488.14	8.510773	-3.18783519	0.944279	-3.37595	0.000736	0.003936

ENSG00000157087.20	7.357208	-3.167767087	1.029512	-3.07696	0.002091	0.009842
ENSG00000121905.10	11.16666	-3.155569078	0.808296	-3.90398	9.46E-05	0.000633
ENSG00000149527.18	19.17477	-3.149936836	0.624286	-5.04567	4.52E-07	4.69E-06
ENSG00000196581.11	75.54451	-3.141865127	0.304557	-10.3162	5.95E-25	2.51E-23
ENSG00000120645.12	9.138291	-3.137602989	0.909524	-3.44972	0.000561	0.003099
ENSG00000112852.7	7.73218	-3.130001042	0.929735	-3.36655	0.000761	0.004052
ENSG00000226087.4	4.560645	-3.128109359	1.231857	-2.53934	0.011106	0.040876
ENSG00000130707.18	242.1211	-3.116790539	0.195862	-15.9132	5.14E-57	8.26E-55
ENSG00000170927.15	4.610138	-3.093926692	1.215585	-2.54522	0.010921	0.040311
ENSG00000277200.2	76.16364	-3.092632166	0.305991	-10.1069	5.15E-24	2.07E-22
ENSG00000169313.10	10.4615	-3.086134384	0.803794	-3.83946	0.000123	0.000805
ENSG00000231672.8	4.668383	-3.063898243	1.237474	-2.47593	0.013289	0.04764
ENSG00000171368.12	10.09261	-3.05878385	0.831413	-3.67902	0.000234	0.001427
ENSG00000213186.8	100.1925	-3.05666509	0.268979	-11.364	6.32E-30	3.44E-28
ENSG00000036565.15	1053.666	-3.037041159	0.101244	-29.9973	1.1E-197	2.7E-194
ENSG00000142549.10	74.61688	-3.032031998	0.305345	-9.92987	3.09E-23	1.18E-21
ENSG00000056487.17	561.4602	-3.0312788	0.144934	-20.915	3.91E-97	1.86E-94
ENSG00000125207.7	67.99681	-3.028389594	0.340605	-8.89122	6.04E-19	1.81E-17
ENSG00000163618.18	30.15889	-3.027843102	0.472611	-6.40663	1.49E-10	2.34E-09
ENSG00000008056.14	54.09298	-3.014621225	0.359288	-8.39055	4.84E-17	1.26E-15
ENSG00000274825.1	5.556286	-3.013707444	1.093844	-2.75515	0.005866	0.023927
ENSG00000118515.12	527.0277	-3.013038971	0.134806	-22.351	1.2E-110	7.3E-108
ENSG00000224090.1	5.744867	-2.988881193	1.062258	-2.81371	0.004897	0.020536
ENSG00000107147.14	14.69343	-2.958148774	0.693641	-4.26467	2E-05	0.000155
ENSG00000116983.13	133.189	-2.956199602	0.22589	-13.0869	3.91E-39	3.09E-37
ENSG00000163531.17	93.66035	-2.951049997	0.288643	-10.2239	1.55E-24	6.42E-23
ENSG00000104722.14	707.8587	-2.924218192	0.131503	-22.2369	1.5E-109	8.8E-107
ENSG00000167654.18	399.4248	-2.923097116	0.147	-19.885	5.49E-88	2.24E-85
ENSG00000149403.13	21.12164	-2.874866537	0.629685	-4.56556	4.98E-06	4.32E-05
ENSG00000156453.14	33.7735	-2.857671776	0.47626	-6.00024	1.97E-09	2.72E-08
ENSG00000261594.4	8.807332	-2.856140159	0.855743	-3.33761	0.000845	0.004454
ENSG00000089558.9	8.744635	-2.847802165	0.859994	-3.31142	0.000928	0.004834
ENSG00000237361.5	6.626022	-2.843792668	0.990907	-2.86989	0.004106	0.017648
ENSG00000136267.14	7.841131	-2.841244618	0.927756	-3.06249	0.002195	0.010266
ENSG00000184156.18	8.844826	-2.826234926	0.878586	-3.2168	0.001296	0.006476
ENSG00000168830.8	12.95262	-2.817854627	0.716577	-3.93238	8.41E-05	0.00057
ENSG00000125966.10	171.5455	-2.806163565	0.199934	-14.0354	9.46E-45	9.77E-43
ENSG00000231754.4	10.23919	-2.804405469	0.777963	-3.60481	0.000312	0.001848
ENSG00000187498.16	727.5345	-2.795075367	0.111244	-25.1257	2.6E-139	2.7E-136
ENSG00000168993.15	121.3427	-2.793067752	0.246914	-11.3119	1.15E-29	6.11E-28
ENSG00000167880.8	9.877314	-2.792521692	0.807012	-3.46032	0.00054	0.002991
ENSG00000289877.1	6.800921	-2.7923188	1.086947	-2.56896	0.010201	0.03805
ENSG00000135914.6	36.14952	-2.782703048	0.436987	-6.36794	1.92E-10	2.98E-09
ENSG00000120738.8	178.9533	-2.772164655	0.206811	-13.4043	5.7E-41	4.97E-39
ENSG00000148798.11	861.0145	-2.770773807	0.106771	-25.9507	1.8E-148	2.6E-145
ENSG00000111348.9	40.85874	-2.76611524	0.423053	-6.53846	6.22E-11	1.02E-09

ENSG00000140022.14	18.10645	-2.759257564	0.59681	-4.62334	3.78E-06	3.35E-05
ENSG00000129009.13	101.5223	-2.756548896	0.271376	-10.1577	3.06E-24	1.25E-22
ENSG00000196361.10	1389.089	-2.741519236	0.099735	-27.4879	2.4E-166	4.2E-163
ENSG00000151136.15	43.44291	-2.739438151	0.405171	-6.76118	1.37E-11	2.38E-10
ENSG00000206557.6	10.98019	-2.715576624	0.74198	-3.6599	0.000252	0.001528
ENSG00000249509.1	7.40839	-2.711089381	0.921571	-2.94181	0.003263	0.014498
ENSG00000258986.7	130.1932	-2.7041586	0.231464	-11.6829	1.56E-31	9.21E-30
ENSG00000113739.11	16.56726	-2.70234038	0.608392	-4.44178	8.92E-06	7.38E-05
ENSG00000128564.8	22484.85	-2.688961096	0.086678	-31.0225	2.7E-211	1.4E-207
ENSG00000177614.11	170.6477	-2.682515456	0.196264	-13.6679	1.58E-42	1.47E-40
ENSG00000119973.6	16.20034	-2.667240882	0.676352	-3.94357	8.03E-05	0.000547
ENSG00000196132.14	91.44231	-2.645164897	0.264038	-10.0181	1.27E-23	4.98E-22
ENSG00000233930.4	59.83503	-2.640300978	0.322362	-8.19049	2.6E-16	6.37E-15
ENSG00000165646.14	29.33916	-2.635004476	0.463482	-5.68524	1.31E-08	1.63E-07
ENSG00000171551.12	4945.002	-2.631777629	0.102819	-25.5962	1.7E-144	1.8E-141
ENSG00000152936.11	38.35577	-2.62770943	0.394347	-6.66345	2.67E-11	4.5E-10
ENSG00000162433.15	258.4512	-2.621810739	0.177138	-14.8009	1.45E-49	1.87E-47
ENSG00000151632.18	10.40518	-2.619079628	0.775616	-3.37677	0.000733	0.003928
ENSG00000197977.4	216.0507	-2.613143155	0.184286	-14.1798	1.22E-45	1.33E-43
ENSG00000112137.19	63.453	-2.611188221	0.330762	-7.89445	2.92E-15	6.66E-14
ENSG00000197465.14	10.15702	-2.610790221	0.75404	-3.4624	0.000535	0.00297
ENSG00000167244.22	1015.792	-2.605261976	0.315133	-8.26718	1.37E-16	3.46E-15
ENSG00000078549.16	29.91995	-2.599664009	0.456961	-5.68903	1.28E-08	1.6E-07
ENSG00000167767.14	76.05934	-2.599327223	0.286106	-9.08518	1.04E-19	3.27E-18
ENSG00000101331.17	5.631474	-2.58457872	1.047168	-2.46816	0.013581	0.048584
ENSG00000166257.10	197.928	-2.58249293	0.22399	-11.5295	9.37E-31	5.29E-29
ENSG00000205426.11	13.09384	-2.576535026	0.763548	-3.37442	0.00074	0.003953
ENSG00000285822.1	90.90979	-2.576383768	0.295944	-8.70564	3.16E-18	8.95E-17
ENSG00000119714.11	78.44471	-2.575628397	0.290823	-8.85633	8.27E-19	2.45E-17
ENSG00000112769.20	375.9089	-2.569841185	0.162037	-15.8596	1.21E-56	1.9E-54
ENSG00000182459.5	6.236885	-2.568004458	1.01297	-2.53512	0.011241	0.04129
ENSG00000287600.1	6.205667	-2.565643122	0.982342	-2.61176	0.009008	0.034366
ENSG00000187122.17	12.46511	-2.56157914	0.721392	-3.55088	0.000384	0.002213
ENSG00000283213.1	6.726043	-2.561547452	0.95126	-2.69279	0.007086	0.028136
ENSG00000186777.13	31.76202	-2.559315536	0.471402	-5.42916	5.66E-08	6.59E-07
ENSG00000225746.13	6.161093	-2.558093185	0.994575	-2.57205	0.01011	0.037816
ENSG00000236056.1	12.01148	-2.554860756	0.689275	-3.70659	0.00021	0.001299
ENSG00000205336.14	8.509992	-2.531748919	0.847494	-2.98734	0.002814	0.012728
ENSG00000170442.12	11.76825	-2.52992969	0.716805	-3.52946	0.000416	0.002376
ENSG00000263874.3	62.04876	-2.529365696	0.348618	-7.25541	4E-13	7.83E-12
ENSG00000118160.14	96.87141	-2.520352043	0.253438	-9.94465	2.66E-23	1.02E-21
ENSG00000127824.15	7.011413	-2.517713986	0.959603	-2.6237	0.008698	0.033376
ENSG00000185551.15	52.15347	-2.513324945	0.344857	-7.28802	3.15E-13	6.2E-12
ENSG00000179292.5	106.2322	-2.512853043	0.243829	-10.3058	6.63E-25	2.78E-23
ENSG00000041353.10	215.1471	-2.507584536	0.227575	-11.0187	3.1E-28	1.57E-26
ENSG00000127863.16	374.8411	-2.506552395	0.141703	-17.6888	5.12E-70	1.26E-67

ENSG00000101542.10	19.19666	-2.500369531	0.551476	-4.53396	5.79E-06	4.97E-05
ENSG00000101098.13	356.5143	-2.483593762	0.141293	-17.5776	3.65E-69	8.68E-67
ENSG00000131015.5	6.514742	-2.469214666	0.97347	-2.53651	0.011196	0.041164
ENSG00000026508.21	20.3089	-2.46101114	0.528098	-4.66015	3.16E-06	2.85E-05
ENSG00000171617.16	467.3948	-2.448000467	0.134308	-18.2267	3.17E-74	8.87E-72
ENSG00000185090.15	342.1247	-2.447093118	0.150318	-16.2794	1.38E-59	2.52E-57
ENSG00000148357.18	425.7429	-2.445305964	0.152468	-16.0382	6.92E-58	1.16E-55
ENSG00000204950.4	26.14055	-2.444344143	0.510734	-4.78595	1.7E-06	1.6E-05
ENSG00000135324.6	86.54895	-2.442356589	0.27006	-9.04376	1.51E-19	4.72E-18
ENSG00000136237.20	266.2438	-2.437807246	0.172613	-14.123	2.74E-45	2.9E-43
ENSG00000134438.10	27.38821	-2.436803919	0.460329	-5.29361	1.2E-07	1.34E-06
ENSG00000135824.13	31.66889	-2.432355565	0.436863	-5.56778	2.58E-08	3.14E-07
ENSG00000143858.12	153.8898	-2.422816235	0.230691	-10.5024	8.42E-26	3.76E-24
ENSG00000287431.1	93.69547	-2.420531537	0.275859	-8.77453	1.72E-18	4.97E-17
ENSG00000287051.1	6.214688	-2.417148019	0.953544	-2.53491	0.011248	0.0413
ENSG00000131650.15	47.24738	-2.412370191	0.368948	-6.53851	6.21E-11	1.02E-09
ENSG00000229205.5	91.74494	-2.409959215	0.259799	-9.27624	1.76E-20	5.82E-19
ENSG00000165105.10	122.6612	-2.405255293	0.237229	-10.139	3.71E-24	1.51E-22
ENSG00000075429.9	10.30059	-2.399094754	0.774579	-3.09729	0.001953	0.009281
ENSG00000224126.2	9.710822	-2.389477769	0.766097	-3.11903	0.001814	0.008693
ENSG00000103546.19	1691.541	-2.387026135	0.111576	-21.3938	1.5E-101	8E-99
ENSG00000140481.15	49.29988	-2.377854228	0.343757	-6.91726	4.6E-12	8.35E-11
ENSG00000214391.3	27.24826	-2.373168311	0.457113	-5.19164	2.08E-07	2.25E-06
ENSG00000231466.2	7.734089	-2.35140845	0.871494	-2.69814	0.006973	0.02776
ENSG00000131398.15	74.58888	-2.342584691	0.306692	-7.63822	2.2E-14	4.73E-13
ENSG00000176406.25	77.2904	-2.341013517	0.304811	-7.68023	1.59E-14	3.45E-13
ENSG00000167619.13	90.17325	-2.336701118	0.272379	-8.57887	9.58E-18	2.62E-16
ENSG00000178796.13	15.7289	-2.335540396	0.605736	-3.85571	0.000115	0.000758
ENSG00000141668.10	56.36384	-2.322970783	0.334049	-6.95398	3.55E-12	6.47E-11
ENSG00000144218.21	92.66739	-2.322215494	0.279913	-8.2962	1.07E-16	2.74E-15
ENSG00000075043.21	4141.446	-2.306357697	0.08969	-25.7148	8E-146	9.1E-143
ENSG00000269951.1	9.102033	-2.306254514	0.816657	-2.82402	0.004743	0.020019
ENSG00000171631.16	33.94409	-2.303732194	0.435549	-5.28926	1.23E-07	1.37E-06
ENSG00000109832.14	31.62909	-2.300602784	0.49399	-4.65719	3.21E-06	2.88E-05
ENSG00000139445.18	41.82743	-2.300486605	0.386817	-5.94722	2.73E-09	3.71E-08
ENSG00000127074.15	121.6974	-2.292739592	0.273882	-8.37127	5.7E-17	1.48E-15
ENSG00000182901.18	41.54771	-2.287627151	0.382663	-5.97817	2.26E-09	3.1E-08
ENSG00000170419.11	40.75527	-2.287459134	0.37624	-6.07978	1.2E-09	1.7E-08
ENSG00000130513.7	61.61727	-2.279733068	0.317194	-7.18719	6.61E-13	1.27E-11
ENSG00000183654.9	57.92604	-2.277779189	0.320854	-7.09912	1.26E-12	2.37E-11
ENSG00000128713.14	66.71066	-2.276607795	0.29811	-7.63681	2.23E-14	4.78E-13
ENSG00000119283.16	11.41676	-2.273985086	0.722828	-3.14596	0.001655	0.008023
ENSG00000102003.12	410.9194	-2.272406871	0.136043	-16.7036	1.23E-62	2.55E-60
ENSG00000248127.1	12.85666	-2.264559221	0.649458	-3.48684	0.000489	0.002746
ENSG00000112343.11	26.3013	-2.263614371	0.454942	-4.97561	6.5E-07	6.6E-06
ENSG00000184545.11	424.8588	-2.261209767	0.154739	-14.6131	2.32E-48	2.84E-46

ENSG00000104112.9	569.5883	-2.260747622	0.12709	-17.7886	8.66E-71	2.21E-68
ENSG00000152377.15	659.0959	-2.247416868	0.115396	-19.4756	1.77E-84	6.02E-82
ENSG00000271524.1	7.576114	-2.247332736	0.900612	-2.49534	0.012584	0.045511
ENSG00000254854.1	37.89551	-2.246377712	0.389753	-5.76359	8.23E-09	1.06E-07
ENSG00000101292.8	8.708411	-2.244503027	0.819346	-2.73938	0.006155	0.024932
ENSG00000237758.1	9.303616	-2.24447258	0.804835	-2.78874	0.005291	0.021928
ENSG00000110723.12	39.55321	-2.233714264	0.396512	-5.63341	1.77E-08	2.17E-07
ENSG00000161940.11	12.76886	-2.231933251	0.656409	-3.40022	0.000673	0.003645
ENSG00000224945.1	10.04194	-2.230699949	0.740024	-3.01436	0.002575	0.011817
ENSG00000265282.1	77.1002	-2.227475616	0.279109	-7.98068	1.46E-15	3.4E-14
ENSG00000293394.1	7.393551	-2.22285111	0.879397	-2.5277	0.011481	0.041987
ENSG00000103723.17	345.0154	-2.217129126	0.158813	-13.9606	2.71E-44	2.73E-42
ENSG00000118513.21	94.30656	-2.204893636	0.282156	-7.81444	5.52E-15	1.24E-13
ENSG00000145147.20	434.1966	-2.201341595	0.146011	-15.0766	2.31E-51	3.25E-49
ENSG00000198134.3	9.48575	-2.200786338	0.7518	-2.92735	0.003419	0.015088
ENSG00000118733.17	51.3768	-2.196057131	0.328262	-6.68995	2.23E-11	3.79E-10
ENSG00000152894.15	826.7491	-2.19402465	0.107566	-20.3971	1.77E-92	7.56E-90
ENSG00000290803.1	43.48755	-2.192344787	0.372928	-5.87873	4.13E-09	5.54E-08
ENSG00000164303.11	17.65248	-2.188650443	0.55175	-3.96674	7.29E-05	0.000501
ENSG00000184144.12	10.74437	-2.174937376	0.728442	-2.98574	0.002829	0.012792
ENSG00000143375.15	30.53213	-2.173298367	0.437518	-4.96733	6.79E-07	6.87E-06
ENSG00000164076.17	62.01667	-2.171012717	0.310262	-6.99734	2.61E-12	4.82E-11
ENSG00000189057.11	350.5085	-2.167751577	0.140857	-15.3897	1.92E-53	2.8E-51
ENSG00000134871.19	1065.586	-2.164885175	0.097076	-22.3009	3.6E-110	2.2E-107
ENSG00000065618.21	24.68066	-2.158421666	0.510092	-4.23143	2.32E-05	0.000178
ENSG00000144488.15	117.4919	-2.155995425	0.224794	-9.59097	8.73E-22	3.08E-20
ENSG00000007171.19	20.50782	-2.153464376	0.543121	-3.96498	7.34E-05	0.000505
ENSG00000157851.17	2332.213	-2.15152558	0.089806	-23.9574	7.7E-127	5.7E-124
ENSG00000231638.1	9.498275	-2.148563121	0.774162	-2.77534	0.005514	0.0227
ENSG00000110092.4	12641.34	-2.142828527	0.074527	-28.7524	8.5E-182	1.9E-178
ENSG00000256663.1	38.58396	-2.133822797	0.386543	-5.52027	3.38E-08	4.05E-07
ENSG00000183307.4	60.25268	-2.131380225	0.316596	-6.73218	1.67E-11	2.87E-10
ENSG00000127564.17	345.9472	-2.12672139	0.153681	-13.8386	1.49E-43	1.47E-41
ENSG00000099998.19	10.42956	-2.126440608	0.793652	-2.67931	0.007377	0.029125
ENSG00000101057.16	1867.254	-2.115913289	0.096555	-21.914	1.9E-106	1.1E-103
ENSG00000288459.1	80.66282	-2.110287034	0.278049	-7.58963	3.21E-14	6.79E-13
ENSG00000091879.14	14.97991	-2.107166744	0.597057	-3.52925	0.000417	0.002376
ENSG00000116981.4	62.47755	-2.105409742	0.294205	-7.15627	8.29E-13	1.58E-11
ENSG00000129910.8	34.41624	-2.103406121	0.401836	-5.23449	1.65E-07	1.81E-06
ENSG00000101412.13	718.3269	-2.099965042	0.105681	-19.8708	7.28E-88	2.92E-85
ENSG00000225077.3	8.391326	-2.085119734	0.793511	-2.62771	0.008596	0.033024
ENSG00000241345.1	11.62675	-2.083550939	0.66983	-3.11057	0.001867	0.008913
ENSG00000260802.2	24.30795	-2.081672175	0.467439	-4.45335	8.45E-06	7.02E-05
ENSG00000133454.16	16.76523	-2.078205874	0.560792	-3.70584	0.000211	0.001301
ENSG00000128594.8	20.51641	-2.07799527	0.506485	-4.10278	4.08E-05	0.000297
ENSG00000116014.10	78.40295	-2.077886574	0.273002	-7.61124	2.71E-14	5.78E-13

ENSG00000078114.19	4521.374	-2.077260472	0.06759	-30.733	2.1E-207	7E-204
ENSG00000103056.12	17.51421	-2.072312277	0.568249	-3.64684	0.000265	0.001603
ENSG00000122012.14	307.0533	-2.070185326	0.152508	-13.5743	5.69E-42	5.12E-40
ENSG00000137868.19	1408.107	-2.066564028	0.091432	-22.6022	4.1E-113	2.7E-110
ENSG00000266236.1	10.84649	-2.060892533	0.697325	-2.95543	0.003122	0.013943
ENSG00000086159.14	13.54653	-2.057522596	0.686517	-2.99704	0.002726	0.012393
ENSG00000119922.11	12.94059	-2.026433411	0.681847	-2.97197	0.002959	0.0133
ENSG00000255571.10	25.55969	-2.022905503	0.554261	-3.64974	0.000263	0.001586
ENSG00000101445.10	514.52	-2.018958652	0.121003	-16.6852	1.68E-62	3.43E-60
ENSG00000199572.1	27.95098	-2.015883506	0.462957	-4.35436	1.33E-05	0.000107
ENSG00000291130.1	183.4024	-2.014764088	0.185474	-10.8628	1.73E-27	8.46E-26
ENSG00000173826.15	34.63977	-2.013741222	0.422851	-4.76229	1.91E-06	1.79E-05
ENSG00000179348.13	4466.654	-2.003937279	0.081412	-24.6148	8.8E-134	6.9E-131
ENSG00000253457.4	39.91586	-1.999378108	0.434676	-4.5997	4.23E-06	3.73E-05
ENSG00000164626.9	39.32902	-1.994388346	0.404477	-4.93079	8.19E-07	8.14E-06
ENSG00000224885.1	10.78003	-1.991091431	0.715285	-2.78363	0.005375	0.022225
ENSG00000171094.18	1147.763	-1.989563628	0.113518	-17.5264	9.02E-69	2.12E-66
ENSG00000169744.13	32.37845	-1.984387995	0.406752	-4.87862	1.07E-06	1.04E-05
ENSG00000188985.6	11.61527	-1.980948204	0.657734	-3.01178	0.002597	0.011897
ENSG00000282390.1	640.3118	-1.976731007	0.106424	-18.5741	5.21E-77	1.54E-74
ENSG00000007968.7	441.158	-1.970789432	0.131957	-14.9351	1.95E-50	2.64E-48
ENSG00000152969.21	16.45912	-1.97017358	0.60791	-3.24089	0.001192	0.006027
ENSG00000159409.15	724.1909	-1.959166781	0.106664	-18.3676	2.39E-75	6.86E-73
ENSG00000171951.5	3605.492	-1.955868902	0.075958	-25.7492	3.3E-146	4E-143
ENSG00000272405.1	32.92301	-1.952986037	0.412259	-4.73728	2.17E-06	2.01E-05
ENSG00000133958.14	166.1613	-1.951042855	0.213847	-9.12355	7.27E-20	2.31E-18
ENSG00000143333.7	411.4149	-1.940812038	0.134733	-14.4049	4.82E-47	5.63E-45
ENSG00000100479.14	147.3245	-1.940055748	0.214757	-9.03371	1.66E-19	5.14E-18
ENSG00000232977.7	23.80328	-1.931224673	0.466284	-4.14174	3.45E-05	0.000255
ENSG00000249741.2	13.94993	-1.925770547	0.609368	-3.16028	0.001576	0.007707
ENSG00000093009.11	415.0442	-1.917637092	0.128906	-14.8762	4.7E-50	6.2E-48
ENSG00000277586.4	1714.019	-1.914190513	0.083281	-22.9848	6.6E-117	4.5E-114
ENSG00000168243.11	2850.472	-1.90999005	0.07394	-25.8317	3.9E-147	5.3E-144
ENSG00000198879.13	536.2189	-1.909976236	0.135993	-14.0446	8.31E-45	8.62E-43
ENSG00000178403.4	346.2192	-1.902021817	0.379042	-5.01797	5.22E-07	5.36E-06
ENSG00000126259.20	29.56842	-1.885210761	0.459945	-4.09878	4.15E-05	0.000302
ENSG00000135519.8	11.03501	-1.884736315	0.694184	-2.71504	0.006627	0.026605
ENSG00000131378.14	1488.038	-1.883661851	0.08871	-21.234	4.6E-100	2.37E-97
ENSG00000171450.6	992.6908	-1.87728163	0.106158	-17.6838	5.59E-70	1.36E-67
ENSG00000065328.17	557.8566	-1.871758633	0.129539	-14.4494	2.53E-47	3E-45
ENSG00000028277.22	11.12104	-1.87014214	0.693128	-2.69812	0.006973	0.02776
ENSG00000072832.15	2229.547	-1.861365086	0.079261	-23.4839	6E-122	4.2E-119
ENSG00000134333.15	6356.495	-1.856946889	0.072097	-25.7563	2.7E-146	3.5E-143
ENSG00000100526.21	373.235	-1.852185738	0.164724	-11.2442	2.47E-29	1.31E-27
ENSG00000169258.7	953.845	-1.848059474	0.12481	-14.807	1.32E-49	1.72E-47
ENSG00000084628.10	437.2849	-1.84708048	0.142032	-13.0047	1.15E-38	9.01E-37

ENSG00000236671.9	76.21055	-1.846621755	0.273953	-6.74064	1.58E-11	2.72E-10
ENSG00000158748.4	71.72098	-1.845393589	0.310107	-5.95083	2.67E-09	3.63E-08
ENSG00000177108.5	11.95405	-1.844683138	0.683786	-2.69775	0.006981	0.027785
ENSG00000078018.22	887.9466	-1.843435313	0.112605	-16.3707	3.09E-60	5.75E-58
ENSG00000143476.18	1121.55	-1.838817402	0.094989	-19.3583	1.74E-83	5.81E-81
ENSG00000125355.16	29.34851	-1.838681531	0.424457	-4.33184	1.48E-05	0.000117
ENSG00000163888.4	112.434	-1.835801792	0.23688	-7.74993	9.19E-15	2.04E-13
ENSG00000094804.12	732.3254	-1.835668659	0.113536	-16.1682	8.45E-59	1.48E-56
ENSG00000276043.5	1522.392	-1.830416773	0.097987	-18.6803	7.16E-78	2.22E-75
ENSG00000171848.16	1367.508	-1.828215212	0.098305	-18.5974	3.37E-77	1.01E-74
ENSG00000293057.1	13.05968	-1.822863769	0.620167	-2.93931	0.003289	0.014596
ENSG00000131153.9	366.0152	-1.821644683	0.135586	-13.4353	3.75E-41	3.29E-39
ENSG00000146592.17	268.5564	-1.820513338	0.181252	-10.0441	9.75E-24	3.86E-22
ENSG00000099256.19	80.55332	-1.81636714	0.2888	-6.28936	3.19E-10	4.79E-09
ENSG00000129159.9	117.3499	-1.815085914	0.215354	-8.42839	3.5E-17	9.26E-16
ENSG00000078725.13	38.72233	-1.814436902	0.372586	-4.86985	1.12E-06	1.08E-05
ENSG00000077152.12	1562.662	-1.813665704	0.122027	-14.8628	5.75E-50	7.53E-48
ENSG00000236603.2	17.75603	-1.807614662	0.558326	-3.23756	0.001206	0.006087
ENSG00000178752.16	275.8548	-1.80755161	0.157871	-11.4496	2.36E-30	1.3E-28
ENSG00000224888.4	41.52066	-1.800043451	0.358494	-5.02113	5.14E-07	5.28E-06
ENSG00000133863.9	70.40454	-1.798551411	0.272706	-6.5952	4.25E-11	7.03E-10
ENSG00000087258.16	397.5536	-1.791830145	0.135698	-13.2046	8.26E-40	6.8E-38
ENSG00000152932.8	329.3213	-1.7897507	0.163408	-10.9526	6.45E-28	3.19E-26
ENSG00000135144.8	172.3653	-1.788485854	0.179403	-9.96908	2.08E-23	8.06E-22
ENSG00000164109.14	1915.3	-1.788342237	0.134526	-13.2937	2.52E-40	2.13E-38
ENSG00000085840.13	252.6084	-1.781984328	0.176847	-10.0764	7.03E-24	2.8E-22
ENSG00000068615.20	39.31131	-1.779532312	0.395773	-4.49634	6.91E-06	5.84E-05
ENSG00000103489.12	1085.644	-1.776793291	0.111692	-15.908	5.58E-57	8.91E-55
ENSG00000144354.15	442.0196	-1.775361334	0.128486	-13.8176	2E-43	1.95E-41
ENSG00000158246.8	34.50287	-1.774800661	0.388584	-4.56736	4.94E-06	4.29E-05
ENSG00000196353.13	15.91626	-1.773372432	0.572988	-3.09496	0.001968	0.009337
ENSG00000112715.26	390.5785	-1.772929251	0.131114	-13.522	1.16E-41	1.04E-39
ENSG00000165244.7	352.0636	-1.772349439	0.135202	-13.1089	2.93E-39	2.33E-37
ENSG00000152256.14	507.6876	-1.77167302	0.13393	-13.2283	6.02E-40	5E-38
ENSG00000124302.13	133.2949	-1.76882638	0.202699	-8.72635	2.63E-18	7.52E-17
ENSG00000237289.10	35.65563	-1.765025971	0.380019	-4.64458	3.41E-06	3.04E-05
ENSG00000051180.17	283.2666	-1.76148537	0.154206	-11.4229	3.21E-30	1.77E-28
ENSG00000138166.6	44.23781	-1.759803789	0.374334	-4.70115	2.59E-06	2.36E-05
ENSG00000162073.14	381.8653	-1.755838272	0.14598	-12.028	2.53E-33	1.59E-31
ENSG00000106327.13	100.6603	-1.751093888	0.231717	-7.55704	4.12E-14	8.64E-13
ENSG00000185559.16	2354.688	-1.748887799	0.118946	-14.7032	6.15E-49	7.66E-47
ENSG00000291174.1	19.83253	-1.748665905	0.571548	-3.05952	0.002217	0.010352
ENSG00000146670.10	1036.142	-1.746875926	0.097766	-17.868	2.1E-71	5.42E-69
ENSG00000107105.15	341.5204	-1.743696768	0.141527	-12.3206	7.02E-35	4.75E-33
ENSG00000181215.17	31.37198	-1.738227318	0.395733	-4.39243	1.12E-05	9.12E-05
ENSG00000091129.22	1825.82	-1.734676766	0.108396	-16.0031	1.22E-57	2.02E-55

ENSG00000168078.10	815.7072	-1.734604807	0.123828	-14.0081	1.39E-44	1.41E-42
ENSG00000196358.12	49.64656	-1.732721983	0.336113	-5.15518	2.53E-07	2.72E-06
ENSG00000290615.1	44.58747	-1.731181269	0.361071	-4.79457	1.63E-06	1.54E-05
ENSG00000107984.10	6760.152	-1.731089226	0.091483	-18.9226	7.43E-80	2.37E-77
ENSG00000171126.8	22.54365	-1.729728852	0.480676	-3.59853	0.00032	0.001886
ENSG00000183780.13	26.10856	-1.727712834	0.520424	-3.31982	0.000901	0.004714
ENSG00000151725.12	998.5831	-1.723663161	0.122412	-14.0808	4.98E-45	5.25E-43
ENSG00000253317.1	34.76058	-1.723061234	0.400247	-4.305	1.67E-05	0.000131
ENSG00000234224.3	1357.691	-1.722520361	0.087171	-19.7602	6.55E-87	2.53E-84
ENSG00000158747.15	1956.821	-1.722132537	0.082827	-20.7919	5.12E-96	2.38E-93
ENSG00000224807.5	73.19838	-1.718758132	0.285994	-6.00977	1.86E-09	2.57E-08
ENSG00000075461.6	3831.306	-1.71522546	0.091255	-18.7959	8.16E-79	2.56E-76
ENSG00000164647.9	18.80509	-1.714362446	0.547865	-3.12917	0.001753	0.008427
ENSG00000143228.13	1082.654	-1.713933441	0.125461	-13.6611	1.73E-42	1.61E-40
ENSG00000230002.3	10.92612	-1.71165684	0.678241	-2.52367	0.011614	0.042387
ENSG00000140323.6	115.8704	-1.709665099	0.228207	-7.49171	6.8E-14	1.4E-12
ENSG00000174371.17	694.6269	-1.703177331	0.113323	-15.0294	4.71E-51	6.6E-49
ENSG00000051341.15	200.1492	-1.702555895	0.176918	-9.62344	6.37E-22	2.27E-20
ENSG00000073111.14	1476.977	-1.701814031	0.098357	-17.3023	4.52E-67	1.03E-64
ENSG00000145861.9	34.37865	-1.70115657	0.41619	-4.08745	4.36E-05	0.000316
ENSG00000255650.6	125.6891	-1.697339475	0.214147	-7.92604	2.26E-15	5.2E-14
ENSG00000176641.11	854.0712	-1.696930106	0.109837	-15.4495	7.6E-54	1.13E-51
ENSG00000204624.8	36.16492	-1.695876704	0.375587	-4.51527	6.32E-06	5.38E-05
ENSG00000175497.17	49.76032	-1.694533486	0.334432	-5.0669	4.04E-07	4.23E-06
ENSG00000170312.17	1944.944	-1.691234803	0.237057	-7.13429	9.73E-13	1.85E-11
ENSG00000178999.13	402.1819	-1.690826012	0.128964	-13.1108	2.86E-39	2.28E-37
ENSG00000176890.16	664.1113	-1.686950593	0.104621	-16.1244	1.72E-58	2.98E-56
ENSG00000231607.15	82.10043	-1.684825443	0.273762	-6.15435	7.54E-10	1.09E-08
ENSG00000149926.14	259.2792	-1.684220849	0.162669	-10.3537	4.03E-25	1.74E-23
ENSG00000196584.4	417.1601	-1.679016531	0.122098	-13.7514	4.99E-43	4.75E-41
ENSG00000126500.4	44.86986	-1.678415091	0.372548	-4.50523	6.63E-06	5.62E-05
ENSG00000186193.9	545.3954	-1.678091867	0.118549	-14.1553	1.73E-45	1.88E-43
ENSG00000253305.3	11.11059	-1.676489776	0.682143	-2.45768	0.013984	0.049738
ENSG00000129810.15	171.5337	-1.676075218	0.183462	-9.13584	6.49E-20	2.08E-18
ENSG00000167900.12	1067.667	-1.674159762	0.094309	-17.7518	1.67E-70	4.21E-68
ENSG00000165480.16	318.9985	-1.670085463	0.145335	-11.4913	1.46E-30	8.15E-29
ENSG00000109805.10	854.8241	-1.667328074	0.099238	-16.8013	2.39E-63	4.98E-61
ENSG00000204653.10	29.1593	-1.666072336	0.438703	-3.79772	0.000146	0.000938
ENSG00000174279.5	40.49043	-1.6659833	0.394914	-4.2186	2.46E-05	0.000187
ENSG00000280079.1	36.99733	-1.66286332	0.379672	-4.37974	1.19E-05	9.6E-05
ENSG00000101447.15	608.4665	-1.655848642	0.113625	-14.5729	4.18E-48	5.08E-46
ENSG00000134198.10	130.8388	-1.655501765	0.228165	-7.25571	4E-13	7.82E-12
ENSG00000258815.1	22.89022	-1.653466298	0.469401	-3.5225	0.000427	0.00243
ENSG00000269397.1	30.79853	-1.652359533	0.409482	-4.03525	5.45E-05	0.000386
ENSG00000164045.12	552.4641	-1.650339378	0.1153	-14.3135	1.8E-46	2.07E-44
ENSG00000103257.9	690.7379	-1.650109248	0.13468	-12.2521	1.64E-34	1.08E-32

ENSG00000165891.16	322.8215	-1.642374163	0.155208	-10.5818	3.62E-26	1.65E-24
ENSG00000287872.1	15.83877	-1.641829729	0.562284	-2.91993	0.003501	0.015395
ENSG00000092853.14	684.3087	-1.637574883	0.128668	-12.7271	4.18E-37	3.08E-35
ENSG00000167513.9	556.2757	-1.636159269	0.116434	-14.0522	7.47E-45	7.79E-43
ENSG00000180875.5	198.2646	-1.635775352	0.176776	-9.2534	2.17E-20	7.16E-19
ENSG00000130287.14	24.78884	-1.631904051	0.449683	-3.62901	0.000285	0.001706
ENSG00000203995.11	60.18053	-1.631069338	0.295442	-5.52078	3.38E-08	4.04E-07
ENSG00000145569.6	603.9769	-1.626117333	0.129764	-12.5314	5.03E-36	3.52E-34
ENSG00000154146.13	184.1587	-1.624854619	0.18331	-8.86396	7.72E-19	2.29E-17
ENSG00000061337.16	293.7462	-1.623996228	0.152399	-10.6562	1.63E-26	7.61E-25
ENSG00000111247.15	492.1937	-1.62380406	0.160679	-10.1059	5.2E-24	2.08E-22
ENSG00000117525.14	121.3415	-1.621244566	0.23772	-6.81996	9.11E-12	1.61E-10
ENSG00000137285.11	4017.703	-1.62085088	0.079161	-20.4754	3.57E-93	1.55E-90
ENSG00000175063.17	989.953	-1.61973149	0.100569	-16.1057	2.33E-58	3.96E-56
ENSG00000166803.14	303.125	-1.617159893	0.171972	-9.40363	5.27E-21	1.8E-19
ENSG00000164362.21	45.64303	-1.610484055	0.357652	-4.50294	6.7E-06	5.68E-05
ENSG00000198056.16	344.3249	-1.610094457	0.155103	-10.3808	3.03E-25	1.32E-23
ENSG00000133401.16	29.51753	-1.606024896	0.410527	-3.91211	9.15E-05	0.000615
ENSG00000155974.14	169.8901	-1.604193719	0.207124	-7.7451	9.55E-15	2.11E-13
ENSG00000131094.4	197.2128	-1.600966285	0.166565	-9.61169	7.14E-22	2.55E-20
ENSG00000102384.14	361.6282	-1.598378457	0.134481	-11.8855	1.41E-32	8.64E-31
ENSG00000158164.7	935.3019	-1.59570061	0.142731	-11.1798	5.12E-29	2.66E-27
ENSG00000197299.13	390.5159	-1.59409394	0.128878	-12.369	3.84E-35	2.62E-33
ENSG00000132692.19	193.5517	-1.592176218	0.171495	-9.28409	1.63E-20	5.42E-19
ENSG00000126787.13	1247.298	-1.591602277	0.117845	-13.5059	1.44E-41	1.29E-39
ENSG00000130558.20	168.5333	-1.590919449	0.193252	-8.23237	1.84E-16	4.58E-15
ENSG00000177602.5	123.0586	-1.587831861	0.212697	-7.46523	8.32E-14	1.71E-12
ENSG00000163497.3	1130.417	-1.58757959	0.095371	-16.6463	3.22E-62	6.51E-60
ENSG00000089685.15	1986.485	-1.583744982	0.079866	-19.8301	1.64E-87	6.43E-85
ENSG00000224411.3	47.7694	-1.579661444	0.34881	-4.52871	5.93E-06	5.08E-05
ENSG00000106537.8	437.5695	-1.577543012	0.122223	-12.9071	4.11E-38	3.16E-36
ENSG00000133101.10	202.197	-1.573619986	0.183029	-8.59763	8.14E-18	2.24E-16
ENSG00000113083.15	50.271	-1.573283059	0.327492	-4.80404	1.55E-06	1.48E-05
ENSG00000115963.14	1076.922	-1.57258719	0.136894	-11.4876	1.52E-30	8.48E-29
ENSG00000152253.9	389.3475	-1.570295986	0.133116	-11.7965	4.07E-32	2.47E-30
ENSG00000289840.1	62.93142	-1.567417291	0.29443	-5.32356	1.02E-07	1.15E-06
ENSG00000177432.8	932.5317	-1.567274133	0.096357	-16.2653	1.74E-59	3.15E-57
ENSG00000075643.6	576.879	-1.562361773	0.113915	-13.7152	8.24E-43	7.76E-41
ENSG00000136997.22	51.94286	-1.560956797	0.318992	-4.89341	9.91E-07	9.68E-06
ENSG00000100297.16	1326.555	-1.559872207	0.089092	-17.5086	1.23E-68	2.86E-66
ENSG00000122952.17	841.1556	-1.552039903	0.101059	-15.3578	3.14E-53	4.55E-51
ENSG00000123219.13	397.3542	-1.551192643	0.158644	-9.77779	1.4E-22	5.18E-21
ENSG00000165304.8	995.1443	-1.550022869	0.093622	-16.5561	1.45E-61	2.82E-59
ENSG00000138346.16	508.5784	-1.549046178	0.123369	-12.5562	3.68E-36	2.6E-34
ENSG00000203877.9	17.25801	-1.548543678	0.553542	-2.79752	0.00515	0.021423
ENSG00000276672.1	25.94194	-1.548274075	0.435374	-3.55619	0.000376	0.002173

ENSG00000188290.11	27.71639	-1.541696919	0.419451	-3.67551	0.000237	0.001445
ENSG00000164082.15	18.13342	-1.540564604	0.548719	-2.80756	0.004992	0.020873
ENSG00000100749.9	375.4485	-1.539997761	0.162231	-9.4926	2.25E-21	7.83E-20
ENSG00000144452.15	99.23808	-1.53723679	0.253914	-6.05416	1.41E-09	1.98E-08
ENSG00000123416.15	7061.778	-1.536866454	0.088127	-17.4391	4.16E-68	9.56E-66
ENSG00000135069.14	738.0004	-1.536597768	0.102582	-14.9792	1E-50	1.37E-48
ENSG00000131351.15	201.2954	-1.532686751	0.167585	-9.14574	5.92E-20	1.9E-18
ENSG00000185168.7	19.71926	-1.532627865	0.538891	-2.84404	0.004455	0.01894
ENSG00000142731.11	727.0227	-1.529857935	0.12694	-12.0518	1.9E-33	1.19E-31
ENSG00000178233.18	426.174	-1.526451912	0.124004	-12.3097	8.03E-35	5.38E-33
ENSG00000065911.13	911.9144	-1.52639379	0.1191	-12.8161	1.33E-37	9.98E-36
ENSG00000286456.3	85.63016	-1.524492587	0.248497	-6.13484	8.52E-10	1.23E-08
ENSG00000187800.14	492.5324	-1.521683927	0.120849	-12.5916	2.35E-36	1.67E-34
ENSG00000198901.14	36.49446	-1.521550084	0.378947	-4.0152	5.94E-05	0.000417
ENSG00000177679.16	332.9959	-1.521314243	0.148252	-10.2616	1.05E-24	4.38E-23
ENSG00000144406.19	105.9856	-1.519539819	0.244024	-6.22702	4.75E-10	7.02E-09
ENSG00000112742.10	828.6849	-1.518307207	0.150164	-10.111	4.94E-24	1.99E-22
ENSG00000174521.8	56.58638	-1.515882603	0.308596	-4.91219	9.01E-07	8.9E-06
ENSG00000106089.12	278.4593	-1.515523776	0.17522	-8.64927	5.18E-18	1.45E-16
ENSG00000167553.17	26.03704	-1.514694616	0.43339	-3.49499	0.000474	0.002669
ENSG00000167670.16	794.2469	-1.51066535	0.114882	-13.1497	1.71E-39	1.38E-37
ENSG00000168496.4	1749.665	-1.510495195	0.085438	-17.6793	6.05E-70	1.45E-67
ENSG00000007237.19	102.8897	-1.508696015	0.232947	-6.47656	9.38E-11	1.51E-09
ENSG00000070882.13	161.6424	-1.50803295	0.18807	-8.01846	1.07E-15	2.52E-14
ENSG00000184113.10	23.78567	-1.5071315	0.475671	-3.16843	0.001533	0.007528
ENSG00000149636.16	477.0696	-1.505358622	0.124811	-12.0611	1.7E-33	1.07E-31
ENSG00000135048.14	41.29281	-1.504546948	0.342658	-4.39081	1.13E-05	9.18E-05
ENSG00000164683.19	900.2649	-1.504106074	0.096272	-15.6235	5.04E-55	7.8E-53
ENSG00000127423.11	66.53427	-1.503920571	0.282641	-5.32096	1.03E-07	1.16E-06
ENSG00000165490.13	399.1827	-1.502847886	0.133171	-11.2851	1.55E-29	8.27E-28
ENSG00000066032.19	353.3905	-1.499860466	0.140396	-10.683	1.22E-26	5.77E-25
ENSG00000185811.21	76.30683	-1.499673365	0.261807	-5.72817	1.02E-08	1.29E-07
ENSG00000282381.2	17.34391	-1.49829175	0.541656	-2.76613	0.005673	0.023273
ENSG00000085999.13	198.7925	-1.497114087	0.183567	-8.15568	3.47E-16	8.45E-15
ENSG00000148082.10	400.3812	-1.497002505	0.13762	-10.8778	1.47E-27	7.19E-26
ENSG00000106034.18	1002.113	-1.489914247	0.091882	-16.2154	3.92E-59	6.91E-57
ENSG00000203760.9	296.588	-1.488977498	0.182935	-8.1394	3.97E-16	9.64E-15
ENSG00000112312.10	780.0196	-1.485664293	0.127973	-11.6092	3.7E-31	2.16E-29
ENSG00000005020.13	102.8487	-1.485556698	0.231317	-6.42218	1.34E-10	2.12E-09
ENSG00000104738.19	2331.127	-1.483573723	0.082375	-18.0101	1.62E-72	4.42E-70
ENSG00000244953.1	78.20182	-1.483386039	0.258895	-5.72968	1.01E-08	1.28E-07
ENSG00000105011.9	492.1025	-1.481571168	0.122716	-12.0732	1.46E-33	9.29E-32
ENSG00000170525.21	650.4111	-1.48148528	0.110973	-13.3499	1.19E-40	1.02E-38
ENSG00000168280.18	974.4566	-1.479361922	0.107413	-13.7727	3.72E-43	3.57E-41
ENSG00000138180.16	746.4506	-1.479228956	0.105489	-14.0225	1.13E-44	1.17E-42
ENSG00000056972.21	240.0318	-1.478722808	0.17166	-8.61426	7.04E-18	1.95E-16

ENSG00000132646.11	2532.634	-1.476490328	0.082564	-17.883	1.6E-71	4.24E-69
ENSG00000183850.14	72.79864	-1.475402367	0.273762	-5.38936	7.07E-08	8.1E-07
ENSG00000130203.10	187.8183	-1.473988781	0.203507	-7.24293	4.39E-13	8.56E-12
ENSG00000109255.11	164.3595	-1.472339419	0.190916	-7.71196	1.24E-14	2.72E-13
ENSG00000129195.16	658.876	-1.468550961	0.124964	-11.7518	6.92E-32	4.16E-30
ENSG00000060656.20	1523.313	-1.468098255	0.093596	-15.6855	1.9E-55	2.97E-53
ENSG00000166292.12	27.77933	-1.466493449	0.428581	-3.42174	0.000622	0.0034
ENSG00000035499.13	356.6897	-1.465163847	0.143567	-10.2054	1.87E-24	7.74E-23
ENSG00000139734.19	466.1547	-1.462617767	0.119437	-12.2459	1.77E-34	1.16E-32
ENSG00000186871.7	317.4545	-1.457535973	0.13909	-10.4791	1.08E-25	4.79E-24
ENSG00000171388.12	23.79535	-1.452007388	0.475201	-3.05557	0.002246	0.010468
ENSG00000125898.13	113.5686	-1.451672757	0.210148	-6.90787	4.92E-12	8.88E-11
ENSG00000146700.9	177.5372	-1.450981094	0.175076	-8.28771	1.15E-16	2.93E-15
ENSG00000161888.12	324.9896	-1.446985511	0.137634	-10.5133	7.5E-26	3.37E-24
ENSG00000197134.13	136.2303	-1.44600103	0.232895	-6.20882	5.34E-10	7.85E-09
ENSG00000072571.20	1107.531	-1.445295268	0.124944	-11.5675	6.02E-31	3.45E-29
ENSG00000123485.12	573.0558	-1.444686869	0.111598	-12.9454	2.49E-38	1.93E-36
ENSG00000171320.15	382.349	-1.441244316	0.129591	-11.1215	9.86E-29	5.1E-27
ENSG00000111846.20	93.66166	-1.441083903	0.245183	-5.87758	4.16E-09	5.57E-08
ENSG00000164032.12	5236.729	-1.43975754	0.241517	-5.9613	2.5E-09	3.41E-08
ENSG00000130540.14	44.00107	-1.438643253	0.35319	-4.07328	4.64E-05	0.000333
ENSG00000142945.13	906.3835	-1.436148141	0.097843	-14.6781	8.9E-49	1.1E-46
ENSG00000278828.2	35.14048	-1.433609702	0.386834	-3.706	0.000211	0.001301
ENSG00000074276.11	49.60024	-1.43154694	0.363762	-3.9354	8.31E-05	0.000564
ENSG00000014138.10	88.23556	-1.429525243	0.24185	-5.9108	3.4E-09	4.59E-08
ENSG00000182010.11	341.2698	-1.42749429	0.131924	-10.8206	2.75E-27	1.33E-25
ENSG00000008735.14	740.7667	-1.426218736	0.104492	-13.6491	2.04E-42	1.88E-40
ENSG00000013810.21	1561.317	-1.425332343	0.087893	-16.2166	3.85E-59	6.84E-57
ENSG00000164611.13	761.9648	-1.42454177	0.12627	-11.2817	1.62E-29	8.58E-28
ENSG00000134690.11	522.3894	-1.424078401	0.1127	-12.636	1.34E-36	9.62E-35
ENSG00000154839.10	259.5869	-1.423967807	0.159725	-8.91511	4.87E-19	1.48E-17
ENSG00000165140.12	32.08702	-1.423310812	0.443582	-3.20868	0.001333	0.006643
ENSG00000139567.13	26.04767	-1.42291272	0.441698	-3.22146	0.001275	0.006395
ENSG00000109674.4	225.8442	-1.422769788	0.16948	-8.39492	4.66E-17	1.22E-15
ENSG00000121152.10	775.3717	-1.42092414	0.100502	-14.1383	2.21E-45	2.37E-43
ENSG00000074966.11	54.78882	-1.420023061	0.358597	-3.95995	7.5E-05	0.000515
ENSG00000237649.8	1058.494	-1.419437579	0.110033	-12.9001	4.5E-38	3.43E-36
ENSG00000242593.8	64.38747	-1.414567338	0.322792	-4.38229	1.17E-05	9.51E-05
ENSG00000154920.15	268.4369	-1.412457245	0.145865	-9.68331	3.55E-22	1.29E-20
ENSG00000075702.19	362.6533	-1.41165687	0.147792	-9.55165	1.28E-21	4.47E-20
ENSG00000119403.15	682.5141	-1.410683774	0.114696	-12.2993	9.13E-35	6.1E-33
ENSG00000116985.12	71.99974	-1.410449109	0.27062	-5.21192	1.87E-07	2.04E-06
ENSG00000117650.13	854.9077	-1.409862429	0.103857	-13.575	5.63E-42	5.09E-40
ENSG00000277496.1	56.76083	-1.407951668	0.310566	-4.53351	5.8E-06	4.97E-05
ENSG00000290382.1	17.02364	-1.407518441	0.550681	-2.55596	0.010589	0.039285
ENSG00000117595.13	23.13655	-1.407140998	0.455483	-3.08934	0.002006	0.0095

ENSG00000121957.15	533.5266	-1.405474041	0.121451	-11.5724	5.69E-31	3.28E-29
ENSG00000272163.2	55.10023	-1.403647118	0.305996	-4.58715	4.49E-06	3.93E-05
ENSG00000278971.1	17.70536	-1.402392938	0.540228	-2.59593	0.009434	0.035717
ENSG00000105825.14	848.3241	-1.401758826	0.101665	-13.7881	3.01E-43	2.91E-41
ENSG00000198554.12	646.9781	-1.401327421	0.111069	-12.6167	1.71E-36	1.22E-34
ENSG00000145386.11	1195.411	-1.400763474	0.09339	-14.9991	7.44E-51	1.03E-48
ENSG00000087586.18	530.243	-1.400459425	0.114102	-12.2738	1.25E-34	8.31E-33
ENSG00000169607.13	535.6904	-1.399854357	0.123454	-11.3391	8.4E-30	4.52E-28
ENSG00000024526.17	774.6099	-1.399287756	0.103686	-13.4954	1.66E-41	1.48E-39
ENSG00000136982.6	329.6999	-1.399098114	0.147896	-9.46002	3.08E-21	1.06E-19
ENSG00000102924.13	25.28738	-1.399067986	0.45371	-3.08361	0.002045	0.009669
ENSG00000235363.1	86.70693	-1.397802385	0.284682	-4.91004	9.11E-07	8.98E-06
ENSG00000162551.14	19.53968	-1.397486092	0.489666	-2.85396	0.004318	0.018435
ENSG00000122641.11	358.8981	-1.396320803	0.138018	-10.1169	4.65E-24	1.87E-22
ENSG00000175305.18	130.4148	-1.395678802	0.221478	-6.30165	2.94E-10	4.48E-09
ENSG00000115163.15	351.193	-1.394820235	0.14104	-9.88954	4.62E-23	1.75E-21
ENSG00000251573.2	17.63487	-1.394786516	0.535639	-2.60397	0.009215	0.035033
ENSG00000179331.3	170.0653	-1.393912735	0.179491	-7.7659	8.11E-15	1.8E-13
ENSG00000187479.9	8882.905	-1.393736317	0.082738	-16.8451	1.14E-63	2.4E-61
ENSG00000133710.17	60.45202	-1.389269124	0.292319	-4.75258	2.01E-06	1.87E-05
ENSG00000148219.18	476.7867	-1.388270065	0.121563	-11.4201	3.32E-30	1.82E-28
ENSG00000167552.15	16745.63	-1.385241667	0.086709	-15.9758	1.89E-57	3.08E-55
ENSG00000138778.14	1038.431	-1.38511914	0.127114	-10.8966	1.2E-27	5.86E-26
ENSG00000076382.17	1046.931	-1.383460347	0.090696	-15.2539	1.55E-52	2.23E-50
ENSG00000176912.7	33.92179	-1.381613991	0.381105	-3.62528	0.000289	0.001729
ENSG00000116852.15	57.59221	-1.380797474	0.292142	-4.72646	2.28E-06	2.11E-05
ENSG00000156970.13	890.4288	-1.378689558	0.10147	-13.5872	4.77E-42	4.33E-40
ENSG00000179431.7	27.16036	-1.377358303	0.441382	-3.12056	0.001805	0.008656
ENSG00000173207.13	1976.195	-1.374423351	0.117941	-11.6535	2.2E-31	1.29E-29
ENSG00000137142.5	2590.011	-1.374135349	0.088416	-15.5417	1.81E-54	2.72E-52
ENSG00000054967.13	148.324	-1.373734418	0.208665	-6.58343	4.6E-11	7.59E-10
ENSG00000197275.14	83.28322	-1.372732686	0.289025	-4.74952	2.04E-06	1.9E-05
ENSG00000183814.16	500.6807	-1.371768517	0.136992	-10.0135	1.33E-23	5.2E-22
ENSG00000196781.16	213.406	-1.371049622	0.167334	-8.19347	2.54E-16	6.23E-15
ENSG00000144554.13	552.2471	-1.369397851	0.129724	-10.5563	4.75E-26	2.16E-24
ENSG00000234423.2	20.75334	-1.369367974	0.492287	-2.78165	0.005408	0.022326
ENSG00000123975.5	740.8965	-1.367760982	0.147765	-9.25631	2.12E-20	6.97E-19
ENSG00000117399.14	1135.217	-1.367673117	0.095607	-14.3051	2.03E-46	2.32E-44
ENSG00000096060.15	398.629	-1.362966442	0.146671	-9.2927	1.5E-20	5.02E-19
ENSG00000269416.7	215.7707	-1.360049999	0.163194	-8.33395	7.82E-17	2.02E-15
ENSG00000091651.9	476.8585	-1.358922139	0.115801	-11.735	8.43E-32	5.05E-30
ENSG00000215784.6	54.17974	-1.358162309	0.329862	-4.11737	3.83E-05	0.000281
ENSG00000080986.13	540.4258	-1.356927778	0.116131	-11.6844	1.53E-31	9.07E-30
ENSG00000186185.14	560.1333	-1.356362297	0.128526	-10.5532	4.91E-26	2.22E-24
ENSG00000131979.20	1199.36	-1.355831985	0.091725	-14.7816	1.93E-49	2.46E-47
ENSG00000155093.20	374.1449	-1.355395002	0.131451	-10.311	6.28E-25	2.64E-23

ENSG00000148773.14	3843.429	-1.353732178	0.108375	-12.4911	8.35E-36	5.8E-34
ENSG00000163918.12	446.6413	-1.352692214	0.123076	-10.9907	4.24E-28	2.12E-26
ENSG00000285095.2	19.40195	-1.350961541	0.492541	-2.74284	0.006091	0.02473
ENSG00000138658.16	325.1565	-1.34992323	0.137979	-9.78353	1.33E-22	4.91E-21
ENSG00000137563.13	1095.826	-1.349503631	0.129978	-10.3826	2.98E-25	1.3E-23
ENSG00000137558.9	1963.406	-1.348921521	0.100044	-13.4833	1.96E-41	1.74E-39
ENSG00000228203.8	148.3952	-1.348222083	0.196406	-6.86445	6.67E-12	1.19E-10
ENSG00000183856.11	1495.56	-1.345389065	0.101078	-13.3104	2.01E-40	1.72E-38
ENSG00000153044.10	868.4683	-1.342665767	0.102116	-13.1485	1.74E-39	1.39E-37
ENSG00000100162.15	81.38863	-1.340633768	0.257538	-5.20557	1.93E-07	2.11E-06
ENSG00000071539.14	688.6187	-1.34045562	0.101546	-13.2005	8.71E-40	7.15E-38
ENSG00000120875.9	382.0497	-1.338793226	0.144327	-9.27611	1.76E-20	5.82E-19
ENSG00000151490.15	141.4311	-1.338421918	0.205103	-6.52561	6.77E-11	1.11E-09
ENSG00000165895.19	1881.399	-1.337970885	0.078629	-17.0163	6.22E-65	1.34E-62
ENSG00000153823.19	32.88617	-1.337503337	0.38932	-3.43549	0.000591	0.003253
ENSG00000139438.6	1681.218	-1.333938322	0.095397	-13.9831	1.98E-44	2E-42
ENSG00000177181.15	378.7559	-1.333810945	0.126054	-10.5813	3.64E-26	1.66E-24
ENSG00000184661.14	366.8873	-1.333011137	0.134472	-9.9129	3.66E-23	1.39E-21
ENSG00000071575.12	432.6501	-1.332928276	0.123972	-10.7518	5.81E-27	2.79E-25
ENSG00000139318.8	1335.093	-1.330134764	0.098786	-13.4649	2.52E-41	2.22E-39
ENSG00000008283.16	5684.87	-1.329769496	0.064384	-20.6537	9.05E-95	4.11E-92
ENSG00000070182.21	37.23798	-1.32825817	0.36503	-3.63877	0.000274	0.001647
ENSG00000106399.11	879.2068	-1.326658761	0.282179	-4.70148	2.58E-06	2.36E-05
ENSG00000116990.12	46.25266	-1.325453814	0.351143	-3.77468	0.00016	0.001017
ENSG00000160716.6	1098.979	-1.323718528	0.102589	-12.9031	4.32E-38	3.31E-36
ENSG00000140534.14	344.3646	-1.323446991	0.151559	-8.73223	2.5E-18	7.17E-17
ENSG00000197744.5	53.77545	-1.32102647	0.309754	-4.26475	2E-05	0.000155
ENSG00000171604.12	2193.127	-1.320877004	0.082052	-16.098	2.64E-58	4.45E-56
ENSG00000204956.6	59.03723	-1.315289237	0.320349	-4.1058	4.03E-05	0.000294
ENSG00000137812.21	815.3486	-1.314751547	0.121935	-10.7824	4.17E-27	2.01E-25
ENSG00000126583.12	23.41475	-1.3122534	0.479411	-2.73722	0.006196	0.025081
ENSG00000175643.10	441.17	-1.310645437	0.121775	-10.7628	5.16E-27	2.48E-25
ENSG00000166073.12	1072.793	-1.309708511	0.101346	-12.9231	3.33E-38	2.57E-36
ENSG00000139514.13	1428.018	-1.30959415	0.107256	-12.21	2.75E-34	1.79E-32
ENSG00000139354.11	318.2897	-1.308962701	0.142342	-9.19588	3.72E-20	1.21E-18
ENSG00000225206.12	23.93865	-1.308277343	0.45504	-2.87508	0.004039	0.01742
ENSG00000076864.21	108.9675	-1.308193688	0.213777	-6.11942	9.39E-10	1.34E-08
ENSG00000131747.15	6734.368	-1.307081947	0.067796	-19.2795	7.98E-83	2.63E-80
ENSG00000139618.18	325.1377	-1.306620729	0.14164	-9.22495	2.84E-20	9.3E-19
ENSG00000158710.15	2849.947	-1.306413305	0.081884	-15.9544	2.65E-57	4.31E-55
ENSG00000117394.24	1336.744	-1.305745507	0.088883	-14.6907	7.4E-49	9.16E-47
ENSG00000258748.1	54.20993	-1.304685528	0.319259	-4.08661	4.38E-05	0.000317
ENSG00000133119.13	868.4855	-1.303750809	0.107646	-12.1114	9.19E-34	5.9E-32
ENSG00000112394.18	43.2385	-1.301490445	0.361972	-3.59556	0.000324	0.001905
ENSG00000129173.13	168.4344	-1.297852178	0.180531	-7.18907	6.52E-13	1.26E-11
ENSG00000161800.13	1765.164	-1.297784862	0.079833	-16.2562	2.02E-59	3.62E-57

ENSG00000135045.7	168.535	-1.296283742	0.210489	-6.15844	7.35E-10	1.07E-08
ENSG00000246465.1	54.61711	-1.295830006	0.30764	-4.21216	2.53E-05	0.000192
ENSG00000223572.10	23.63374	-1.293532308	0.446795	-2.89513	0.00379	0.016481
ENSG00000076356.7	1184.624	-1.293064728	0.114003	-11.3424	8.09E-30	4.36E-28
ENSG00000012048.25	691.35	-1.29018213	0.105537	-12.2249	2.29E-34	1.49E-32
ENSG00000188573.8	412.2893	-1.288827469	0.143194	-9.00055	2.25E-19	6.89E-18
ENSG00000109205.17	23.16481	-1.287087337	0.468169	-2.7492	0.005974	0.024323
ENSG00000233270.1	23.09791	-1.286534545	0.45837	-2.80676	0.005004	0.020916
ENSG00000248538.10	127.4579	-1.286450959	0.207222	-6.20809	5.36E-10	7.88E-09
ENSG00000118276.12	1109.759	-1.286106006	0.092152	-13.9564	2.88E-44	2.88E-42
ENSG00000163376.11	77.30138	-1.285535342	0.276516	-4.64905	3.33E-06	2.99E-05
ENSG00000233871.2	24.54713	-1.283796036	0.452095	-2.83966	0.004516	0.019166
ENSG00000167664.9	24.38616	-1.281789826	0.455998	-2.81095	0.004939	0.020687
ENSG00000135451.13	362.4079	-1.279517316	0.152395	-8.39605	4.62E-17	1.21E-15
ENSG00000111674.9	1922.607	-1.275396205	0.084183	-15.1503	7.55E-52	1.07E-49
ENSG00000121621.7	507.9941	-1.272524435	0.140425	-9.06196	1.28E-19	4.01E-18
ENSG00000134057.15	1485.173	-1.271388603	0.085987	-14.7858	1.81E-49	2.32E-47
ENSG00000151067.23	88.9963	-1.270635054	0.244116	-5.20504	1.94E-07	2.11E-06
ENSG00000181544.16	79.5587	-1.26864187	0.278112	-4.56162	5.08E-06	4.4E-05
ENSG00000118193.12	859.2731	-1.267419313	0.098628	-12.8505	8.55E-38	6.42E-36
ENSG00000111206.13	1561.749	-1.26736172	0.091954	-13.7826	3.24E-43	3.13E-41
ENSG00000134594.5	106.9479	-1.264684057	0.240524	-5.25804	1.46E-07	1.61E-06
ENSG00000092470.12	483.7608	-1.263318435	0.119219	-10.5966	3.09E-26	1.42E-24
ENSG00000204301.6	335.8631	-1.263005467	0.159121	-7.93738	2.06E-15	4.77E-14
ENSG00000144485.11	1646.818	-1.262287622	0.122277	-10.3232	5.54E-25	2.35E-23
ENSG00000183671.13	109.5256	-1.262120225	0.217128	-5.81278	6.14E-09	8.05E-08
ENSG00000136842.14	2713.064	-1.259977446	0.075777	-16.6274	4.41E-62	8.84E-60
ENSG00000204856.12	108.9071	-1.259878532	0.230064	-5.47622	4.35E-08	5.12E-07
ENSG00000055163.21	2529.41	-1.259230512	0.088959	-14.1552	1.73E-45	1.88E-43
ENSG00000135476.12	912.2644	-1.258842127	0.110944	-11.3466	7.71E-30	4.18E-28
ENSG00000168490.14	37.25044	-1.258420006	0.394143	-3.1928	0.001409	0.006986
ENSG00000164070.12	213.1451	-1.257250775	0.157133	-8.00119	1.23E-15	2.89E-14
ENSG00000166851.15	985.9911	-1.253284021	0.100808	-12.4324	1.74E-35	1.2E-33
ENSG00000198826.11	602.4175	-1.25308168	0.113214	-11.0683	1.79E-28	9.14E-27
ENSG00000101003.12	670.4729	-1.252097316	0.101679	-12.3142	7.59E-35	5.1E-33
ENSG00000224738.2	43.04785	-1.251722499	0.334727	-3.73953	0.000184	0.001155
ENSG00000159399.10	1785.786	-1.251686414	0.08647	-14.4754	1.73E-47	2.08E-45
ENSG00000104689.10	19.87721	-1.25152433	0.493198	-2.53757	0.011163	0.041046
ENSG00000171451.15	148.2954	-1.250363298	0.185796	-6.72976	1.7E-11	2.92E-10
ENSG00000169184.7	63.20438	-1.249254041	0.317991	-3.92858	8.54E-05	0.000578
ENSG00000146410.12	92.28672	-1.246576803	0.257673	-4.83782	1.31E-06	1.26E-05
ENSG00000120539.15	663.093	-1.246201807	0.108839	-11.45	2.35E-30	1.3E-28
ENSG00000132026.15	38.31821	-1.245902021	0.389786	-3.19637	0.001392	0.006906
ENSG00000197961.12	313.3988	-1.245885235	0.144586	-8.61693	6.88E-18	1.91E-16
ENSG00000089199.10	2697.954	-1.245382614	0.093807	-13.2761	3.19E-40	2.68E-38
ENSG00000213967.12	82.02895	-1.244875144	0.257167	-4.84073	1.29E-06	1.25E-05

ENSG00000158050.5	73.87	-1.243472772	0.268569	-4.62999	3.66E-06	3.26E-05
ENSG00000068078.20	51.95071	-1.243299749	0.31037	-4.00587	6.18E-05	0.000432
ENSG00000289194.1	106.4404	-1.236206981	0.219209	-5.6394	1.71E-08	2.1E-07
ENSG00000108370.17	102.7744	-1.235949396	0.234629	-5.26767	1.38E-07	1.53E-06
ENSG00000138092.11	491.2267	-1.235559251	0.116435	-10.6116	2.63E-26	1.21E-24
ENSG00000225783.9	59.61397	-1.233718782	0.301515	-4.09173	4.28E-05	0.000311
ENSG00000049130.16	260.8101	-1.233457165	0.176275	-6.99736	2.61E-12	4.82E-11
ENSG00000136492.10	1031.478	-1.233438467	0.093021	-13.2597	3.96E-40	3.32E-38
ENSG00000048540.16	60.30994	-1.232428545	0.282276	-4.36604	1.27E-05	0.000102
ENSG00000185614.7	73.71718	-1.232145837	0.292494	-4.21255	2.53E-05	0.000191
ENSG00000069482.7	806.8759	-1.227523874	0.114285	-10.7409	6.54E-27	3.13E-25
ENSG00000185950.9	3329.777	-1.227142015	0.092855	-13.2157	7.12E-40	5.89E-38
ENSG00000075218.19	1293.785	-1.222348952	0.089496	-13.6581	1.81E-42	1.67E-40
ENSG00000053747.17	45.68176	-1.220743843	0.361622	-3.37575	0.000736	0.003936
ENSG00000090889.12	1346.614	-1.219882784	0.087831	-13.8889	7.39E-44	7.3E-42
ENSG00000100167.21	311.1096	-1.219378356	0.13467	-9.05456	1.37E-19	4.28E-18
ENSG00000114405.11	387.4205	-1.219209407	0.145756	-8.36471	6.03E-17	1.57E-15
ENSG00000267886.1	67.66864	-1.21888169	0.297253	-4.10048	4.12E-05	0.0003
ENSG00000158402.21	228.6185	-1.218029672	0.185801	-6.55555	5.54E-11	9.13E-10
ENSG00000074800.16	14550.86	-1.217158768	0.074558	-16.325	6.55E-60	1.21E-57
ENSG00000006625.18	539.3105	-1.215360571	0.117339	-10.3577	3.86E-25	1.68E-23
ENSG00000176208.9	234.9224	-1.214098723	0.157324	-7.71717	1.19E-14	2.62E-13
ENSG00000168389.18	128.3135	-1.213456832	0.206626	-5.87273	4.29E-09	5.72E-08
ENSG00000164687.11	297.5575	-1.212186683	0.152801	-7.93312	2.14E-15	4.93E-14
ENSG00000111602.12	2244.932	-1.210521102	0.083886	-14.4305	3.33E-47	3.91E-45
ENSG00000171241.9	262.5457	-1.209719989	0.163326	-7.4068	1.29E-13	2.63E-12
ENSG00000250979.1	142.1848	-1.209630287	0.201203	-6.01199	1.83E-09	2.54E-08
ENSG00000138316.11	40.40146	-1.208921734	0.369979	-3.26755	0.001085	0.00555
ENSG00000237989.2	86.04623	-1.207811932	0.254262	-4.75027	2.03E-06	1.89E-05
ENSG00000059915.17	187.3281	-1.206846338	0.173441	-6.95825	3.45E-12	6.3E-11
ENSG00000157456.8	828.9589	-1.206800965	0.109195	-11.0518	2.15E-28	1.09E-26
ENSG00000019505.8	20.49825	-1.204674886	0.475211	-2.53503	0.011244	0.041293
ENSG00000164087.8	288.3138	-1.202613931	0.139805	-8.6021	7.83E-18	2.16E-16
ENSG00000119333.12	1918.341	-1.201998672	0.104072	-11.5497	7.41E-31	4.22E-29
ENSG00000156802.13	1410.293	-1.201677142	0.087754	-13.6937	1.11E-42	1.04E-40
ENSG00000124207.17	3250.463	-1.198727457	0.120833	-9.92054	3.39E-23	1.29E-21
ENSG00000135472.9	178.7523	-1.198622778	0.171674	-6.98197	2.91E-12	5.35E-11
ENSG00000111445.14	414.5095	-1.198275165	0.120529	-9.94177	2.74E-23	1.05E-21
ENSG00000166451.14	612.2272	-1.197943098	0.126923	-9.43835	3.79E-21	1.3E-19
ENSG00000144834.14	388.1552	-1.196044181	0.13795	-8.6701	4.32E-18	1.22E-16
ENSG00000173848.19	459.5047	-1.195336018	0.121255	-9.85802	6.33E-23	2.38E-21
ENSG00000137310.13	533.9976	-1.193605376	0.110445	-10.8072	3.18E-27	1.54E-25
ENSG00000284946.1	61.69134	-1.193180828	0.287015	-4.15721	3.22E-05	0.000239
ENSG00000125885.13	314.6082	-1.192616535	0.150126	-7.94411	1.96E-15	4.54E-14
ENSG00000133302.13	914.9958	-1.19250352	0.12416	-9.60453	7.65E-22	2.71E-20
ENSG00000011422.12	278.4916	-1.19038859	0.150611	-7.90374	2.71E-15	6.2E-14

ENSG00000112530.13	83.7702	-1.18935019	0.267036	-4.4539	8.43E-06	7.01E-05
ENSG00000185480.12	327.3565	-1.188558939	0.155524	-7.6423	2.13E-14	4.6E-13
ENSG00000076003.5	1605.213	-1.186773518	0.078021	-15.211	2.99E-52	4.27E-50
ENSG00000137267.7	582.0381	-1.186644915	0.107211	-11.0684	1.79E-28	9.14E-27
ENSG00000140525.20	1670.792	-1.186510533	0.086	-13.7967	2.67E-43	2.6E-41
ENSG00000095370.20	155.7427	-1.183898514	0.184539	-6.41544	1.4E-10	2.21E-09
ENSG00000149328.16	79.34323	-1.182572032	0.256586	-4.60887	4.05E-06	3.58E-05
ENSG00000139174.12	317.5618	-1.180874506	0.138543	-8.52354	1.55E-17	4.17E-16
ENSG00000113368.12	2692.416	-1.180075907	0.078738	-14.9874	8.88E-51	1.22E-48
ENSG00000008196.13	4227.672	-1.179585202	0.075725	-15.5772	1.04E-54	1.59E-52
ENSG00000185347.19	296.7024	-1.17320106	0.142433	-8.23686	1.77E-16	4.42E-15
ENSG00000184635.17	228.2665	-1.172361164	0.17389	-6.74196	1.56E-11	2.69E-10
ENSG00000082684.16	253.4127	-1.172200251	0.17454	-6.71594	1.87E-11	3.19E-10
ENSG00000288840.1	43.86658	-1.171414719	0.355578	-3.29439	0.000986	0.005102
ENSG00000143933.20	22000.79	-1.170803191	0.235673	-4.96792	6.77E-07	6.86E-06
ENSG00000188321.14	378.6428	-1.17035642	0.139037	-8.41758	3.84E-17	1.01E-15
ENSG00000167136.7	581.4007	-1.169539355	0.128129	-9.12779	6.99E-20	2.23E-18
ENSG00000166165.14	3357.601	-1.167744087	0.078335	-14.907	2.97E-50	3.96E-48
ENSG00000242419.6	287.212	-1.16636094	0.156813	-7.43792	1.02E-13	2.09E-12
ENSG00000244242.2	23.16282	-1.165689392	0.448416	-2.59957	0.009334	0.035406
ENSG00000286125.3	24.38756	-1.164026411	0.44856	-2.59503	0.009458	0.035791
ENSG00000087495.17	56.41758	-1.16202694	0.29289	-3.96745	7.26E-05	0.000501
ENSG00000181026.15	597.4842	-1.16194367	0.120644	-9.63117	5.91E-22	2.12E-20
ENSG00000265763.4	30.77483	-1.160912057	0.400223	-2.90066	0.003724	0.01621
ENSG00000174939.11	128.7441	-1.160760192	0.213068	-5.44783	5.1E-08	5.96E-07
ENSG00000070950.10	405.7401	-1.156891231	0.12297	-9.40795	5.06E-21	1.73E-19
ENSG00000169679.15	845.4547	-1.156212044	0.102075	-11.3271	9.63E-30	5.15E-28
ENSG00000164104.12	6982.532	-1.155417617	0.108327	-10.666	1.47E-26	6.87E-25
ENSG00000231185.9	64.0523	-1.152664179	0.28497	-4.04486	5.24E-05	0.000372
ENSG00000263513.6	46.45289	-1.152496028	0.321726	-3.58223	0.000341	0.00199
ENSG00000185697.17	124.6671	-1.152443682	0.210198	-5.48266	4.19E-08	4.95E-07
ENSG00000105750.15	189.6706	-1.151914215	0.176501	-6.52638	6.74E-11	1.1E-09
ENSG00000228709.3	283.5885	-1.151719049	0.159411	-7.22485	5.02E-13	9.74E-12
ENSG00000184445.12	957.7487	-1.151151711	0.095376	-12.0697	1.53E-33	9.66E-32
ENSG00000068489.13	2072.603	-1.147466675	0.07694	-14.9137	2.68E-50	3.61E-48
ENSG00000183092.18	995.0465	-1.146158615	0.113699	-10.0807	6.73E-24	2.69E-22
ENSG00000267383.9	49.51212	-1.146147135	0.364071	-3.14815	0.001643	0.007973
ENSG00000170345.10	28.45937	-1.143723938	0.402393	-2.8423	0.004479	0.019025
ENSG00000159259.8	476.5012	-1.142917426	0.118306	-9.66065	4.43E-22	1.6E-20
ENSG00000116329.12	229.4112	-1.142727419	0.192927	-5.92312	3.16E-09	4.27E-08
ENSG00000241472.8	85.41491	-1.142497872	0.236259	-4.83578	1.33E-06	1.27E-05
ENSG00000130475.16	32.67717	-1.142430362	0.390578	-2.92498	0.003445	0.01518
ENSG00000162374.18	3569.477	-1.141359798	0.071375	-15.991	1.48E-57	2.43E-55
ENSG00000164418.22	126.5358	-1.139922637	0.2013	-5.66281	1.49E-08	1.85E-07
ENSG00000204899.6	1185.106	-1.139423752	0.113596	-10.0305	1.12E-23	4.41E-22
ENSG00000229676.3	83.00637	-1.138381107	0.284805	-3.99705	6.41E-05	0.000447

ENSG00000100077.16	2176.099	-1.138027512	0.089709	-12.6858	7.09E-37	5.19E-35
ENSG00000134508.13	95.70429	-1.137746279	0.237169	-4.79719	1.61E-06	1.53E-05
ENSG00000130600.22	53.39003	-1.1366579	0.366999	-3.09717	0.001954	0.009282
ENSG00000165704.15	482.2836	-1.135982887	0.129209	-8.79179	1.47E-18	4.28E-17
ENSG00000088325.16	1850.593	-1.134327571	0.088058	-12.8816	5.71E-38	4.34E-36
ENSG00000013619.15	29.92192	-1.133866268	0.44031	-2.57515	0.01002	0.037522
ENSG00000171649.12	399.2186	-1.133479959	0.133456	-8.4933	2.01E-17	5.38E-16
ENSG00000158859.10	550.7148	-1.133291143	0.118428	-9.56949	1.07E-21	3.78E-20
ENSG00000180537.13	363.4032	-1.133270482	0.128862	-8.79442	1.44E-18	4.19E-17
ENSG00000131242.18	551.75	-1.132733406	0.121877	-9.29405	1.49E-20	4.97E-19
ENSG00000110400.11	1578.498	-1.132713442	0.096297	-11.7627	6.08E-32	3.66E-30
ENSG00000170779.11	801.402	-1.130457081	0.121294	-9.31996	1.16E-20	3.91E-19
ENSG00000176749.9	371.7475	-1.129507221	0.128834	-8.76715	1.83E-18	5.3E-17
ENSG00000226415.1	77.52621	-1.128923232	0.263711	-4.28091	1.86E-05	0.000145
ENSG00000177606.9	381.588	-1.128766219	0.12791	-8.82471	1.1E-18	3.22E-17
ENSG00000167325.15	1757.047	-1.128615637	0.082699	-13.6472	2.1E-42	1.92E-40
ENSG00000111665.12	250.8887	-1.124509661	0.167645	-6.70767	1.98E-11	3.38E-10
ENSG00000146918.20	1816.452	-1.124063213	0.09061	-12.4055	2.44E-35	1.67E-33
ENSG00000264350.1	51.53849	-1.12394723	0.332383	-3.38148	0.000721	0.00387
ENSG00000106948.17	76.76168	-1.122704936	0.2607	-4.30651	1.66E-05	0.00013
ENSG00000166840.14	27.55588	-1.121886551	0.417375	-2.68796	0.007189	0.028485
ENSG00000119547.6	86.11619	-1.121286464	0.25838	-4.33969	1.43E-05	0.000114
ENSG00000086730.17	42.42568	-1.120096263	0.337143	-3.32232	0.000893	0.004676
ENSG00000120802.14	3491.297	-1.118233595	0.076048	-14.7042	6.06E-49	7.59E-47
ENSG00000268089.3	92.14427	-1.116818168	0.234145	-4.76977	1.84E-06	1.73E-05
ENSG00000015133.20	568.6383	-1.115591053	0.129712	-8.60052	7.94E-18	2.19E-16
ENSG00000068028.18	357.1872	-1.115247851	0.138811	-8.03432	9.41E-16	2.23E-14
ENSG00000188312.15	96.14966	-1.114086463	0.225536	-4.93972	7.82E-07	7.8E-06
ENSG00000143882.12	25.14092	-1.113125402	0.439493	-2.53275	0.011317	0.041533
ENSG00000085831.16	55.49855	-1.108937302	0.301222	-3.68146	0.000232	0.001417
ENSG00000172244.9	142.5907	-1.108150599	0.193745	-5.71962	1.07E-08	1.36E-07
ENSG00000106025.9	347.6864	-1.107252013	0.140241	-7.89536	2.89E-15	6.62E-14
ENSG00000127920.6	1567.713	-1.104504719	0.135572	-8.14698	3.73E-16	9.07E-15
ENSG00000162755.14	26.84123	-1.104322467	0.427936	-2.58058	0.009864	0.03707
ENSG00000006634.8	710.0226	-1.104299848	0.10064	-10.9728	5.17E-28	2.57E-26
ENSG00000151617.17	651.2644	-1.103135038	0.110321	-9.99932	1.53E-23	5.97E-22
ENSG00000293339.1	27.66161	-1.102635421	0.429066	-2.56985	0.010174	0.037989
ENSG00000164061.5	188.2245	-1.099531069	0.189785	-5.79358	6.89E-09	8.97E-08
ENSG00000108179.14	733.6654	-1.098871832	0.103215	-10.6464	1.81E-26	8.44E-25
ENSG00000163535.18	936.4941	-1.096965065	0.119573	-9.17401	4.56E-20	1.48E-18
ENSG00000174442.12	1307.674	-1.09517562	0.104952	-10.435	1.72E-25	7.52E-24
ENSG00000171223.6	413.9769	-1.094820635	0.140529	-7.79068	6.66E-15	1.49E-13
ENSG00000122966.18	523.8439	-1.09127129	0.137462	-7.9387	2.04E-15	4.73E-14
ENSG00000187741.16	403.9877	-1.090201609	0.136393	-7.99307	1.32E-15	3.08E-14
ENSG00000010278.15	110.9395	-1.089642734	0.224604	-4.85139	1.23E-06	1.18E-05
ENSG00000215218.4	260.6612	-1.089387078	0.146703	-7.42582	1.12E-13	2.29E-12

ENSG00000169871.13	848.3434	-1.088794335	0.108077	-10.0742	7.18E-24	2.85E-22
ENSG00000128944.14	795.6144	-1.088150844	0.107302	-10.141	3.63E-24	1.48E-22
ENSG00000175183.10	1460.894	-1.087727793	0.123831	-8.78399	1.58E-18	4.58E-17
ENSG00000197472.15	44.44268	-1.087606888	0.331011	-3.28571	0.001017	0.00524
ENSG00000138182.15	881.068	-1.086103167	0.115852	-9.37493	6.92E-21	2.36E-19
ENSG00000163254.5	27.17579	-1.085984777	0.435061	-2.49616	0.012554	0.045413
ENSG00000164649.20	810.7547	-1.084014399	0.104994	-10.3245	5.46E-25	2.32E-23
ENSG00000146263.12	422.0365	-1.080557663	0.1259	-8.58264	9.27E-18	2.54E-16
ENSG00000153721.19	285.1052	-1.08044468	0.157487	-6.86055	6.86E-12	1.22E-10
ENSG00000181938.14	244.1876	-1.079257321	0.147246	-7.3296	2.31E-13	4.58E-12
ENSG00000175344.19	129.4558	-1.077744032	0.198301	-5.43488	5.48E-08	6.39E-07
ENSG00000140479.18	111.9098	-1.07654611	0.217936	-4.93973	7.82E-07	7.8E-06
ENSG00000116991.12	4663.285	-1.076226855	0.076123	-14.1381	2.21E-45	2.37E-43
ENSG00000128708.13	1398.096	-1.075145134	0.240293	-4.4743	7.67E-06	6.42E-05
ENSG00000130021.15	224.0815	-1.070936774	0.169761	-6.30848	2.82E-10	4.3E-09
ENSG00000081181.8	75.31009	-1.070080578	0.258337	-4.14219	3.44E-05	0.000254
ENSG00000168917.9	170.7092	-1.07004479	0.186418	-5.74003	9.47E-09	1.22E-07
ENSG00000101945.17	169.2442	-1.069848959	0.174998	-6.11351	9.75E-10	1.39E-08
ENSG00000126803.10	257.0973	-1.06825077	0.148335	-7.20161	5.95E-13	1.15E-11
ENSG00000105290.13	2480.237	-1.068156993	0.076227	-14.0129	1.3E-44	1.33E-42
ENSG00000289438.1	49.11874	-1.068122439	0.328833	-3.24822	0.001161	0.005894
ENSG00000224940.10	2631.595	-1.068068438	0.074701	-14.2979	2.26E-46	2.55E-44
ENSG00000183763.9	200.6274	-1.06764592	0.169782	-6.28835	3.21E-10	4.82E-09
ENSG00000188229.6	4615.247	-1.066106865	0.074573	-14.2961	2.31E-46	2.6E-44
ENSG00000163808.17	831.001	-1.06505656	0.097461	-10.928	8.47E-28	4.17E-26
ENSG00000049541.11	1063.075	-1.063166613	0.101656	-10.4585	1.34E-25	5.93E-24
ENSG00000128050.9	4524.57	-1.062401337	0.085631	-12.4068	2.4E-35	1.65E-33
ENSG00000108106.14	3927.591	-1.060987391	0.087734	-12.0933	1.15E-33	7.3E-32
ENSG00000109084.14	683.3553	-1.058937801	0.102301	-10.3512	4.13E-25	1.78E-23
ENSG00000113196.3	818.1433	-1.057940203	0.101036	-10.4709	1.17E-25	5.21E-24
ENSG00000232892.1	33.6661	-1.057290812	0.40708	-2.59726	0.009397	0.035612
ENSG00000007402.12	11503.78	-1.056841348	0.080148	-13.1862	1.05E-39	8.58E-38
ENSG00000138587.6	109.2461	-1.056661027	0.258081	-4.09429	4.23E-05	0.000308
ENSG00000229891.4	41.40091	-1.055089133	0.380044	-2.77623	0.005499	0.022647
ENSG00000123136.15	1211.385	-1.052161881	0.091882	-11.4512	2.32E-30	1.28E-28
ENSG00000162407.9	95.84994	-1.048604458	0.239259	-4.38272	1.17E-05	9.49E-05
ENSG00000141682.12	759.9305	-1.04831948	0.135912	-7.71321	1.23E-14	2.7E-13
ENSG00000196230.14	25589.42	-1.04582325	0.06355	-16.4568	7.5E-61	1.43E-58
ENSG00000149548.15	117.9136	-1.04401723	0.216065	-4.83195	1.35E-06	1.3E-05
ENSG00000233966.1	53.66813	-1.043392258	0.304247	-3.42942	0.000605	0.003317
ENSG00000157680.17	183.8042	-1.043113012	0.173253	-6.02075	1.74E-09	2.41E-08
ENSG00000226976.3	51.72722	-1.042811877	0.301503	-3.45871	0.000543	0.003008
ENSG00000188042.8	209.252	-1.042790276	0.179976	-5.79406	6.87E-09	8.95E-08
ENSG00000137727.13	163.9641	-1.042702201	0.205051	-5.08508	3.67E-07	3.86E-06
ENSG00000112984.12	1157.66	-1.040972347	0.091773	-11.3429	8.05E-30	4.35E-28
ENSG00000274276.6	96.76504	-1.040759996	0.248611	-4.18629	2.84E-05	0.000213

ENSG00000184613.11	118.7828	-1.040726225	0.208688	-4.98699	6.13E-07	6.24E-06
ENSG00000040275.17	626.1895	-1.038784971	0.111478	-9.31828	1.18E-20	3.97E-19
ENSG00000116183.11	50.97041	-1.036150438	0.329931	-3.14051	0.001687	0.008145
ENSG00000109107.14	632.4963	-1.035701888	0.106709	-9.70587	2.85E-22	1.04E-20
ENSG00000132872.12	3948.577	-1.035397576	0.094214	-10.9899	4.28E-28	2.14E-26
ENSG00000166483.12	870.1607	-1.035001787	0.092517	-11.1871	4.71E-29	2.46E-27
ENSG00000154553.17	526.2653	-1.034102364	0.127811	-8.0909	5.92E-16	1.42E-14
ENSG00000114346.14	1693.52	-1.033088952	0.084506	-12.2251	2.28E-34	1.49E-32
ENSG00000198959.12	5058.561	-1.032499951	0.074176	-13.9196	4.81E-44	4.8E-42
ENSG00000125354.24	6420.505	-1.03187037	0.07503	-13.7527	4.91E-43	4.68E-41
ENSG00000166016.6	680.5092	-1.03170902	0.100911	-10.2239	1.55E-24	6.42E-23
ENSG00000005108.17	502.3944	-1.031061051	0.131698	-7.82897	4.92E-15	1.11E-13
ENSG00000245694.12	501.3879	-1.029887287	0.130216	-7.90905	2.59E-15	5.95E-14
ENSG00000112118.20	3150.719	-1.029462994	0.071874	-14.3232	1.57E-46	1.81E-44
ENSG00000166562.9	443.6738	-1.028739253	0.147833	-6.95877	3.43E-12	6.29E-11
ENSG00000082929.8	268.6366	-1.027658021	0.179694	-5.71895	1.07E-08	1.36E-07
ENSG00000175175.6	480.202	-1.0250688	0.118588	-8.64395	5.43E-18	1.52E-16
ENSG00000188486.4	3595.344	-1.023601954	0.080092	-12.7803	2.11E-37	1.57E-35
ENSG00000011332.22	212.9382	-1.023405347	0.161012	-6.35608	2.07E-10	3.2E-09
ENSG00000075340.23	474.1584	-1.023342847	0.147968	-6.91599	4.65E-12	8.4E-11
ENSG00000214826.5	77.88004	-1.020935784	0.289101	-3.53141	0.000413	0.002359
ENSG00000170017.12	5244.358	-1.019460587	0.076629	-13.3039	2.2E-40	1.87E-38
ENSG00000198088.10	32.38738	-1.018911269	0.38983	-2.61373	0.008956	0.034226
ENSG00000162607.13	1800.456	-1.01880023	0.09075	-11.2264	3.02E-29	1.59E-27
ENSG00000135905.21	224.4218	-1.018682912	0.160357	-6.3526	2.12E-10	3.27E-09
ENSG00000067840.13	1282.811	-1.018440561	0.083007	-12.2694	1.32E-34	8.75E-33
ENSG00000166582.11	1013.95	-1.0148881	0.104116	-9.74768	1.89E-22	6.92E-21
ENSG00000163923.10	308.1227	-1.014617809	0.184899	-5.48741	4.08E-08	4.83E-07
ENSG00000188610.12	101.1747	-1.014582131	0.231168	-4.38894	1.14E-05	9.25E-05
ENSG00000160307.10	31.73106	-1.014479475	0.38772	-2.61652	0.008883	0.03399
ENSG00000213347.11	345.3425	-1.014094662	0.137631	-7.36823	1.73E-13	3.47E-12
ENSG00000138160.7	1583.311	-1.01362052	0.08572	-11.8248	2.91E-32	1.77E-30
ENSG00000138495.7	322.4664	-1.00891504	0.186105	-5.42121	5.92E-08	6.87E-07
ENSG00000137193.14	257.5909	-1.008614894	0.146796	-6.87084	6.38E-12	1.14E-10
ENSG00000117724.15	3943.193	-1.007551116	0.084379	-11.9408	7.26E-33	4.51E-31
ENSG00000170860.4	937.1549	-1.006994301	0.136042	-7.4021	1.34E-13	2.72E-12
ENSG00000182481.10	4093.201	-1.006606249	0.078254	-12.8634	7.24E-38	5.46E-36
ENSG00000079616.14	1158.874	-1.00488135	0.087361	-11.5026	1.28E-30	7.19E-29
ENSG00000154640.15	599.9825	-1.004758661	0.13189	-7.61816	2.57E-14	5.49E-13
ENSG00000113356.13	62.87206	-1.004188499	0.282812	-3.55072	0.000384	0.002214
ENSG00000105486.15	1032.023	-1.001344032	0.090972	-11.0071	3.53E-28	1.78E-26

Significantly upregulated genes immature LIPT2-AS1$^{-/-}$ vs WT

gene_id	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj
ENSG00000164294.14	522.4739	8.753644	0.722635	12.11351	8.96E-34	8.78E-30
ENSG00000150457.9	65.07349	8.633233	1.243678	6.941692	3.87E-12	1.43E-09
ENSG00000251003.9	22.94859	8.09763	1.334344	6.068624	1.29E-09	3.2E-07
ENSG00000164023.15	15.33343	7.516423	1.44215	5.211957	1.87E-07	2.63E-05
ENSG00000156535.15	72.96543	7.341602	0.955732	7.681652	1.57E-14	8.55E-12
ENSG00000218233.1	23.73198	7.172168	1.296962	5.529974	3.2E-08	5.71E-06
ENSG00000258053.1	11.78337	7.131744	1.389809	5.131456	2.88E-07	3.78E-05
ENSG00000170160.19	22.22897	7.080798	1.335336	5.302632	1.14E-07	1.71E-05
ENSG00000261213.1	10.34213	6.949116	1.56392	4.443397	8.85E-06	0.000754
ENSG00000146233.8	8.349058	6.636252	1.433816	4.628385	3.69E-06	0.000354
ENSG00000290976.1	7.588823	6.500038	1.440861	4.511219	6.45E-06	0.000577
ENSG00000112280.18	27.77717	6.461962	1.179827	5.477041	4.32E-08	7.31E-06
ENSG00000196427.14	6.537768	6.285398	1.484319	4.234534	2.29E-05	0.001733
ENSG00000237283.2	12.81347	6.275947	1.409671	4.452065	8.5E-06	0.000731
ENSG00000135678.12	11.6618	6.142918	1.397525	4.395569	1.1E-05	0.00091
ENSG00000255794.11	5.577003	6.05089	1.747612	3.462377	0.000535	0.026033
ENSG00000186086.19	5.292931	5.976295	1.557001	3.838338	0.000124	0.007491
ENSG00000131620.18	4.839044	5.855311	1.675165	3.495363	0.000473	0.023785
ENSG00000168899.5	16.17636	5.730429	1.228058	4.666254	3.07E-06	0.000302
ENSG00000250305.9	4.425139	5.717565	1.639016	3.488412	0.000486	0.024103
ENSG00000176788.9	225.4491	5.109184	0.456092	11.2021	3.98E-29	1.11E-25
ENSG00000197935.7	15.48787	5.093681	1.081876	4.708193	2.5E-06	0.000252
ENSG00000162624.16	126.0099	4.886159	0.611753	7.987144	1.38E-15	9.33E-13
ENSG00000113532.13	29.69345	4.809388	0.83986	5.726417	1.03E-08	1.93E-06
ENSG00000116711.10	21.58793	4.779831	0.864969	5.526016	3.28E-08	5.78E-06
ENSG00000112837.17	305.5082	4.69564	0.409265	11.47336	1.8E-30	7.04E-27
ENSG00000188313.13	7.593298	4.571017	1.316323	3.472565	0.000516	0.025189
ENSG00000111728.11	48.70283	4.554945	0.590087	7.719114	1.17E-14	6.75E-12
ENSG00000122862.5	20.38318	4.414702	1.251284	3.528138	0.000418	0.02141
ENSG00000143341.12	9.316484	4.331766	1.110398	3.901094	9.58E-05	0.006132
ENSG00000186340.17	13.45254	4.309503	1.181491	3.647512	0.000265	0.014372
ENSG00000169064.13	8.107233	4.073213	1.100518	3.701179	0.000215	0.01198
ENSG00000164342.14	17.51918	3.743506	0.79421	4.713498	2.44E-06	0.000251
ENSG00000251493.5	65.55952	3.681965	1.087694	3.385112	0.000711	0.033593
ENSG00000175984.16	17.43021	3.420384	0.715129	4.782888	1.73E-06	0.000186
ENSG00000180530.11	19.96406	3.316718	0.713845	4.646273	3.38E-06	0.000326
ENSG00000205413.8	174.7671	3.116283	0.527688	5.90554	3.51E-09	7.65E-07
ENSG00000153291.16	25.46609	3.094689	0.660967	4.68206	2.84E-06	0.000284
ENSG00000204335.4	33.36855	3.083807	0.88587	3.481104	0.000499	0.024646
ENSG00000046604.14	14.1162	3.067169	0.810233	3.78554	0.000153	0.008953
ENSG00000139117.14	19.83545	3.035929	0.839701	3.615489	0.0003	0.016137
ENSG00000073756.13	41.19673	2.942332	0.61047	4.819778	1.44E-06	0.000158
ENSG00000101000.6	24.6913	2.856363	0.673212	4.24289	2.21E-05	0.001676

ENSG00000144227.5	26.4885	2.811126	0.632041	4.447697	8.68E-06	0.000743
ENSG00000279675.1	15.50582	2.795514	0.797079	3.507197	0.000453	0.022947
ENSG00000123496.8	146.0429	2.738451	0.485062	5.645572	1.65E-08	3.04E-06
ENSG00000021826.18	114.5771	2.725439	0.357488	7.623854	2.46E-14	1.27E-11
ENSG00000173698.18	103.8786	2.707788	0.417107	6.491827	8.48E-11	2.6E-08
ENSG00000270959.2	21.65339	2.688981	0.641552	4.191367	2.77E-05	0.002058
ENSG00000122824.11	47.74928	2.66887	0.464709	5.743094	9.3E-09	1.79E-06
ENSG00000078579.9	41.26684	2.654147	0.623575	4.256342	2.08E-05	0.001603
ENSG00000074410.14	21.44274	2.642386	0.607846	4.347128	1.38E-05	0.001085
ENSG00000184304.17	79.17808	2.62098	0.397759	6.589358	4.42E-11	1.42E-08
ENSG00000110002.16	22.31296	2.532288	0.564422	4.486514	7.24E-06	0.000636
ENSG00000292998.1	23.41343	2.468616	0.542548	4.550043	5.36E-06	0.000493
ENSG00000164442.11	13.53201	2.452058	0.712281	3.442546	0.000576	0.027607
ENSG00000189171.15	50.63656	2.449749	0.526339	4.654316	3.25E-06	0.000315
ENSG00000104760.17	149.3291	2.44503	0.404708	6.041472	1.53E-09	3.69E-07
ENSG00000182022.18	92.07811	2.44333	0.363753	6.717002	1.86E-11	6.38E-09
ENSG00000129422.15	800.4744	2.437745	0.366233	6.656261	2.81E-11	9.33E-09
ENSG00000128606.13	108.0325	2.424144	0.432716	5.602163	2.12E-08	3.84E-06
ENSG00000148948.8	159.0926	2.396572	0.333583	7.184341	6.75E-13	2.94E-10
ENSG00000170396.9	102.9095	2.386599	0.439845	5.425998	5.76E-08	9.25E-06
ENSG00000153993.14	48.51516	2.332902	0.566109	4.12094	3.77E-05	0.002748
ENSG00000073792.16	339.1084	2.292945	0.315593	7.265508	3.72E-13	1.78E-10
ENSG00000120594.17	42.10803	2.271282	0.592153	3.835632	0.000125	0.007551
ENSG00000187323.13	72.68916	2.264658	0.381595	5.934717	2.94E-09	6.48E-07
ENSG00000115361.8	24.0327	2.250203	0.560548	4.014289	5.96E-05	0.004114
ENSG00000119684.16	316.1532	2.164087	0.315506	6.859107	6.93E-12	2.42E-09
ENSG00000154645.14	15.57658	2.154416	0.645636	3.336888	0.000847	0.03906
ENSG00000072952.21	111.1524	2.125781	0.460268	4.618576	3.86E-06	0.000369
ENSG00000078098.15	117.599	2.101234	0.347081	6.054019	1.41E-09	3.46E-07
ENSG00000164100.9	35.35852	2.095351	0.636356	3.292734	0.000992	0.044089
ENSG00000276399.1	26.18994	2.046957	0.585944	3.493437	0.000477	0.02385
ENSG00000145012.15	509.9579	2.04367	0.476624	4.287802	1.8E-05	0.001409
ENSG00000132872.12	4129.161	1.995038	0.412669	4.83447	1.34E-06	0.000149
ENSG00000245067.8	46.11305	1.970767	0.482395	4.085384	4.4E-05	0.003124
ENSG00000162631.20	169.1436	1.938054	0.345882	5.603226	2.1E-08	3.84E-06
ENSG00000176697.20	273.8347	1.935011	0.527472	3.668465	0.000244	0.013393
ENSG00000156869.14	128.4201	1.922155	0.387888	4.955444	7.22E-07	8.67E-05
ENSG00000173930.9	64.95456	1.909952	0.389831	4.899438	9.61E-07	0.000113
ENSG00000251011.6	137.9785	1.902222	0.412137	4.615513	3.92E-06	0.000369
ENSG00000152104.12	473.9898	1.864705	0.344425	5.413973	6.16E-08	9.74E-06
ENSG00000104219.13	1940.762	1.861821	0.367048	5.072418	3.93E-07	5E-05
ENSG00000197747.9	223.9108	1.85572	0.266512	6.962999	3.33E-12	1.28E-09
ENSG00000226530.4	63.72842	1.851781	0.466819	3.96681	7.28E-05	0.004855
ENSG00000113805.9	177.0479	1.837521	0.417953	4.396478	1.1E-05	0.00091
ENSG00000169760.18	153.144	1.831006	0.531006	3.448182	0.000564	0.027237
ENSG00000140285.12	255.9337	1.814834	0.402865	4.504817	6.64E-06	0.00059

ENSG00000254535.4	157.7605	1.776145	0.506461	3.506975	0.000453	0.022947
ENSG00000169432.19	1144.686	1.770009	0.425501	4.159827	3.18E-05	0.002355
ENSG00000255248.11	47.23943	1.754571	0.478017	3.670523	0.000242	0.013323
ENSG00000075223.14	1627.05	1.747295	0.378572	4.615487	3.92E-06	0.000369
ENSG00000142700.12	231.9623	1.742866	0.370121	4.708904	2.49E-06	0.000252
ENSG00000120820.13	45.02289	1.732337	0.466635	3.712403	0.000205	0.011493
ENSG00000198795.11	108.452	1.712711	0.434957	3.937655	8.23E-05	0.005392
ENSG00000076716.9	123.1848	1.706843	0.3091	5.521969	3.35E-08	5.86E-06
ENSG00000174799.12	1247.383	1.704674	0.435008	3.91872	8.9E-05	0.005738
ENSG00000214106.9	75.15135	1.682384	0.345393	4.870922	1.11E-06	0.000127
ENSG00000138650.9	71.09274	1.643064	0.405133	4.055619	5E-05	0.003505
ENSG00000079931.15	364.0198	1.641981	0.336609	4.878001	1.07E-06	0.000124
ENSG00000231887.9	32.28986	1.638345	0.485314	3.375844	0.000736	0.034661
ENSG00000196549.13	134.1309	1.600099	0.450957	3.548231	0.000388	0.019934
ENSG00000007062.12	50.92603	1.582628	0.417955	3.7866	0.000153	0.008953
ENSG00000182667.15	76.33298	1.564814	0.403435	3.878723	0.000105	0.006564
ENSG00000197928.11	323.5016	1.555498	0.400332	3.885522	0.000102	0.006433
ENSG00000125398.8	229.0021	1.549415	0.301485	5.139274	2.76E-07	3.65E-05
ENSG00000290901.1	52.22793	1.540571	0.428932	3.591644	0.000329	0.017292
ENSG00000166147.15	5158.917	1.512452	0.274802	5.50378	3.72E-08	6.39E-06
ENSG00000119508.18	141.2759	1.503015	0.297571	5.050949	4.4E-07	5.52E-05
ENSG00000198121.15	520.0432	1.497743	0.254802	5.878077	4.15E-09	8.74E-07
ENSG00000112319.21	919.0715	1.497507	0.324648	4.612714	3.97E-06	0.000373
ENSG00000114251.15	83.96125	1.496718	0.375853	3.98219	6.83E-05	0.004582
ENSG00000169860.7	44.2521	1.47857	0.43507	3.398469	0.000678	0.032072
ENSG00000168938.6	434.8919	1.474446	0.362868	4.063317	4.84E-05	0.00341
ENSG00000139329.5	1118.899	1.45499	0.443795	3.278519	0.001044	0.045948
ENSG00000113594.10	573.2229	1.447431	0.430504	3.362175	0.000773	0.036249
ENSG00000132334.17	95.44466	1.440253	0.339025	4.248223	2.15E-05	0.001649
ENSG00000134602.16	1755.635	1.428724	0.370157	3.859783	0.000113	0.007015
ENSG00000165025.15	138.3302	1.399833	0.297446	4.706168	2.52E-06	0.000254
ENSG00000112902.12	183.3934	1.394822	0.309656	4.504422	6.66E-06	0.00059
ENSG00000187398.12	128.1777	1.388707	0.337306	4.117049	3.84E-05	0.002785
ENSG00000151491.14	495.2251	1.387251	0.348882	3.976271	7E-05	0.004681
ENSG00000154721.15	334.3393	1.379334	0.351967	3.918927	8.89E-05	0.005738
ENSG00000183778.19	56.4239	1.376702	0.394178	3.492589	0.000478	0.02385
ENSG00000196277.17	235.693	1.360461	0.312938	4.347389	1.38E-05	0.001085
ENSG00000137693.14	740.9689	1.302781	0.297461	4.379673	1.19E-05	0.000966
ENSG00000026025.16	14598.27	1.302244	0.298292	4.365664	1.27E-05	0.001018
ENSG00000112186.13	601.7475	1.283347	0.349447	3.67251	0.00024	0.013257
ENSG00000125848.10	314.0189	1.264846	0.367101	3.445497	0.00057	0.027442
ENSG00000135116.9	630.9575	1.234628	0.246314	5.012419	5.37E-07	6.58E-05
ENSG00000008311.16	79.33978	1.222599	0.347923	3.513994	0.000441	0.022466
ENSG00000156140.11	1975.369	1.214308	0.340946	3.561584	0.000369	0.019058
ENSG00000173705.9	99.51137	1.199542	0.319806	3.750841	0.000176	0.010127
ENSG00000196159.14	994.4055	1.184187	0.32839	3.606041	0.000311	0.016554

ENSG00000139292.13	7563.298	1.167748	0.311699	3.746393	0.000179	0.010278
ENSG00000144057.16	342.3703	1.15743	0.348867	3.317685	0.000908	0.041234
ENSG00000147257.16	1954.379	1.151691	0.288331	3.994338	6.49E-05	0.004399
ENSG00000114541.15	906.3614	1.150687	0.350441	3.283543	0.001025	0.045351
ENSG00000071967.12	284.6183	1.133428	0.292877	3.869978	0.000109	0.006771
ENSG00000287038.3	120.8925	1.097205	0.294902	3.720579	0.000199	0.011159
ENSG00000124942.14	2093.823	1.073615	0.307525	3.491149	0.000481	0.023918
ENSG00000011465.18	197.7865	1.058004	0.278182	3.80328	0.000143	0.008427
ENSG00000106066.15	257.5351	1.057717	0.264071	4.005423	6.19E-05	0.004212
ENSG00000152229.18	105.2504	1.048827	0.304789	3.441154	0.000579	0.027682
ENSG00000120324.9	212.9942	1.048665	0.310719	3.374959	0.000738	0.03469
ENSG00000170500.13	380.3275	1.047602	0.271693	3.855829	0.000115	0.007062
ENSG00000163430.13	5166.492	1.009549	0.289029	3.492895	0.000478	0.02385
ENSG00000174482.11	618.4849	1.004547	0.262749	3.823215	0.000132	0.007893

Significantly downregulated genes immature LIPT2-AS1$^{-/-}$ vs WT

gene_id	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj
ENSG00000131089.17	162.7014	-10.7276	1.229792	-8.72308	2.71E-18	2.74E-15
ENSG00000161681.17	43.22939	-8.81645	1.268196	-6.95196	3.6E-12	1.36E-09
ENSG00000233608.5	24.06909	-7.97266	1.304146	-6.11332	9.76E-10	2.52E-07
ENSG00000101282.9	16.89897	-7.46051	1.316762	-5.6658	1.46E-08	2.73E-06
ENSG00000215397.4	32.83646	-7.44877	1.265542	-5.88583	3.96E-09	8.44E-07
ENSG00000123609.11	12.17729	-6.99042	1.364942	-5.12141	3.03E-07	3.96E-05
ENSG00000248837.7	10.36761	-6.75167	1.397356	-4.83175	1.35E-06	0.00015
ENSG00000225968.8	9.564869	-6.641	1.395057	-4.76038	1.93E-06	0.000206
ENSG00000112343.11	17.23058	-6.51269	1.314862	-4.95314	7.3E-07	8.72E-05
ENSG00000261594.4	8.758054	-6.50928	1.440268	-4.51949	6.2E-06	0.000562
ENSG00000293490.1	8.631332	-6.4937	1.420246	-4.57223	4.83E-06	0.000448
ENSG00000253477.5	8.612611	-6.49073	1.452519	-4.4686	7.87E-06	0.000683
ENSG00000063015.21	54.21274	-6.25702	0.872381	-7.17235	7.37E-13	3.14E-10
ENSG00000186642.16	39.37911	-6.20772	0.982531	-6.31809	2.65E-10	7.74E-08
ENSG00000054356.14	250.7086	-6.1063	0.70031	-8.71943	2.8E-18	2.74E-15
ENSG00000144583.5	64.91166	-5.92323	0.726132	-8.15724	3.43E-16	2.49E-13
ENSG00000101180.17	20.78983	-5.86574	1.166114	-5.03016	4.9E-07	6.04E-05
ENSG00000155265.12	41.50949	-5.86381	0.849405	-6.90343	5.08E-12	1.81E-09
ENSG00000271830.1	5.246379	-5.76471	1.582306	-3.64323	0.000269	0.014573
ENSG00000104833.12	103.1892	-5.69671	0.676045	-8.42652	3.56E-17	3.03E-14
ENSG00000095713.14	4.844769	-5.66563	1.691508	-3.34945	0.00081	0.037596
ENSG00000163531.17	45.0262	-5.65449	0.781625	-7.23427	4.68E-13	2.18E-10

ENSG00000178171.12	34.89404	-5.60704	0.863556	-6.49297	8.42E-11	2.6E-08
ENSG00000157087.20	24.88723	-5.53357	1.009406	-5.482	4.21E-08	7.17E-06
ENSG00000166111.10	39.63735	-5.46633	0.819576	-6.6697	2.56E-11	8.66E-09
ENSG00000100095.19	61.24917	-5.40491	0.703504	-7.68283	1.56E-14	8.55E-12
ENSG00000116147.18	37.39918	-5.37614	0.883089	-6.08788	1.14E-09	2.91E-07
ENSG00000145920.15	202.8388	-5.3457	0.765625	-6.98214	2.91E-12	1.14E-09
ENSG00000196353.13	13.24491	-5.20299	1.184567	-4.39231	1.12E-05	0.000919
ENSG00000053438.11	186.4356	-4.91937	0.80451	-6.11474	9.67E-10	2.52E-07
ENSG00000112038.18	175.9927	-4.83762	0.414135	-11.6812	1.59E-31	7.79E-28
ENSG00000149527.18	105.1659	-4.83506	0.620095	-7.7973	6.32E-15	4E-12
ENSG00000272449.3	10.31328	-4.82829	1.238789	-3.89759	9.72E-05	0.006201
ENSG00000134107.5	20.43057	-4.80649	0.923224	-5.2062	1.93E-07	2.68E-05
ENSG00000105409.19	111.5163	-4.76101	0.790658	-6.02158	1.73E-09	4.03E-07
ENSG00000120645.12	14.44581	-4.71847	1.07055	-4.40752	1.05E-05	0.000868
ENSG00000013293.6	51.79695	-4.7087	0.5781	-8.14514	3.79E-16	2.65E-13
ENSG00000260903.3	205.1025	-4.66603	0.84641	-5.51273	3.53E-08	6.13E-06
ENSG00000116254.18	13.63742	-4.62739	1.099338	-4.20925	2.56E-05	0.001924
ENSG00000110076.21	155.8478	-4.6078	0.803793	-5.73256	9.89E-09	1.88E-06
ENSG00000171303.8	94.43751	-4.58653	0.737505	-6.21898	5E-10	1.42E-07
ENSG00000125851.10	33.67597	-4.50794	0.775556	-5.81253	6.15E-09	1.27E-06
ENSG00000154118.13	76.17102	-4.49417	0.775666	-5.79396	6.87E-09	1.39E-06
ENSG00000130054.5	36.74023	-4.46165	0.743407	-6.00163	1.95E-09	4.45E-07
ENSG00000170091.11	361.6546	-4.45047	0.3696	-12.0413	2.15E-33	1.41E-29
ENSG00000261678.3	97.31591	-4.4269	0.819161	-5.40418	6.51E-08	1.02E-05
ENSG00000164176.13	47.07335	-4.41196	0.625728	-7.05092	1.78E-12	7.11E-10
ENSG00000155052.15	62.05847	-4.41051	0.538529	-8.18992	2.61E-16	1.97E-13
ENSG00000124194.17	59.74858	-4.40741	0.758568	-5.81017	6.24E-09	1.27E-06
ENSG00000187714.7	49.8928	-4.40674	0.668004	-6.59687	4.2E-11	1.37E-08
ENSG00000183654.9	50.51707	-4.40652	0.605486	-7.27766	3.4E-13	1.66E-10
ENSG00000029534.22	47.06728	-4.39811	0.716471	-6.13858	8.33E-10	2.23E-07
ENSG00000180818.5	96.63755	-4.35638	0.757649	-5.74987	8.93E-09	1.73E-06
ENSG00000168993.15	147.1543	-4.35077	0.835622	-5.20663	1.92E-07	2.68E-05
ENSG00000174672.16	112.7323	-4.3089	0.801932	-5.37316	7.74E-08	1.18E-05
ENSG00000109956.13	87.71049	-4.2972	0.574364	-7.48167	7.34E-14	3.69E-11
ENSG00000130294.18	1060.429	-4.27696	0.480203	-8.90657	5.26E-19	6.07E-16
ENSG00000157064.11	144.6119	-4.27036	0.384439	-11.108	1.15E-28	2.81E-25
ENSG00000162105.21	35.68425	-4.25366	0.807941	-5.26481	1.4E-07	2.05E-05
ENSG00000198910.14	1008.971	-4.20077	0.582341	-7.2136	5.45E-13	2.48E-10
ENSG00000101489.20	106.4183	-4.15936	0.499562	-8.32601	8.36E-17	6.83E-14
ENSG00000277200.2	122.0602	-4.14783	0.849866	-4.88057	1.06E-06	0.000123
ENSG00000077080.10	174.2077	-4.13094	0.897544	-4.60249	4.17E-06	0.00039
ENSG00000205744.10	9.754551	-4.1286	1.080222	-3.822	0.000132	0.007908
ENSG00000139767.10	118.4101	-4.11528	0.66649	-6.17455	6.64E-10	1.86E-07
ENSG00000123610.5	12.2581	-4.09562	1.056036	-3.87829	0.000105	0.006564
ENSG00000101144.13	109.7679	-4.08494	0.70907	-5.76098	8.36E-09	1.64E-06
ENSG00000148408.14	174.7951	-4.08432	0.469109	-8.70654	3.13E-18	2.92E-15

ENSG00000198028.4	12.50868	-4.05698	1.000635	-4.0544	5.03E-05	0.003505
ENSG00000124191.18	143.4558	-4.05427	0.870369	-4.65811	3.19E-06	0.000311
ENSG00000167880.8	12.35122	-4.04906	0.989764	-4.09094	4.3E-05	0.003084
ENSG00000015592.17	33.95504	-4.03101	0.837907	-4.81081	1.5E-06	0.000163
ENSG00000288638.1	8.738268	-3.95542	1.150177	-3.43897	0.000584	0.027839
ENSG00000089169.15	172.2393	-3.95286	0.352642	-11.2093	3.67E-29	1.11E-25
ENSG00000118515.12	197.5715	-3.93785	0.549699	-7.16365	7.86E-13	3.27E-10
ENSG00000130477.16	629.0015	-3.91632	0.46061	-8.50246	1.86E-17	1.65E-14
ENSG00000077279.21	1632.287	-3.90047	0.258364	-15.0968	1.7E-51	3.33E-47
ENSG00000196092.14	131.3029	-3.88435	0.391128	-9.93115	3.05E-23	5.43E-20
ENSG00000111199.12	13.912	-3.88123	0.914392	-4.2446	2.19E-05	0.00167
ENSG00000144488.15	164.8695	-3.87736	0.912355	-4.24984	2.14E-05	0.001644
ENSG00000008056.14	73.58241	-3.871	0.876474	-4.41656	1E-05	0.00084
ENSG00000188488.14	13.57389	-3.84601	0.964542	-3.9874	6.68E-05	0.004498
ENSG00000120251.22	15.76407	-3.80995	0.807624	-4.71748	2.39E-06	0.000248
ENSG00000187957.8	60.60572	-3.76991	0.47473	-7.94117	2E-15	1.31E-12
ENSG00000142549.10	60.35585	-3.76114	0.793972	-4.73712	2.17E-06	0.00023
ENSG00000126259.20	12.86434	-3.76102	0.988427	-3.80505	0.000142	0.008392
ENSG00000100678.20	568.8698	-3.74563	0.406011	-9.22543	2.82E-20	3.95E-17
ENSG00000199572.1	25.1832	-3.71353	0.82815	-4.48412	7.32E-06	0.00064
ENSG00000087085.16	12.48364	-3.70933	0.959866	-3.86442	0.000111	0.006905
ENSG00000113763.12	38.09353	-3.70813	0.939428	-3.94722	7.91E-05	0.005199
ENSG00000259803.8	14.3569	-3.64929	0.900008	-4.05473	5.02E-05	0.003505
ENSG00000186487.21	115.7772	-3.62694	0.370412	-9.79163	1.22E-22	2E-19
ENSG00000177511.6	368.0602	-3.60399	0.377895	-9.53702	1.47E-21	2.22E-18
ENSG00000089558.9	25.79529	-3.60242	0.771198	-4.6712	2.99E-06	0.000298
ENSG00000124302.13	274.3363	-3.60061	0.915537	-3.93279	8.4E-05	0.005466
ENSG00000185551.15	42.1393	-3.5918	0.760364	-4.72379	2.31E-06	0.000241
ENSG00000129990.15	122.8125	-3.5558	0.930725	-3.82046	0.000133	0.007933
ENSG00000174521.8	36.3938	-3.55342	0.761211	-4.66811	3.04E-06	0.000301
ENSG00000011347.10	69.91471	-3.54585	0.579396	-6.1199	9.36E-10	2.48E-07
ENSG00000158089.15	158.1602	-3.54136	0.576688	-6.14087	8.21E-10	2.23E-07
ENSG00000274588.2	12.86716	-3.53228	0.897464	-3.93585	8.29E-05	0.005415
ENSG00000143217.9	8.716213	-3.51848	0.998697	-3.52307	0.000427	0.021767
ENSG00000279692.1	28.6339	-3.50012	0.964575	-3.62867	0.000285	0.015378
ENSG00000260802.2	42.53679	-3.4997	0.922535	-3.79356	0.000148	0.008738
ENSG00000133958.14	157.5441	-3.49291	0.335757	-10.4031	2.4E-25	4.7E-22
ENSG00000185615.16	89.1214	-3.46345	0.889598	-3.89328	9.89E-05	0.006271
ENSG00000177614.11	224.5273	-3.41518	0.662205	-5.15728	2.51E-07	3.36E-05
ENSG00000125966.10	150.1367	-3.39727	0.542776	-6.25906	3.87E-10	1.12E-07
ENSG00000149295.14	137.4318	-3.36685	0.944846	-3.56338	0.000366	0.018978
ENSG00000075043.21	3502.369	-3.3508	0.645165	-5.19371	2.06E-07	2.84E-05
ENSG00000131398.15	89.85303	-3.33977	0.437247	-7.63816	2.2E-14	1.17E-11
ENSG00000132821.12	233.0295	-3.3352	0.8053	-4.14156	3.45E-05	0.002541
ENSG00000135127.12	79.5895	-3.29932	0.661827	-4.98517	6.19E-07	7.53E-05
ENSG00000173826.15	52.64992	-3.29605	0.822716	-4.00631	6.17E-05	0.004211

ENSG00000070886.12	26.70244	-3.26927	0.865647	-3.77667	0.000159	0.009214
ENSG00000124557.14	11.05774	-3.2624	0.848359	-3.84554	0.00012	0.007319
ENSG00000167654.18	373.912	-3.25596	0.630034	-5.16791	2.37E-07	3.22E-05
ENSG00000101331.17	13.10187	-3.25378	0.870222	-3.73902	0.000185	0.010523
ENSG00000266074.11	149.3668	-3.25325	0.985447	-3.3013	0.000962	0.043151
ENSG00000121966.8	16.54369	-3.22167	0.874	-3.68612	0.000228	0.012603
ENSG00000100314.4	43.35232	-3.20554	0.621336	-5.15912	2.48E-07	3.35E-05
ENSG00000263874.3	36.03913	-3.18318	0.622737	-5.1116	3.19E-07	4.12E-05
ENSG00000056487.17	646.5482	-3.17553	0.348053	-9.1237	7.26E-20	9.48E-17
ENSG00000187122.17	17.87346	-3.17489	0.808733	-3.92576	8.65E-05	0.005609
ENSG00000156453.14	91.24038	-3.12177	0.802255	-3.89124	9.97E-05	0.006304
ENSG00000171450.6	479.0151	-3.07452	0.750296	-4.09774	4.17E-05	0.003005
ENSG00000255571.10	49.19754	-3.06273	0.921482	-3.3237	0.000888	0.040763
ENSG00000125878.7	28.92702	-3.05355	0.606464	-5.035	4.78E-07	5.93E-05
ENSG00000163618.18	29.5474	-3.01483	0.597007	-5.04991	4.42E-07	5.52E-05
ENSG00000156959.9	445.0534	-3.01284	0.78446	-3.84065	0.000123	0.007444
ENSG00000135144.8	227.3408	-2.99088	0.673661	-4.43974	9.01E-06	0.000764
ENSG00000128564.8	2799.923	-2.98179	0.806167	-3.69872	0.000217	0.012042
ENSG00000118160.14	158.5927	-2.9531	0.495976	-5.95411	2.61E-09	5.82E-07
ENSG00000083454.23	15.55956	-2.93943	0.889827	-3.30338	0.000955	0.04303
ENSG00000101210.14	2637.138	-2.93465	0.699269	-4.19674	2.71E-05	0.002017
ENSG00000112137.19	66.86494	-2.93403	0.462424	-6.34488	2.23E-10	6.61E-08
ENSG00000256982.3	28.29784	-2.92319	0.62083	-4.70853	2.5E-06	0.000252
ENSG00000139445.18	70.52368	-2.91859	0.483552	-6.03572	1.58E-09	3.78E-07
ENSG00000119973.6	62.35409	-2.89124	0.479388	-6.03112	1.63E-09	3.84E-07
ENSG00000100604.13	19848.48	-2.88631	0.54871	-5.26018	1.44E-07	2.08E-05
ENSG00000187416.13	19.65554	-2.88548	0.800166	-3.6061	0.000311	0.016554
ENSG00000126583.12	25.20683	-2.85804	0.817036	-3.49806	0.000469	0.023668
ENSG00000103056.12	32.48873	-2.84861	0.833202	-3.41887	0.000629	0.029906
ENSG00000143595.13	46.70739	-2.8238	0.464592	-6.07803	1.22E-09	3.06E-07
ENSG00000050767.18	23.04511	-2.81303	0.847287	-3.32004	0.0009	0.041013
ENSG00000134871.19	1383.816	-2.81185	0.722035	-3.89433	9.85E-05	0.006264
ENSG00000101098.13	593.6324	-2.81155	0.643945	-4.36613	1.26E-05	0.001018
ENSG00000119283.16	16.41724	-2.79318	0.679995	-4.10766	4E-05	0.00289
ENSG00000198959.12	294.0115	-2.79154	0.517545	-5.39381	6.9E-08	1.07E-05
ENSG00000158258.16	693.5215	-2.78565	0.253533	-10.9873	4.4E-28	9.57E-25
ENSG00000144218.21	111.175	-2.7226	0.378036	-7.20196	5.94E-13	2.64E-10
ENSG00000129993.15	70.03922	-2.68583	0.540655	-4.96774	6.77E-07	8.19E-05
ENSG00000087116.17	26.99462	-2.67569	0.589877	-4.53602	5.73E-06	0.000522
ENSG00000116544.12	72.64041	-2.66736	0.691644	-3.85655	0.000115	0.007062
ENSG00000285822.1	82.80056	-2.62673	0.539241	-4.87116	1.11E-06	0.000127
ENSG00000036565.15	1475.147	-2.59669	0.587804	-4.4176	9.98E-06	0.000839
ENSG00000221890.6	76.76601	-2.56925	0.588744	-4.36395	1.28E-05	0.001022
ENSG00000181790.13	82.86617	-2.53741	0.620642	-4.08837	4.34E-05	0.003107
ENSG00000159753.15	120.4117	-2.50541	0.726143	-3.45031	0.00056	0.027091
ENSG00000164076.17	275.055	-2.50186	0.559319	-4.47304	7.71E-06	0.000672

ENSG00000087258.16	723.5131	-2.49492	0.473642	-5.26753	1.38E-07	2.04E-05
ENSG00000100234.12	532.6479	-2.48671	0.284216	-8.74936	2.15E-18	2.34E-15
ENSG00000187498.16	988.4114	-2.45933	0.317694	-7.74119	9.85E-15	5.85E-12
ENSG00000137699.17	83.9085	-2.45191	0.701347	-3.496	0.000472	0.023785
ENSG00000113504.21	94.11708	-2.44814	0.53013	-4.61799	3.87E-06	0.000369
ENSG00000175206.11	29.55031	-2.43852	0.617344	-3.95002	7.81E-05	0.005155
ENSG00000129159.9	142.0249	-2.42379	0.674928	-3.59118	0.000329	0.017292
ENSG00000116983.13	258.2817	-2.41249	0.29221	-8.25603	1.51E-16	1.18E-13
ENSG00000130540.14	50.52492	-2.40122	0.456545	-5.25954	1.44E-07	2.08E-05
ENSG00000130222.11	102.004	-2.39277	0.620471	-3.85638	0.000115	0.007062
ENSG00000072832.15	2468.299	-2.35256	0.406228	-5.79124	6.99E-09	1.4E-06
ENSG00000059915.17	155.2732	-2.34361	0.528445	-4.43491	9.21E-06	0.000778
ENSG00000171631.16	57.87166	-2.32734	0.428445	-5.43206	5.57E-08	9.02E-06
ENSG00000287431.1	90.29029	-2.32576	0.582701	-3.99135	6.57E-05	0.004439
ENSG00000185090.15	631.9094	-2.32081	0.614713	-3.77544	0.00016	0.009232
ENSG00000179431.7	225.2621	-2.27085	0.629968	-3.6047	0.000313	0.016594
ENSG00000078549.16	327.0996	-2.25771	0.436879	-5.16781	2.37E-07	3.22E-05
ENSG00000164082.15	20.72211	-2.24996	0.599641	-3.75219	0.000175	0.010102
ENSG00000099994.11	177.9262	-2.23967	0.625351	-3.58146	0.000342	0.017901
ENSG00000169871.13	418.8818	-2.23252	0.466882	-4.78176	1.74E-06	0.000186
ENSG00000131055.5	60.97328	-2.16212	0.474858	-4.55319	5.28E-06	0.000488
ENSG00000172020.13	300.1811	-2.15987	0.278703	-7.74973	9.21E-15	5.64E-12
ENSG00000154556.20	237.926	-2.14477	0.29984	-7.15306	8.49E-13	3.46E-10
ENSG00000095739.11	175.1246	-2.13012	0.354625	-6.00668	1.89E-09	4.37E-07
ENSG00000182912.6	237.3011	-2.1084	0.532826	-3.95702	7.59E-05	0.005024
ENSG00000176406.25	96.13817	-2.08658	0.547194	-3.81325	0.000137	0.008144
ENSG00000166257.10	203.0275	-2.07252	0.29949	-6.92016	4.51E-12	1.64E-09
ENSG00000148798.11	1618.116	-2.06892	0.231411	-8.94049	3.87E-19	4.74E-16
ENSG00000167619.13	207.6973	-2.06669	0.395629	-5.2238	1.75E-07	2.49E-05
ENSG00000166073.12	847.8745	-2.06144	0.402656	-5.11961	3.06E-07	3.97E-05
ENSG00000135048.14	36.73232	-2.05658	0.441383	-4.6594	3.17E-06	0.000311
ENSG00000225783.9	45.31358	-2.04041	0.532179	-3.83406	0.000126	0.007576
ENSG00000155093.20	400.4949	-2.03366	0.394575	-5.15406	2.55E-07	3.4E-05
ENSG00000151136.15	59.498	-2.01388	0.384906	-5.23213	1.68E-07	2.4E-05
ENSG00000054803.4	450.4806	-2.00347	0.342729	-5.84564	5.05E-09	1.05E-06
ENSG00000067842.19	48.17784	-1.99412	0.441825	-4.51336	6.38E-06	0.000574
ENSG00000196132.14	198.6177	-1.98054	0.330329	-5.99567	2.03E-09	4.56E-07
ENSG00000110328.6	205.2742	-1.97447	0.4784	-4.12723	3.67E-05	0.002684
ENSG00000080031.10	92.63215	-1.96677	0.519583	-3.78529	0.000154	0.008953
ENSG00000258986.7	241.0052	-1.94309	0.337094	-5.76424	8.2E-09	1.62E-06
ENSG00000122254.7	393.9486	-1.93309	0.397603	-4.86187	1.16E-06	0.000132
ENSG00000144406.19	140.7877	-1.92233	0.352499	-5.45344	4.94E-08	8.13E-06
ENSG00000028277.22	18.97048	-1.91742	0.579095	-3.31107	0.000929	0.042058
ENSG00000140481.15	69.06559	-1.90628	0.421993	-4.51732	6.26E-06	0.000565
ENSG00000091129.22	1891.792	-1.90574	0.377344	-5.0504	4.41E-07	5.52E-05
ENSG00000112715.26	775.6784	-1.90101	0.322654	-5.8918	3.82E-09	8.23E-07

ENSG00000107742.14	51.05245	-1.89815	0.402888	-4.71135	2.46E-06	0.000252
ENSG00000107338.10	167.4303	-1.89487	0.529927	-3.57572	0.000349	0.018161
ENSG00000106089.12	305.3241	-1.88864	0.543513	-3.47488	0.000511	0.025098
ENSG00000140323.6	208.9019	-1.88858	0.567318	-3.32897	0.000872	0.040094
ENSG00000138083.5	430.8581	-1.88344	0.388544	-4.84742	1.25E-06	0.00014
ENSG00000204653.10	30.98997	-1.86175	0.520684	-3.57559	0.000349	0.018161
ENSG00000103489.12	1186.834	-1.85205	0.338674	-5.46855	4.54E-08	7.53E-06
ENSG00000184613.11	57.48307	-1.85199	0.441271	-4.19694	2.71E-05	0.002017
ENSG00000109832.14	60.95292	-1.82581	0.490306	-3.72382	0.000196	0.011049
ENSG00000119714.11	103.7524	-1.81325	0.429324	-4.2235	2.41E-05	0.001813
ENSG00000276168.1	49.12755	-1.81283	0.502502	-3.6076	0.000309	0.016545
ENSG00000116981.4	711.7714	-1.79854	0.40385	-4.45347	8.45E-06	0.000729
ENSG00000129757.15	827.7734	-1.79209	0.410917	-4.3612	1.29E-05	0.00103
ENSG00000152954.12	120.4505	-1.78565	0.329171	-5.42469	5.81E-08	9.25E-06
ENSG00000168209.6	1182.677	-1.78374	0.540159	-3.30226	0.000959	0.043102
ENSG00000183780.13	42.81582	-1.78258	0.436127	-4.0873	4.36E-05	0.003109
ENSG00000196361.10	2493.247	-1.77424	0.362774	-4.89075	1E-06	0.000118
ENSG00000118733.17	77.09168	-1.7682	0.41529	-4.25776	2.06E-05	0.001599
ENSG00000185008.19	220.3863	-1.75272	0.284262	-6.16586	7.01E-10	1.93E-07
ENSG00000166292.12	282.0628	-1.74831	0.319614	-5.47007	4.5E-08	7.53E-06
ENSG00000147642.17	58.40451	-1.71291	0.388213	-4.41229	1.02E-05	0.000853
ENSG00000290803.1	122.1466	-1.69109	0.302136	-5.59713	2.18E-08	3.92E-06
ENSG00000139832.5	57.6796	-1.68846	0.50812	-3.32297	0.000891	0.040775
ENSG00000113739.11	73.51942	-1.66725	0.371097	-4.49276	7.03E-06	0.000621
ENSG00000164061.5	367.2584	-1.65175	0.474793	-3.4789	0.000503	0.024787
ENSG00000198914.5	445.4096	-1.61216	0.334934	-4.81337	1.48E-06	0.000162
ENSG00000143333.7	476.2862	-1.57047	0.29216	-5.37537	7.64E-08	1.18E-05
ENSG00000104722.14	1090.646	-1.56703	0.244748	-6.40265	1.53E-10	4.6E-08
ENSG00000111859.17	417.9634	-1.55605	0.291221	-5.34321	9.13E-08	1.38E-05
ENSG00000154319.16	1196.518	-1.55176	0.341325	-4.54629	5.46E-06	0.0005
ENSG00000148082.10	603.0908	-1.5424	0.288694	-5.3427	9.16E-08	1.38E-05
ENSG00000076356.7	540.7942	-1.53928	0.413012	-3.72697	0.000194	0.010975
ENSG00000162433.15	416.5372	-1.5382	0.282652	-5.44205	5.27E-08	8.6E-06
ENSG00000282381.2	32.44695	-1.52976	0.44429	-3.44315	0.000575	0.027607
ENSG00000140479.18	262.1836	-1.52599	0.354423	-4.30556	1.67E-05	0.001305
ENSG00000102048.16	130.4495	-1.51887	0.298592	-5.08677	3.64E-07	4.66E-05
ENSG00000111249.14	241.5155	-1.5005	0.344225	-4.35906	1.31E-05	0.001036
ENSG00000171951.5	452.0707	-1.47865	0.312274	-4.73512	2.19E-06	0.000231
ENSG00000152894.15	1560.796	-1.47657	0.336263	-4.39111	1.13E-05	0.000921
ENSG00000165868.16	413.227	-1.43945	0.295853	-4.86542	1.14E-06	0.00013
ENSG00000204956.6	48.67342	-1.43454	0.412995	-3.47352	0.000514	0.025163
ENSG00000143858.12	92.61553	-1.43432	0.400589	-3.58052	0.000343	0.017917
ENSG00000132437.18	2693.72	-1.41129	0.216087	-6.53114	6.53E-11	2.06E-08
ENSG00000259439.3	95.22746	-1.38213	0.344563	-4.01125	6.04E-05	0.004152
ENSG00000248538.10	52.10382	-1.3774	0.381793	-3.60772	0.000309	0.016545
ENSG00000066032.19	296.7852	-1.37236	0.259487	-5.28873	1.23E-07	1.83E-05

ENSG00000103723.17	599.6136	-1.34573	0.359631	-3.74196	0.000183	0.010431
ENSG00000056972.21	105.5244	-1.33252	0.341125	-3.90626	9.37E-05	0.006022
ENSG00000133401.16	109.8886	-1.31538	0.338613	-3.88463	0.000102	0.006436
ENSG00000130558.20	196.2454	-1.31158	0.272349	-4.81579	1.47E-06	0.00016
ENSG00000144668.12	383.5604	-1.29707	0.264181	-4.90979	9.12E-07	0.000108
ENSG00000100285.10	710.1115	-1.29083	0.335449	-3.84805	0.000119	0.007268
ENSG00000179348.13	11464.41	-1.24878	0.380395	-3.28284	0.001028	0.045351
ENSG00000134594.5	198.2172	-1.22809	0.374021	-3.28347	0.001025	0.045351
ENSG00000131378.14	2113.786	-1.21668	0.250463	-4.85773	1.19E-06	0.000134
ENSG00000171533.12	445.3843	-1.21334	0.306082	-3.96412	7.37E-05	0.004893
ENSG00000106236.4	1019.156	-1.18971	0.335349	-3.54769	0.000389	0.019934
ENSG00000104044.16	85.22795	-1.1826	0.353319	-3.34711	0.000817	0.037826
ENSG00000125207.7	207.6411	-1.17448	0.326565	-3.59647	0.000323	0.017036
ENSG00000171867.18	79.44889	-1.14467	0.344545	-3.32226	0.000893	0.040782
ENSG00000075643.6	263.1963	-1.12217	0.275098	-4.07917	4.52E-05	0.003197
ENSG00000197381.18	313.0887	-1.12124	0.261721	-4.28412	1.83E-05	0.001426
ENSG00000221866.10	230.8032	-1.1163	0.304688	-3.66376	0.000249	0.013603
ENSG00000180287.17	244.4934	-1.11574	0.301681	-3.69842	0.000217	0.012042
ENSG00000168824.15	794.1461	-1.09058	0.230781	-4.72563	2.29E-06	0.00024
ENSG00000087495.17	175.2242	-1.0389	0.299583	-3.46783	0.000525	0.025574
ENSG00000198879.13	573.5777	-1.02085	0.279756	-3.64907	0.000263	0.014325
ENSG00000171617.16	1019.668	-1.01432	0.25242	-4.01836	5.86E-05	0.004058
ENSG00000125869.10	622.3402	-1.00809	0.301905	-3.3391	0.000841	0.038842

Proteins enriched >1.5 log_2Fold in LIPT2-AS1 pulldown vs FLAG alone pulldown

ADD1	MAP1LC3B	H2AC14	PTN
AP2M1	MAP1LC3B2	H2AC16	RAD23B
ARF5	MOGS	H2AC18	RAP1A
ATP5MG	MPDU1	H2AC20	RAP1B
BAG2	MRPL21	H2AC25	RCL1
BLMH	MRPL30	H2AC6	RPA3
C7orf50	MRPS18B	H2AC7	RPL22
CCDC124	MTHFR	H2AC8	RPS14
CDC5L	MYEF2	H2AJ	RPSA
CHCHD3	NAMPT	H2AX	RTN4
CLTA	OCIAD1	HNRNPH3	S100A6

CNIH4	PFN2	IPO8	S100A8
COPG1	POLG2	KARS1	S100A9
COX4I1	POLR2H	KATNAL2	SHROOM3
COX6C	PPIH	LSM12	SLC25A10
DDX1	PPIL3	LSM14A	SORCS1
E2F7	PPP2CA	LZTS3	TMPO
EPDR1	PPP2CB	SRSF10	TRIP11
FARSA	PRKAR1A	SSBP1	UPF1
GADD45GIP1	PRPF4	SSRP1	WDR83
H1-3	PSMB3	STBD1	YWHAE
H2AC1	PSMC2	SUPT16H	YWHAH
H2AC12	PSPC1	TAB3	ZC2HC1A

References

Ackerman, S. (1992). *Discovering the Brain.*

Ahrens, C. H., Wade, J. T., Champion, M. M., & Langer, J. D. (2022). A Practical Guide to Small Protein Discovery and Characterization Using Mass Spectrometry. *Journal of Bacteriology*, *204*(1). https://doi.org/10.1128/jb.00353-21

Albert, M., & Huttner, W. B. (2018). Epigenetic and Transcriptional Pre-patterning—An Emerging Theme in Cortical Neurogenesis. *Frontiers in Neuroscience*, *12*. https://doi.org/10.3389/fnins.2018.00359

Almagro Armenteros, J. J., Salvatore, M., Emanuelsson, O., Winther, O., von Heijne, G., Elofsson, A., & Nielsen, H. (2019). Detecting sequence signals in targeting peptides using deep learning. *Life Science Alliance*, *2*(5), e201900429. https://doi.org/10.26508/lsa.201900429

An, N. A., Zhang, J., Mo, F., Luan, X., Tian, L., Shen, Q. S., Li, X., Li, C., Zhou, F., Zhang, B., Ji, M., Qi, J., Zhou, W.-Z., Ding, W., Chen, J.-Y., Yu, J., Zhang, L., Shu, S., Hu, B., & Li, C.-Y. (2023). De novo genes with an lncRNA origin encode unique human brain developmental functionality. *Nature Ecology & Evolution*. https://doi.org/10.1038/s41559-022-01925-6

Andreev DE, Loughran G, Fedorova AD, Mikhaylova MS, Shatsky IN, Baranov PV. Non-AUG translation initiation in mammals. Genome Biol. 2022;23(1):111. Published 2022 May 9. doi:10.1186/s13059-022-02674-2

Anderson, D. M., Anderson, K. M., Chang, C.-L., Makarewich, C. A., Nelson, B. R., McAnally, J. R., Kasaragod, P., Shelton, J. M., Liou, J., Bassel-Duby, R., & Olson, E. N. (2015). A Micropeptide Encoded by a Putative Long Noncoding RNA Regulates Muscle Performance. *Cell*, *160*(4), 595–606. https://doi.org/10.1016/j.cell.2015.01.009

Asmann, Y. W., Necela, B. M., Kalari, K. R., Hossain, A., Baker, T. R., Carr, J. M., Davis, C., Getz, J. E., Hostetter, G., Li, X., McLaughlin, S. A., Radisky, D. C., Schroth, G. P., Cunliffe, H. E., Perez, E. A., & Thompson, E. A. (2012). Detection of Redundant Fusion Transcripts as Biomarkers or Disease-Specific Therapeutic Targets in Breast Cancer. *Cancer Research*, *72*(8), 1921–1928. https://doi.org/10.1158/0008-5472.CAN-11-3142

Aspden, J. L., Eyre-Walker, Y. C., Phillips, R. J., Amin, U., Mumtaz, M. A. S., Brocard, M., & Couso, J.-P. (2014). Extensive translation of small Open Reading Frames revealed by Poly-Ribo-Seq. *ELife*, *3*. https://doi.org/10.7554/eLife.03528

Bao, Z., Yang, Z., Huang, Z., Zhou, Y., Cui, Q., & Dong, D. (2019). LncRNADisease 2.0: an updated database of long non-coding RNA-associated diseases. *Nucleic Acids Research*, *47*(D1), D1034–D1037. https://doi.org/10.1093/nar/gky905

Barbosa-Morais, N. L., Irimia, M., Pan, Q., Xiong, H. Y., Gueroussov, S., Lee, L. J., Slobodeniuc, V., Kutter, C., Watt, S., Çolak, R., Kim, T., Misquitta-Ali, C. M., Wilson, M. D., Kim, P. M., Odom, D. T., Frey, B. J., & Blencowe, B. J. (2012). The Evolutionary Landscape of

Alternative Splicing in Vertebrate Species. *Science*, *338*(6114), 1587–1593. https://doi.org/10.1126/science.1230612

Basrai, M. A., Hieter, P., & Boeke, J. D. (1997). Small Open Reading Frames: Beautiful Needles in the Haystack. *Genome Research*, *7*(8), 768–771. https://doi.org/10.1101/gr.7.8.768

Bazzini, A. A., Johnstone, T. G., Christiano, R., Mackowiak, S. D., Obermayer, B., Fleming, E. S., Vejnar, C. E., Lee, M. T., Rajewsky, N., Walther, T. C., & Giraldez, A. J. (2014). Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. *The EMBO Journal*, *33*(9), 981–993. https://doi.org/10.1002/embj.201488411

Beadle, G. W., & Tatum, E. L. (1941). Genetic Control of Biochemical Reactions in Neurospora. *Proceedings of the National Academy of Sciences*, *27*(11), 499–506. https://doi.org/10.1073/pnas.27.11.499

Bell, M., & Zempel, H. (2022). SH-SY5Y-derived neurons: a human neuronal model system for investigating TAU sorting and neuronal subtype-specific TAU vulnerability. *Reviews in the Neurosciences*, *33*(1), 1–15. https://doi.org/10.1515/revneuro-2020-0152

Bell, N., Hann, V., Redfern, C. P. F., & Cheek, T. R. (2013). Store-operated Ca2+ entry in proliferating and retinoic acid-differentiated N- and S-type neuroblastoma cells. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1833*(3), 643–651. https://doi.org/10.1016/j.bbamcr.2012.11.025

Biever, A., Glock, C., Tushev, G., Ciirdaeva, E., Dalmay, T., Langer, J. D., & Schuman, E. M. (2020). Monosomes actively translate synaptic mRNAs in neuronal processes. *Science*, *367*(6477). https://doi.org/10.1126/science.aay4991

BOLTE, S., & CORDELIÈRES, F. P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. *Journal of Microscopy*, *224*(3), 213–232. https://doi.org/10.1111/j.1365-2818.2006.01706.x

Brunet, M. A., Leblanc, S., & Roucou, X. (2020). Reconsidering proteomic diversity with functional investigation of small ORFs and alternative ORFs. *Experimental Cell Research*, *393*(1), 112057. https://doi.org/10.1016/j.yexcr.2020.112057

Burke, E. E., Chenoweth, J. G., Shin, J. H., Collado-Torres, L., Kim, S.-K., Micali, N., Wang, Y., Colantuoni, C., Straub, R. E., Hoeppner, D. J., Chen, H.-Y., Sellers, A., Shibbani, K., Hamersky, G. R., Diaz Bustamante, M., Phan, B. N., Ulrich, W. S., Valencia, C., Jaishankar, A., … Jaffe, A. E. (2020). Dissecting transcriptomic signatures of neuronal differentiation and maturation using iPSCs. *Nature Communications*, *11*(1), 462. https://doi.org/10.1038/s41467-019-14266-z

Butti, Z., & Patten, S. A. (2019). RNA Dysregulation in Amyotrophic Lateral Sclerosis. *Frontiers in Genetics*, *9*. https://doi.org/10.3389/fgene.2018.00712

Byrne, A., Beaudin, A. E., Olsen, H. E., Jain, M., Cole, C., Palmer, T., DuBois, R. M., Forsberg, E. C., Akeson, M., & Vollmers, C. (2017). Nanopore long-read RNAseq reveals widespread transcriptional variation among the surface receptors of individual B cells. *Nature Communications*, *8*(1), 16027. https://doi.org/10.1038/ncomms16027

Cai, C.-L., Martin, J. C., Sun, Y., Cui, L., Wang, L., Ouyang, K., Yang, L., Bu, L., Liang, X., Zhang, X., Stallcup, W. B., Denton, C. P., McCulloch, A., Chen, J., & Evans, S. M. (2008). A myocardial lineage derives from Tbx18 epicardial cells. *Nature*, *454*(7200), 104–108. https://doi.org/10.1038/nature06969

Cai, X., & Cullen, B. R. (2007). The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA*, *13*(3), 313–316. https://doi.org/10.1261/rna.351707

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

Campalans, A., Kondorosi, A., & Crespi, M. (2004). Enod40, a Short Open Reading Frame–Containing mRNA, Induces Cytoplasmic Localization of a Nuclear RNA Binding Protein in Medicago truncatula. *The Plant Cell*, *16*(4), 1047–1059. https://doi.org/10.1105/tpc.019406

Candeias, M. M., Malbert-Colas, L., Powell, D. J., Daskalogianni, C., Maslon, M. M., Naski, N., Bourougaa, K., Calvo, F., & Fåhraeus, R. (2008). p53 mRNA controls p53 activity by managing Mdm2 functions. *Nature Cell Biology*, *10*(9), 1098–1105. https://doi.org/10.1038/ncb1770

Cao X, Slavoff SA. Non-AUG start codons: Expanding and regulating the small and alternative ORFeome. Exp Cell Res. 2020;391(1):111973. doi:10.1016/j.yexcr.2020.111973

Cao, B., Wang, T., Qu, Q., Kang, T., & Yang, Q. (2018). Long Noncoding RNA SNHG1 Promotes Neuroinflammation in Parkinson's Disease via Regulating miR-7/NLRP3 Pathway. *Neuroscience*, *388*, 118–127. https://doi.org/10.1016/j.neuroscience.2018.07.019

Cardoso-Moreira, M., Halbert, J., Valloton, D., Velten, B., Chen, C., Shao, Y., Liechti, A., Ascenção, K., Rummel, C., Ovchinnikova, S., Mazin, P. V., Xenarios, I., Harshman, K., Mort, M., Cooper, D. N., Sandi, C., Soares, M. J., Ferreira, P. G., Afonso, S., … Kaessmann, H. (2019). Gene expression across mammalian organ development. *Nature*, *571*(7766), 505–509. https://doi.org/10.1038/s41586-019-1338-5

Carlevaro-Fita, J., Lanzós, A., Feuerbach, L., Hong, C., Mas-Ponte, D., Pedersen, J. S., Abascal, F., Amin, S. B., Bader, G. D., Barenboim, J., Beroukhim, R., Bertl, J., Boroevich, K. A., Brunak, S., Campbell, P. J., Carlevaro-Fita, J., Chakravarty, D., Chan, C. W. Y., Chen, K., … von Mering, C. (2020). Cancer LncRNA Census reveals evidence for deep functional conservation of long noncoding RNAs in tumorigenesis. *Communications Biology*, *3*(1), 56. https://doi.org/10.1038/s42003-019-0741-7

Carlevaro-Fita, J., Rahim, A., Guigó, R., Vardy, L. A., & Johnson, R. (2016). Cytoplasmic long noncoding RNAs are frequently bound to and degraded at ribosomes in human cells. *RNA*, *22*(6), 867–882. https://doi.org/10.1261/rna.053561.115

Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., Fedele, S., Pesce, E., Ferrer, I., Collavin, L., Santoro, C., Forrest, A. R. R., Carninci, P., Biffo, S., Stupka, E., & Gustincich, S. (2012). Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature*, *491*(7424), 454–457. https://doi.org/10.1038/nature11508

Chen, W., Zhao, H., & Li, Y. (2023). Mitochondrial dynamics in health and disease: mechanisms and potential targets. *Signal Transduction and Targeted Therapy*, *8*(1), 333. https://doi.org/10.1038/s41392-023-01547-9

Cheng, W.-Y., Kandel, J. J., Yamashiro, D. J., Canoll, P., & Anastassiou, D. (2012). A Multi-Cancer Mesenchymal Transition Gene Expression Signature Is Associated with Prolonged Time to Recurrence in Glioblastoma. *PLoS ONE*, *7*(4), e34705. https://doi.org/10.1371/journal.pone.0034705

Chong, C., Müller, M., Pak, H., Harnett, D., Huber, F., Grun, D., Leleu, M., Auger, A., Arnaud, M., Stevenson, B. J., Michaux, J., Bilic, I., Hirsekorn, A., Calviello, L., Simó-Riudalbas, L., Planet, E., Lubiński, J., Bryśkiewicz, M., Wiznerowicz, M., … Bassani-Sternberg, M. (2020). Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. *Nature Communications*, *11*(1), 1293. https://doi.org/10.1038/s41467-020-14968-9

Chu, Q., Martinez, T. F., Novak, S. W., Donaldson, C. J., Tan, D., Vaughan, J. M., Chang, T., Diedrich, J. K., Andrade, L., Kim, A., Zhang, T., Manor, U., & Saghatelian, A. (2019). Regulation of the ER stress response by a mitochondrial microprotein. *Nature Communications*, *10*(1), 4883. https://doi.org/10.1038/s41467-019-12816-z

Chung, D. W., Rudnicki, D. D., Yu, L., & Margolis, R. L. (2011). A natural antisense transcript at the Huntington's disease repeat locus regulates HTT expression. *Human Molecular Genetics*, *20*(17), 3467–3477. https://doi.org/10.1093/hmg/ddr263

Clauwaert, J., McVey, Z., Gupta, R., & Menschaert, G. (2023). TIS Transformer: remapping the human proteome using deep learning. *NAR Genomics and Bioinformatics*, *5*(1). https://doi.org/10.1093/nargab/lqad021

Cohen, N., Betts, D. R., Rechavi, G., Amariglio, N., & Trakhtenbrot, L. (2003). Clonal expansion and not cell interconversion is the basis for the neuroblast and nonneuronal types of the SK-N-SH neuroblastoma cell line. *Cancer Genetics and Cytogenetics*, *143*(1), 80–84. https://doi.org/10.1016/S0165-4608(02)00835-X

Couso, J.-P., & Patraquim, P. (2017). Classification and function of small open reading frames. *Nature Reviews Molecular Cell Biology*, *18*(9), 575–589. https://doi.org/10.1038/nrm.2017.58

Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, *26*(12), 1367–1372. https://doi.org/10.1038/nbt.1511

de Groot, J. W. B., Links, T. P., Plukker, J. T. M., Lips, C. J. M., & Hofstra, R. M. W. (2006). RET as a Diagnostic and Therapeutic Target in Sporadic and Hereditary Endocrine Tumors. *Endocrine Reviews*, *27*(5), 535–560. https://doi.org/10.1210/er.2006-0017

Deng, B., Cheng, X., Li, H., Qin, J., Tian, M., & Jin, G. (2017). Microarray expression profiling in the denervated hippocampus identifies long noncoding RNAs functionally involved in neurogenesis. *BMC Molecular Biology*, *18*(1), 15. https://doi.org/10.1186/s12867-017-0091-2

Deng, J., Xu, W., Jie, Y., & Chong, Y. (2023). Subcellular localization and relevant mechanisms of human cancer-related micropeptides. *The FASEB Journal*, *37*(12). https://doi.org/10.1096/fj.202301019RR

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D. G., Lagarde, J., Veeravalli, L., Ruan, X., Ruan, Y., Lassmann, T., Carninci, P., Brown, J. B., Lipovich, L., Gonzalez, J. M., … Guigó, R. (2012). The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Research*, *22*(9), 1775–1789. https://doi.org/10.1101/gr.132159.111

Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., Xue, C., Marinov, G. K., Khatun, J., Williams, B. A., Zaleski, C., Rozowsky, J., Röder, M., Kokocinski, F., Abdelhamid, R. F., … Gingeras, T. R. (2012). Landscape of transcription in human cells. *Nature*, *489*(7414), 101–108. https://doi.org/10.1038/nature11233

Dong, F., Liu, Y., Yan, W., Meng, Q., Song, X., Cheng, B., & Yao, R. (2023). Netrin-4: Focus on Its Role in Axon Guidance, Tissue Stability, Angiogenesis and Tumors. *Cellular and Molecular Neurobiology*, *43*(5), 1663–1683. https://doi.org/10.1007/s10571-022-01279-4

Douka, K., Birds, I., Wang, D., Kosteletos, A., Clayton, S., Byford, A., Vasconcelos, E. J. R., O'Connell, M. J., Deuchars, J., Whitehouse, A., & Aspden, J. L. (2021). Cytoplasmic long noncoding RNAs are differentially regulated and translated during human neuronal differentiation. *RNA*, *27*(9), 1082–1101. https://doi.org/10.1261/rna.078782.121

Dravid, A., Raos, B., Svirskis, D., & O'Carroll, S. J. (2021). Optimised techniques for high-throughput screening of differentiated SH-SY5Y cells and application for neurite outgrowth assays. *Scientific Reports*, *11*(1), 23935. https://doi.org/10.1038/s41598-021-03442-1

Dwane, S., Durack, E., & Kiely, P. A. (2013). Optimising parameters for the differentiation of SH-SY5Y cells to study cell adhesion and cell migration. *BMC Research Notes*, *6*(1), 366. https://doi.org/10.1186/1756-0500-6-366

Eigenhuis, K. N., Somsen, H. B., van der Kroeg, M., Smeenk, H., Korporaal, A. L., Kushner, S. A., de Vrij, F. M. S., & van den Berg, D. L. C. (2023). A simplified protocol for the generation of cortical brain organoids. *Frontiers in Cellular Neuroscience*, *17*. https://doi.org/10.3389/fncel.2023.1114420

Erb, C. (2006). Embryology and Teratology. In *The Laboratory Rat* (pp. 817–846). Elsevier. https://doi.org/10.1016/B978-012074903-4/50031-5

Fabre, B., Combier, J.-P., & Plaza, S. (2021). Recent advances in mass spectrometry–based peptidomics workflows to identify short-open-reading-frame-encoded peptides and explore their functions. *Current Opinion in Chemical Biology*, *60*, 122–130. https://doi.org/10.1016/j.cbpa.2020.12.002

Fedorova AD, Kiniry SJ, Andreev DE, Mudge JM, Baranov PV. Thousands of human non-AUG extended proteoforms lack evidence of evolutionary selection among mammals. Nat Commun. 2022;13(1):7910. Published 2022 Dec 23. doi:10.1038/s41467-022-35595-6

Feles, S., Overath, C., Reichardt, S., Diegeler, S., Schmitz, C., Kronenberg, J., Baumstark-Khan, C., Hemmersbach, R., Hellweg, C. E., & Liemersdorf, C. (2022). Streamlining Culture Conditions for the Neuroblastoma Cell Line SH-SY5Y: A Prerequisite for Functional Studies. *Methods and Protocols*, *5*(4), 58. https://doi.org/10.3390/mps5040058

Fernández, V., Llinares-Benadero, C., & Borrell, V. (2016). Cerebral cortex expansion and folding: what have we learned? *The EMBO Journal*, *35*(10), 1021–1044. https://doi.org/10.15252/embj.201593701

Flippo, K. H., & Strack, S. (2017). Mitochondrial dynamics in neuronal injury, development and plasticity. *Journal of Cell Science*. https://doi.org/10.1242/jcs.171017

Formosa, T., & Winston, F. (2020). The role of FACT in managing chromatin: disruption, assembly, or repair? *Nucleic Acids Research*, *48*(21), 11929–11941. https://doi.org/10.1093/nar/gkaa912

Forster, J. I., Köglsberger, S., Trefois, C., Boyd, O., Baumuratov, A. S., Buck, L., Balling, R., & Antony, P. M. A. (2016). Characterization of Differentiated SH-SY5Y as Neuronal Screening Model Reveals Increased Oxidative Vulnerability. *SLAS Discovery*, *21*(5), 496–509. https://doi.org/10.1177/1087057115625190

Fotuhi, S. N., Khalaj-Kondori, M., Hoseinpour Feizi, M. A., & Talebi, M. (2019). Long Non-coding RNA BACE1-AS May Serve as an Alzheimer's Disease Blood-Based Biomarker. *Journal of Molecular Neuroscience*, *69*(3), 351–359. https://doi.org/10.1007/s12031-019-01364-2

Fu, H., Wang, T., Kong, X., Yan, K., Yang, Y., Cao, J., Yuan, Y., Wang, N., Kee, K., Lu, Z. J., & Xi, Q. (2022). A Nodal enhanced micropeptide NEMEP regulates glucose uptake during mesendoderm differentiation of embryonic stem cells. *Nature Communications*, *13*(1), 3984. https://doi.org/10.1038/s41467-022-31762-x

Fujii, K., Shi, Z., Zhulyn, O., Denans, N., & Barna, M. (2017). Pervasive translational regulation of the cell signalling circuitry underlies mammalian development. *Nature Communications*, *8*(1), 14443. https://doi.org/10.1038/ncomms14443

Gao, B., Wang, Y., Diaby, M., Zong, W., Shen, D., Wang, S., Chen, C., Wang, X., & Song, C. (2020). Evolution of pogo, a separate superfamily of IS630-Tc1-mariner transposons, revealing recurrent domestication events in vertebrates. *Mobile DNA*, *11*(1), 25. https://doi.org/10.1186/s13100-020-00220-0

Gezer, U., Özgür, E., Cetinkaya, M., Isin, M., & Dalay, N. (2014). Long non-coding RNAs with low expression levels in cells are enriched in secreted exosomes. *Cell Biology International*, *38*(9), 1076–1079. https://doi.org/10.1002/cbin.10301

Gilbert, SF. (2000). *Developmental Biology. : Vol. 6th edition.*

Gong, C., & Maquat, L. E. (2011). lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature*, *470*(7333), 284–288. https://doi.org/10.1038/nature09701

Goudarzi, M., Berg, K., Pieper, L. M., & Schier, A. F. (2019). Individual long non-coding RNAs have no overt functions in zebrafish embryogenesis, viability and fertility. *ELife*, *8*. https://doi.org/10.7554/eLife.40815

Grinman, E., Nakahata, Y., Avchalumov, Y., Espadas, I., Swarnkar, S., Yasuda, R., & Puthanveettil, S. V. (2021). Activity-regulated synaptic targeting of lncRNA ADEPTR mediates structural plasticity by localizing Sptn1 and AnkB in dendrites. *Science Advances*, *7*(16). https://doi.org/10.1126/sciadv.abf0605

Guan, H., Shang, G., Cui, Y., Liu, J., Sun, X., Cao, W., Wang, Y., & Li, Y. (2019). Long noncoding RNA APTR contributes to osteosarcoma progression through repression of miR-132-3p and upregulation of yes-associated protein 1. *Journal of Cellular Physiology*, *234*(6), 8998–9007. https://doi.org/10.1002/jcp.27572

Gui, Y., Liu, H., Zhang, L., Lv, W., & Hu, X. (2015). Altered microRNA profiles in cerebrospinal fluid exosome in Parkinson disease and Alzheimer disease. *Oncotarget*, *6*(35), 37043–37053. https://doi.org/10.18632/oncotarget.6158

Guillén, G., Díaz-Camino, C., Loyola-Torres, C. A., Aparicio-Fabre, R., Hernández-López, A., Díaz-Sánchez, M., & Sanchez, F. (2013). Detailed analysis of putative genes encoding small proteins in legume genomes. *Frontiers in Plant Science*, *4*. https://doi.org/10.3389/fpls.2013.00208

Guo, L., Zhang, W., Zhang, X., Wang, J., Nie, J., Jin, X., Ma, Y., Wang, S., Zhou, X., Zhang, Y., Xu, Y., Tanaka, Y., Yuan, J., Liao, X.-H., Gong, Y., & Su, L. (2023). A novel transcription factor SIPA1: identification and verification in triple-negative breast cancer. *Oncogene*, *42*(35), 2641–2654. https://doi.org/10.1038/s41388-023-02787-3

Haimes, J. D., & Kelley, M. (2015). *Demonstration of a ΔΔCq Calculation Method to Compute Relative Gene Expression from qPCR Data*.

Hangauer, M. J., Vaughn, I. W., & McManus, M. T. (2013). Pervasive Transcription of the Human Genome Produces Thousands of Previously Unidentified Long Intergenic Noncoding RNAs. *PLoS Genetics*, *9*(6), e1003569. https://doi.org/10.1371/journal.pgen.1003569

Hann, S. R., & Eisenman, R. N. (1984). Proteins Encoded by the Human c-myc Oncogene: Differential Expression in Neoplastic Cells. Molecular and Cellular Biology, 4(11), 2486–2497. https://doi.org/10.1128/mcb.4.11.2486-2497.1984

Hassel, K. R., Brito-Estrada, O., & Makarewich, C. A. (2023). Microproteins: Overlooked regulators of physiology and disease. *IScience*, *26*(6), 106781. https://doi.org/10.1016/j.isci.2023.106781

Hayano, Y., Sasaki, K., Ohmura, N., Takemoto, M., Maeda, Y., Yamashita, T., Hata, Y., Kitada, K., & Yamamoto, N. (2014). Netrin-4 regulates thalamocortical axon branching in an activity-dependent fashion. *Proceedings of the National Academy of Sciences*, *111*(42), 15226–15231. https://doi.org/10.1073/pnas.1402095111

Hazra, R., Utama, R., Naik, P., Dobin, A., & Spector, D. L. (2023). Identification of glioblastoma stem cell-associated lncRNAs using single-cell RNA sequencing datasets. *Stem Cell Reports*, *18*(11), 2056–2070. https://doi.org/10.1016/j.stemcr.2023.10.004

Herrera-Úbeda, C., Marín-Barba, M., Navas-Pérez, E., Gravemeyer, J., Albuixech-Crespo, B., Wheeler, G. N., & Garcia-Fernàndez, J. (2019). Microsyntenic Clusters Reveal Conservation of lncRNAs in Chordates Despite Absence of Sequence Conservation. *Biology*, *8*(3), 61. https://doi.org/10.3390/biology8030061

Hershberg, R., & Petrov, D. A. (2009). General Rules for Optimal Codon Choice. *PLoS Genetics*, *5*(7), e1000556. https://doi.org/10.1371/journal.pgen.1000556

Hetz, C., & Saxena, S. (2017). ER stress and the unfolded protein response in neurodegeneration. *Nature Reviews Neurology*, *13*(8), 477–491. https://doi.org/10.1038/nrneurol.2017.99

Hofacker, I. L. (2003). Vienna RNA secondary structure server. *Nucleic Acids Research*, *31*(13), 3429–3431. https://doi.org/10.1093/nar/gkg599

Hofman, D. A., Ruiz-Orera, J., Yannuzzi, I., Murugesan, R., Brown, A., Clauser, K. R., Condurat, A. L., van Dinter, J. T., Engels, S. A. G., Goodale, A., van der Lugt, J., Abid, T., Wang, L., Zhou, K. N., Vogelzang, J., Ligon, K. L., Phoenix, T. N., Roth, J. A., Root, D. E., … Prensner, J. R. (2024). Translation of non-canonical open reading frames as a cancer cell survival mechanism in childhood medulloblastoma. *Molecular Cell*. https://doi.org/10.1016/j.molcel.2023.12.003

Hopkins BD, Fine B, Steinbach N, et al. A secreted PTEN phosphatase that enters cells to alter signaling and survival. Science. 2013;341(6144):399-402. doi:10.1126/science.1234907

Housden, B. E., Valvezan, A. J., Kelley, C., Sopko, R., Hu, Y., Roesel, C., Lin, S., Buckner, M., Tao, R., Yilmazel, B., Mohr, S. E., Manning, B. D., & Perrimon, N. (2015). Identification of potential drug targets for tuberous sclerosis complex by synthetic screens combining CRISPR-based knockouts with RNAi. *Science Signaling*, *8*(393). https://doi.org/10.1126/scisignal.aab3729

Huang, Y., Wang, J., Zhao, Y., Wang, H., Liu, T., Li, Y., Cui, T., Li, W., Feng, Y., Luo, J., Gong, J., Ning, L., Zhang, Y., Wang, D., & Zhang, Y. (2021). cncRNAdb: a manually curated resource of experimentally supported RNAs with both protein-coding and noncoding function. *Nucleic Acids Research*, *49*(D1), D65–D70. https://doi.org/10.1093/nar/gkaa791

Hutton, S. R., & Pevny, L. H. (2011). SOX2 expression levels distinguish between neural progenitor populations of the developing dorsal telencephalon. *Developmental Biology*, *352*(1), 40–47. https://doi.org/10.1016/j.ydbio.2011.01.015

Ingolia NT, Lareau LF, Weissman JS. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell. 2011;147(4):789-802. doi:10.1016/j.cell.2011.10.002

Isin, M., Uysaler, E., Gezer, Uäÿ., & Dalay, N. (2015). Exosomal lncRNA-p21 levels may help to distinguish prostate cancer from benign disease. *Frontiers in Genetics*, *6*. https://doi.org/10.3389/fgene.2015.00168

Ivanyi-Nagy, R., Ahmed, S. M., Peter, S., Ramani, P. D., Ong, P. F., Dreesen, O., & Dröge, P. (2018). The RNA interactome of human telomerase RNA reveals a coding-independent role for a histone mRNA in telomere homeostasis. *ELife*, *7*. https://doi.org/10.7554/eLife.40037

Iwata, R., Casimir, P., & Vanderhaeghen, P. (2020). Mitochondrial dynamics in postmitotic cells regulate neurogenesis. *Science*, *369*(6505), 858–862. https://doi.org/10.1126/science.aba9760

Jackson, R. J., Hellen, C. U. T., & Pestova, T. V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature Reviews Molecular Cell Biology*, *11*(2), 113–127. https://doi.org/10.1038/nrm2838

Jan, Y.-N., & Jan, L. Y. (2010). Branching out: mechanisms of dendritic arborization. *Nature Reviews Neuroscience*, *11*(5), 316–328. https://doi.org/10.1038/nrn2836

Jasin, M., & Rothstein, R. (2013). Repair of Strand Breaks by Homologous Recombination. *Cold Spring Harbor Perspectives in Biology*, *5*(11), a012740–a012740. https://doi.org/10.1101/cshperspect.a012740

Jiang, X., Li, Q., Zhang, S., Song, C., & Zheng, P. (2019). Long noncoding RNA GIHCG induces cancer progression and chemoresistance and indicates poor prognosis in colorectal cancer. *OncoTargets and Therapy*, *Volume 12*, 1059–1070. https://doi.org/10.2147/OTT.S192290

Joglekar, A., Hu, W., Zhang, B., Narykov, O., Diekhans, M., Marrocco, J., Balacco, J., Ndhlovu, L. C., Milner, T. A., Fedrigo, O., Jarvis, E. D., Sheynkman, G., Korkin, D., Ross, M. E., & Tilgner, H. U. (2024). Single-cell long-read sequencing-based mapping reveals specialized splicing patterns in developing and adult mouse and human brain. *Nature Neuroscience*. https://doi.org/10.1038/s41593-024-01616-4

Kapp, L. D., & Lorsch, J. R. (2004). The Molecular Mechanics of Eukaryotic Translation. *Annual Review of Biochemistry*, *73*(1), 657–704. https://doi.org/10.1146/annurev.biochem.73.030403.080419

Katsushima, K., Natsume, A., Ohka, F., Shinjo, K., Hatanaka, A., Ichimura, N., Sato, S., Takahashi, S., Kimura, H., Totoki, Y., Shibata, T., Naito, M., Kim, H. J., Miyata, K., Kataoka, K., & Kondo, Y. (2016). Targeting the Notch-regulated non-coding RNA TUG1 for glioma treatment. *Nature Communications*, *7*(1), 13616. https://doi.org/10.1038/ncomms13616

Kawaji, H., Kasukawa, T., Forrest, A., Carninci, P., & Hayashizaki, Y. (2017). The FANTOM5 collection, a data series underpinning mammalian transcriptome atlases in diverse cell types. *Scientific Data*, *4*(1), 170113. https://doi.org/10.1038/sdata.2017.113

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, *10*(6), 845–858. https://doi.org/10.1038/nprot.2015.053

Khacho, M., Clark, A., Svoboda, D. S., Azzi, J., MacLaurin, J. G., Meghaizel, C., Sesaki, H., Lagace, D. C., Germain, M., Harper, M.-E., Park, D. S., & Slack, R. S. (2016). Mitochondrial Dynamics Impacts Stem Cell Identity and Fate Decisions by Regulating a Nuclear Transcriptional Program. *Cell Stem Cell*, *19*(2), 232–247. https://doi.org/10.1016/j.stem.2016.04.015

Kondo, T., Hashimoto, Y., Kato, K., Inagaki, S., Hayashi, S., & Kageyama, Y. (2007). Small peptide regulators of actin-based cell morphogenesis encoded by a polycistronic mRNA. *Nature Cell Biology*, *9*(6), 660–665. https://doi.org/10.1038/ncb1595

Korecka, J. A., van Kesteren, R. E., Blaas, E., Spitzer, S. O., Kamstra, J. H., Smit, A. B., Swaab, D. F., Verhaagen, J., & Bossers, K. (2013). Phenotypic Characterization of Retinoic Acid Differentiated SH-SY5Y Cells by Transcriptional Profiling. *PLoS ONE*, *8*(5), e63862. https://doi.org/10.1371/journal.pone.0063862

Kovalevich, J., & Langford, D. (2013). *Considerations for the Use of SH-SY5Y Neuroblastoma Cells in Neurobiology* (pp. 9–21). https://doi.org/10.1007/978-1-62703-640-5_2

Kragness, S., Clark, Z., Mullin, A., Guidry, J., & Earls, L. R. (2022). An Rtn4/Nogo-A-interacting micropeptide modulates synaptic plasticity with age. *PLOS ONE*, *17*(6), e0269404. https://doi.org/10.1371/journal.pone.0269404

Krishna, A., Biryukov, M., Trefois, C., Antony, P. M., Hussong, R., Lin, J., Heinäniemi, M., Glusman, G., Köglsberger, S., Boyd, O., van den Berg, B. H., Linke, D., Huang, D., Wang, K., Hood, L., Tholey, A., Schneider, R., Galas, D. J., Balling, R., & May, P. (2014). Systems

genomics evaluation of the SH-SY5Y neuroblastoma cell line as a model for Parkinson's disease. *BMC Genomics*, *15*(1), 1154. https://doi.org/10.1186/1471-2164-15-1154

Lancaster, M. A., Renner, M., Martin, C.-A., Wenzel, D., Bicknell, L. S., Hurles, M. E., Homfray, T., Penninger, J. M., Jackson, A. P., & Knoblich, J. A. (2013). Cerebral organoids model human brain development and microcephaly. *Nature*, *501*(7467), 373–379. https://doi.org/10.1038/nature12517

Leading causes of death, UK - Office for National Statistics. (2021).

Lee S, Liu B, Lee S, Huang SX, Shen B, Qian SB. Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. Proc Natl Acad Sci U S A. 2012;109(37):E2424-E2432. doi:10.1073/pnas.1207846109

Lee, C., Zeng, J., Drew, B. G., Sallam, T., Martin-Montalvo, A., Wan, J., Kim, S.-J., Mehta, H., Hevener, A. L., de Cabo, R., & Cohen, P. (2015). The Mitochondrial-Derived Peptide MOTS-c Promotes Metabolic Homeostasis and Reduces Obesity and Insulin Resistance. *Cell Metabolism*, *21*(3), 443–454. https://doi.org/10.1016/j.cmet.2015.02.009

Lee, W. J., Shin, C. H., Ji, H., Jeong, S. D., Park, M.-S., Won, H.-H., Pandey, P. R., Tsitsipatis, D., Gorospe, M., & Kim, H. H. (2021). hnRNPK-regulated LINC00263 promotes malignant phenotypes through miR-147a/CAPN2. *Cell Death & Disease*, *12*(4), 290. https://doi.org/10.1038/s41419-021-03575-1

Lewandowski, J. P., Dumbović, G., Watson, A. R., Hwang, T., Jacobs-Palmer, E., Chang, N., Much, C., Turner, K. M., Kirby, C., Rubinstein, N. D., Groff, A. F., Liapis, S. C., Gerhardinger, C., Bester, A., Pandolfi, P. P., Clohessy, J. G., Hoekstra, H. E., Sauvageau, M., & Rinn, J. L. (2020). The Tug1 lncRNA locus is essential for male fertility. *Genome Biology*, *21*(1), 237. https://doi.org/10.1186/s13059-020-02081-5

Li, C.-Y., Zhang, Y., Wang, Z., Zhang, Y., Cao, C., Zhang, P.-W., Lu, S.-J., Li, X.-M., Yu, Q., Zheng, X., Du, Q., Uhl, G. R., Liu, Q.-R., & Wei, L. (2010). A Human-Specific De Novo Protein-Coding Gene Associated with Human Brain Functions. *PLoS Computational Biology*, *6*(3), e1000734. https://doi.org/10.1371/journal.pcbi.1000734

Li, D.-X., Fei, X.-R., Dong, Y.-F., Cheng, C.-D., Yang, Y., Deng, X.-F., Huang, H.-L., Niu, W.-X., Zhou, C.-X., Xia, C.-Y., & Niu, C.-S. (2017). The long non-coding RNA CRNDE acts as a ceRNA and promotes glioma malignancy by preventing miR-136-5p-mediated downregulation of Bcl-2 and Wnt2. *Oncotarget*, *8*(50), 88163–88178. https://doi.org/10.18632/oncotarget.21513

Li, F., Wang, Y., Yang, H., Xu, Y., Zhou, X., Zhang, X., Xie, Z., & Bi, J. (2019). The effect of BACE1-AS on β-amyloid generation by regulating BACE1 mRNA expression. *BMC Molecular Biology*, *20*(1), 23. https://doi.org/10.1186/s12867-019-0140-0

Li, J., Wang, Y., Li, L., Or, P. M. -Y., Wai Wong, C., Liu, T., Ho, W. L. H., & Chan, A. M. (2021). Tumour-derived substrate-adherent cells promote neuroblastoma survival through secreted trophic factors. *Molecular Oncology*, *15*(8), 2011–2025. https://doi.org/10.1002/1878-0261.12969

Li, J., Zhang, M., An, G., & Ma, Q. (2016). LncRNA TUG1 acts as a tumor suppressor in human glioma by promoting cell apoptosis. *Experimental Biology and Medicine*, *241*(6), 644–649. https://doi.org/10.1177/1535370215622708

Li, J.-R., Sun, C.-H., Li, W., Chao, R.-F., Huang, C.-C., Zhou, X. J., & Liu, C.-C. (2016). Cancer RNA-Seq Nexus: a database of phenotype-specific transcriptome profiling in cancer cells. *Nucleic Acids Research*, *44*(D1), D944–D951. https://doi.org/10.1093/nar/gkv1282

Li, L., Zhuang, Y., Zhao, X., & Li, X. (2019). Long Non-coding RNA in Neuronal Development and Neurological Disorders. *Frontiers in Genetics*, *9*. https://doi.org/10.3389/fgene.2018.00744

Li, S., Jiang, X., Guan, M., Zhang, Y., Cao, Y., & Zhang, L. (2022). The overexpression of GPX8 is correlated with poor prognosis in GBM patients. *Frontiers in Genetics*, *13*. https://doi.org/10.3389/fgene.2022.898204

Li, X. L., Pongor, L., Tang, W., Das, S., Muys, B. R., Jones, M. F., Lazar, S. B., Dangelmaier, E. A., Hartford, C. C., Grammatikakis, I., Hao, Q., Sun, Q., Schetter, A., Martindale, J. L., Tang, B., Jenkins, L. M., Robles, A. I., Walker, R. L., Ambs, S., … Lal, A. (2020). A small protein encoded by a putative lncRNA regulates apoptosis and tumorigenicity in human colorectal cancer cells. *ELife*, *9*. https://doi.org/10.7554/eLife.53734

Lin, M., Pedrosa, E., Shah, A., Hrabovsky, A., Maqbool, S., Zheng, D., & Lachman, H. M. (2011). RNA-Seq of Human Neurons Derived from iPS Cells Reveals Candidate Long Non-Coding RNAs Involved in Neurogenesis and Neuropsychiatric Disorders. *PLoS ONE*, *6*(9), e23356. https://doi.org/10.1371/journal.pone.0023356

Lin, N., Chang, K.-Y., Li, Z., Gates, K., Rana, Z. A., Dang, J., Zhang, D., Han, T., Yang, C.-S., Cunningham, T. J., Head, S. R., Duester, G., Dong, P. D. S., & Rana, T. M. (2014). An Evolutionarily Conserved Long Noncoding RNA TUNA Controls Pluripotency and Neural Lineage Commitment. *Molecular Cell*, *53*(6), 1005–1019. https://doi.org/10.1016/j.molcel.2014.01.021

Liu, G., Jiang, Z., Qiao, M., & Wang, F. (2019). Lnc-GIHCG promotes cell proliferation and migration in gastric cancer through miR- 1281 adsorption. *Molecular Genetics & Genomic Medicine*, *7*(6). https://doi.org/10.1002/mgg3.711

Liu, H., Zong, C., Sun, J., Li, H., Qin, G., Wang, X., Zhu, J., Yang, Y., Xue, Q., & Liu, X. (2022). Bioinformatics analysis of lncRNAs in the occurrence and development of osteosarcoma. *Translational Pediatrics*, *11*(7), 1182–1198. https://doi.org/10.21037/tp-22-253

Liu, Q., Allen, T. D., Song, W., Wada, Y., Lobe, C. G., & Liu, J. (2019). Notch1 activates angiogenic regulator Netrin4 in endothelial cells. *Journal of Cellular and Molecular Medicine*, *23*(5), 3762–3766. https://doi.org/10.1111/jcmm.14240

Liu, S., Harmston, N., Glaser, T. L., Wong, Y., Zhong, Z., Madan, B., Virshup, D. M., & Petretto, E. (2020). Wnt-regulated lncRNA discovery enhanced by in vivo identification and

CRISPRi functional validation. *Genome Medicine*, *12*(1), 89. https://doi.org/10.1186/s13073-020-00788-5

Liu, S. J., Horlbeck, M. A., Cho, S. W., Birk, H. S., Malatesta, M., He, D., Attenello, F. J., Villalta, J. E., Cho, M. Y., Chen, Y., Mandegar, M. A., Olvera, M. P., Gilbert, L. A., Conklin, B. R., Chang, H. Y., Weissman, J. S., & Lim, D. A. (2017). CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science*, *355*(6320). https://doi.org/10.1126/science.aah7111

Liu, T., Zou, B., He, M., Hu, Y., Dou, Y., Cui, T., Tan, P., Li, S., Rao, S., Huang, Y., Liu, S., Cai, K., & Wang, D. (2023). LncReader: identification of dual functional long noncoding RNAs using a multi-head self-attention mechanism. *Briefings in Bioinformatics*, *24*(1). https://doi.org/10.1093/bib/bbac579

Liu, W., Seto, J., Donovan, G., & Toth, M. (2002). Jerky, a Protein Deficient in a Mouse Epilepsy Model, Is Associated with Translationally Inactive mRNA in Neurons. *The Journal of Neuroscience*, *22*(1), 176–182. https://doi.org/10.1523/JNEUROSCI.22-01-00176.2002

Liu, W., Seto, J., Sibille, E., & Toth, M. (2003). The RNA Binding Domain of Jerky Consists of Tandemly Arranged Helix-Turn-Helix/Homeodomain-Like Motifs and Binds Specific Sets of mRNAs. *Molecular and Cellular Biology*, *23*(12), 4083–4093. https://doi.org/10.1128/MCB.23.12.4083-4093.2003

Liu, X., Cheng, W., Li, H., & Song, Y. (2022). Identification and validation of cuproptosis-related LncRNA signatures as a novel prognostic model for head and neck squamous cell cancer. *Cancer Cell International*, *22*(1), 345. https://doi.org/10.1186/s12935-022-02762-0

Lloyd, J., French, C. E., & Brenner, S. E. (2020). Polysome fractionation analysis reveals features important for human nonsense-mediated mRNA decay. *BioRxiv*.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

MacLennan, D. H., & Kranias, E. G. (2003). Phospholamban: a crucial regulator of cardiac contractility. *Nature Reviews Molecular Cell Biology*, *4*(7), 566–577. https://doi.org/10.1038/nrm1151

Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., & Lopez, R. (2022). Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research*, *50*(W1), W276–W279. https://doi.org/10.1093/nar/gkac240

Maden, M. (2007). Retinoic acid in the development, regeneration and maintenance of the nervous system. *Nature Reviews Neuroscience*, *8*(10), 755–765. https://doi.org/10.1038/nrn2212

Mahato, A. K., & Sidorova, Y. A. (2020). RET Receptor Tyrosine Kinase: Role in Neurodegeneration, Obesity, and Cancer. *International Journal of Molecular Sciences*, *21*(19), 7108. https://doi.org/10.3390/ijms21197108

Makarewich, C. A. (2020). The hidden world of membrane microproteins. *Experimental Cell Research*, *388*(2), 111853. https://doi.org/10.1016/j.yexcr.2020.111853

Makarewich, C. A., & Olson, E. N. (2017). Mining for Micropeptides. *Trends in Cell Biology*, *27*(9), 685–696. https://doi.org/10.1016/j.tcb.2017.04.006

Malabat, C., Feuerbach, F., Ma, L., Saveanu, C., & Jacquier, A. (2015). Quality control of transcription start site selection by nonsense-mediated-mRNA decay. *ELife*, *4*. https://doi.org/10.7554/eLife.06722

Martin, F. J., Amode, M. R., Aneja, A., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Becker, A., Bennett, R., Berry, A., Bhai, J., Bhurji, S. K., Bignell, A., Boddu, S., Branco Lins, P. R., Brooks, L., Ramaraju, S. B., Charkhchi, M., Cockburn, A., Da Rin Fiorretto, L., … Flicek, P. (2023). Ensembl 2023. *Nucleic Acids Research*, *51*(D1), D933–D941. https://doi.org/10.1093/nar/gkac958

Martínez-Cerdeño, V., & Noctor, S. C. (2018). Neural Progenitor Cell Terminology. *Frontiers in Neuroanatomy*, *12*. https://doi.org/10.3389/fnana.2018.00104

Mas-Ponte, D., Carlevaro-Fita, J., Palumbo, E., Hermoso Pulido, T., Guigo, R., & Johnson, R. (2017). LncATLAS database for subcellular localization of long noncoding RNAs. *RNA*, *23*(7), 1080–1087. https://doi.org/10.1261/rna.060814.117

McDonald, E. F., Jones, T., Plate, L., Meiler, J., & Gulsevin, A. (2023). Benchmarking AlphaFold2 on peptide structure prediction. *Structure*, *31*(1), 111-119.e2. https://doi.org/10.1016/j.str.2022.11.012

Mercurio, S., Serra, L., & Nicolis, S. K. (2019). More than just Stem Cells: Functional Roles of the Transcription Factor Sox2 in Differentiated Glia and Neurons. *International Journal of Molecular Sciences*, *20*(18), 4540. https://doi.org/10.3390/ijms20184540

Miao, W., Li, N., Gu, B., Yi, G., Su, Z., & Cheng, H. (2020). LncRNA DLGAP1-AS2 modulates glioma development by up-regulating YAP1 expression. *The Journal of Biochemistry*, *167*(4), 411–418. https://doi.org/10.1093/jb/mvz108

Mikheeva, S. A., Mikheev, A. M., Petit, A., Beyer, R., Oxford, R. G., Khorasani, L., Maxwell, J.-P., Glackin, C. A., Wakimoto, H., González-Herrero, I., Sánchez-García, I., Silber, J. R., Horner, P. J., & Rostomily, R. C. (2010). TWIST1 promotes invasion through mesenchymal change in human glioblastoma. *Molecular Cancer*, *9*(1), 194. https://doi.org/10.1186/1476-4598-9-194

Miller, B., Kim, S.-J., Mehta, H. H., Cao, K., Kumagai, H., Thumaty, N., Leelaprachakul, N., Braniff, R. G., Jiao, H., Vaughan, J., Diedrich, J., Saghatelian, A., Arpawong, T. E., Crimmins, E. M., Ertekin-Taner, N., Tubi, M. A., Hare, E. T., Braskie, M. N., Décarie-Spain, L., … Cohen, P. (2023). Mitochondrial DNA variation in Alzheimer's disease

reveals a unique microprotein called SHMOOSE. *Molecular Psychiatry*, *28*(4), 1813–1826. https://doi.org/10.1038/s41380-022-01769-3

Mills, J. D., Kavanagh, T., Kim, W. S., Chen, B. J., Waters, P. D., Halliday, G. M., & Janitz, M. (2015). High expression of long intervening non-coding RNA OLMALINC in the human cortical white matter is associated with regulation of oligodendrocyte maturation. *Molecular Brain*, *8*(1), 2. https://doi.org/10.1186/s13041-014-0091-9

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nature Methods*, *19*(6), 679–682. https://doi.org/10.1038/s41592-022-01488-1

Mudge, J. M., Ruiz-Orera, J., Prensner, J. R., Brunet, M. A., Calvet, F., Jungreis, I., Gonzalez, J. M., Magrane, M., Martinez, T. F., Schulz, J. F., Yang, Y. T., Albà, M. M., Aspden, J. L., Baranov, P. V., Bazzini, A. A., Bruford, E., Martin, M. J., Calviello, L., Carvunis, A.-R., … van Heesch, S. (2022). Standardized annotation of translated open reading frames. *Nature Biotechnology*, *40*(7), 994–999. https://doi.org/10.1038/s41587-022-01369-0

Muñoz-Sanjuán, I., & Brivanlou, A. H. (2002). Neural induction, the default model and embryonic stem cells. *Nature Reviews Neuroscience*, *3*(4), 271–280. https://doi.org/10.1038/nrn786

Muralidharan, V., Tretiakova, A., Steplewski, A., Haas, S., Amini, S., Johnson, E., & Khalili, K. (1997). Evidence for inhibition of MyEF-2 binding to MBP promoter by MEF-1/Pur α. *Journal of Cellular Biochemistry*, *66*(4), 524–531. https://doi.org/10.1002/(SICI)1097-4644(19970915)66:4<524::AID-JCB11>3.0.CO;2-B

Mus, E., Hof, P. R., & Tiedge, H. (2007). Dendritic BC200 RNA in aging and in Alzheimer's disease. *Proceedings of the National Academy of Sciences*, *104*(25), 10679–10684. https://doi.org/10.1073/pnas.0701532104

Nagpal, H., Ali-Ahmad, A., Hirano, Y., Cai, W., Halic, M., Fukagawa, T., Sekulić, N., & Fierz, B. (2023). CENP-A and CENP-B collaborate to create an open centromeric chromatin state. *Nature Communications*, *14*(1), 8227. https://doi.org/10.1038/s41467-023-43739-5

Necsulea, A., Soumillon, M., Warnefors, M., Liechti, A., Daish, T., Zeller, U., Baker, J. C., Grützner, F., & Kaessmann, H. (2014). The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature*, *505*(7485), 635–640. https://doi.org/10.1038/nature12943

Ng, S.-Y., Bogu, G. K., Soh, B. S., & Stanton, L. W. (2013). The Long Noncoding RNA RMST Interacts with SOX2 to Regulate Neurogenesis. *Molecular Cell*, *51*(3), 349–359. https://doi.org/10.1016/j.molcel.2013.07.017

Nichols, M. L., Allen, B. J., Rogers, S. D., Ghilardi, J. R., Honore, P., Luger, N. M., Finke, M. P., Li, J., Lappi, D. A., Simone, D. A., & Mantyh, P. W. (1999). Transmission of Chronic Nociception by Spinal Neurons Expressing the Substance P Receptor. *Science*, *286*(5444), 1558–1561. https://doi.org/10.1126/science.286.5444.1558

Ntini, E., Louloupi, A., Liz, J., Muino, J. M., Marsico, A., & Ørom, U. A. V. (2018). Long ncRNA A-ROD activates its target gene DKK1 at its release from chromatin. *Nature Communications*, *9*(1), 1636. https://doi.org/10.1038/s41467-018-04100-3

O'Leary, N. A., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., … Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*, *44*(D1), D733–D745. https://doi.org/10.1093/nar/gkv1189

Orr, M. W., Mao, Y., Storz, G., & Qian, S.-B. (2020). Alternative ORFs and small ORFs: shedding light on the dark proteome. *Nucleic Acids Research*, *48*(3), 1029–1042. https://doi.org/10.1093/nar/gkz734

Pandey, M., Karmakar, V., Majie, A., Dwivedi, M., Md, S., & Gorain, B. (2024). The SH-SY5Y cell line: a valuable tool for Parkinson's disease drug discovery. *Expert Opinion on Drug Discovery*, *19*(3), 303–316. https://doi.org/10.1080/17460441.2023.2293158

Pant, M., Bal, Naresh. C., & Periasamy, M. (2016). Sarcolipin: A Key Thermogenic and Metabolic Regulator in Skeletal Muscle. *Trends in Endocrinology & Metabolism*, *27*(12), 881–892. https://doi.org/10.1016/j.tem.2016.08.006

Papadopoulos, C., & Albà, M. M. (2023). Newly evolved genes in the human lineage are functional. *Trends in Genetics*, *39*(4), 235–236. https://doi.org/10.1016/j.tig.2023.02.001

Papatheodorou, I., Moreno, P., Manning, J., Fuentes, A. M.-P., George, N., Fexova, S., Fonseca, N. A., Füllgrabe, A., Green, M., Huang, N., Huerta, L., Iqbal, H., Jianu, M., Mohammed, S., Zhao, L., Jarnuczak, A. F., Jupp, S., Marioni, J., Meyer, K., … Brazma, A. (2019). Expression Atlas update: from tissues to single cells. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkz947

Paşca, A. M., Sloan, S. A., Clarke, L. E., Tian, Y., Makinson, C. D., Huber, N., Kim, C. H., Park, J.-Y., O'Rourke, N. A., Nguyen, K. D., Smith, S. J., Huguenard, J. R., Geschwind, D. H., Barres, B. A., & Paşca, S. P. (2015). Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture. *Nature Methods*, *12*(7), 671–678. https://doi.org/10.1038/nmeth.3415

Patraquim, P., Magny, E. G., Pueyo, J. I., Platero, A. I., & Couso, J. P. (2022). Translation and natural selection of micropeptides from long non-canonical RNAs. *Nature Communications*, *13*(1), 6515. https://doi.org/10.1038/s41467-022-34094-y

Peng, J., & Zhao, L. (2024). The origin and structural evolution of de novo genes in Drosophila. *Nature Communications*, *15*(1), 810. https://doi.org/10.1038/s41467-024-45028-1

Pereira, I. T., Spangenberg, L., Cabrera, G., & Dallagiovanna, B. (2020). Polysome-associated lncRNAs during cardiomyogenesis of hESCs. *Molecular and Cellular Biochemistry*, *468*(1–2), 35–45. https://doi.org/10.1007/s11010-020-03709-7

Perrier, A. L., Tabar, V., Barberi, T., Rubio, M. E., Bruses, J., Topf, N., Harrison, N. L., & Studer, L. (2004). Derivation of midbrain dopamine neurons from human embryonic stem cells. *Proceedings of the National Academy of Sciences*, *101*(34), 12543–12548. https://doi.org/10.1073/pnas.0404700101

Perry, R. B.-T., Hezroni, H., Goldrich, M. J., & Ulitsky, I. (2018). Regulation of Neuroregeneration by Long Noncoding RNAs. *Molecular Cell*, *72*(3), 553-567.e5. https://doi.org/10.1016/j.molcel.2018.09.021

Polleux, F., Ince-Dunn, G., & Ghosh, A. (2007). Transcriptional regulation of vertebrate axon guidance and synapse formation. *Nature Reviews Neuroscience*, *8*(5), 331–340. https://doi.org/10.1038/nrn2118

Prensner, J. R., Abelin, J. G., Kok, L. W., Clauser, K. R., Mudge, J. M., Ruiz-Orera, J., Bassani-Sternberg, M., Moritz, R. L., Deutsch, E. W., & van Heesch, S. (2023). What Can Ribo-Seq, Immunopeptidomics, and Proteomics Tell Us About the Noncanonical Proteome? *Molecular & Cellular Proteomics*, *22*(9), 100631. https://doi.org/10.1016/j.mcpro.2023.100631

Prensner, J. R., Enache, O. M., Luria, V., Krug, K., Clauser, K. R., Dempster, J. M., Karger, A., Wang, L., Stumbraite, K., Wang, V. M., Botta, G., Lyons, N. J., Goodale, A., Kalani, Z., Fritchman, B., Brown, A., Alan, D., Green, T., Yang, X., … Golub, T. R. (2021). Noncanonical open reading frames encode functional proteins essential for cancer cell survival. *Nature Biotechnology*, *39*(6), 697–704. https://doi.org/10.1038/s41587-020-00806-2

Qian, C., Ye, Y., Mao, H., Yao, L., Sun, X., Wang, B., Zhang, H., Xie, L., Zhang, H., Zhang, Y., Zhang, S., & He, X. (2019). Downregulated lncRNA-SNHG1 enhances autophagy and prevents cell death through the miR-221/222 /p27/mTOR pathway in Parkinson's disease. *Experimental Cell Research*, *384*(1), 111614. https://doi.org/10.1016/j.yexcr.2019.111614

Ramilowski, J. A., Yip, C. W., Agrawal, S., Chang, J.-C., Ciani, Y., Kulakovskiy, I. V., Mendez, M., Ooi, J. L. C., Ouyang, J. F., Parkinson, N., Petri, A., Roos, L., Severin, J., Yasuzawa, K., Abugessaisa, I., Akalin, A., Antonov, I. V., Arner, E., Bonetti, A., … Carninci, P. (2020). Functional annotation of human long noncoding RNAs via molecular phenotyping. *Genome Research*, *30*(7), 1060–1072. https://doi.org/10.1101/gr.254219.119

Rani, N., Nowakowski, T. J., Zhou, H., Godshalk, S. E., Lisi, V., Kriegstein, A. R., & Kosik, K. S. (2016). A Primate lncRNA Mediates Notch Signaling during Neuronal Development by Sequestering miRNA. *Neuron*, *90*(6), 1174–1188. https://doi.org/10.1016/j.neuron.2016.05.005

Reimand, J., Kull, M., Peterson, H., Hansen, J., & Vilo, J. (2007). g:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Research*, *35*(suppl_2), W193–W200. https://doi.org/10.1093/nar/gkm226

Ren, A., Li, Q., Guo, Y., Cui, X., Wang, L., Huo, Y., Chen, H., Liu, H., & Huang, H. (2022). Low expression of lncRNA APTR promotes gastric cancer progression. *Translational Oncology*, *25*, 101506. https://doi.org/10.1016/j.tranon.2022.101506

Rosso, S. B., & Inestrosa, N. C. (2013). WNT signaling in neuronal maturation and synaptogenesis. *Frontiers in Cellular Neuroscience*, *7*. https://doi.org/10.3389/fncel.2013.00103

Ruiz-Orera, J., Verdaguer-Grau, P., Villanueva-Cañas, J. L., Messeguer, X., & Albà, M. M. (2018). Translation of neutrally evolving peptides provides a basis for de novo gene evolution. *Nature Ecology & Evolution*, *2*(5), 890–896. https://doi.org/10.1038/s41559-018-0506-6

Şahin, M., Öncü, G., Yılmaz, M. A., Özkan, D., & Saybaşılı, H. (2021). Transformation of SH-SY5Y cell line into neuron-like cells: Investigation of electrophysiological and biomechanical changes. *Neuroscience Letters*, *745*, 135628. https://doi.org/10.1016/j.neulet.2021.135628

Sandmann, C.-L., Schulz, J. F., Ruiz-Orera, J., Kirchner, M., Ziehm, M., Adami, E., Marczenke, M., Christ, A., Liebe, N., Greiner, J., Schoenenberger, A., Muecke, M. B., Liang, N., Moritz, R. L., Sun, Z., Deutsch, E. W., Gotthardt, M., Mudge, J. M., Prensner, J. R., … Hubner, N. (2023a). Evolutionary origins and interactomes of human, young microproteins and small peptides translated from short open reading frames. *Molecular Cell*, *83*(6), 994-1011.e18. https://doi.org/10.1016/j.molcel.2023.01.023

Sandmann, C.-L., Schulz, J. F., Ruiz-Orera, J., Kirchner, M., Ziehm, M., Adami, E., Marczenke, M., Christ, A., Liebe, N., Greiner, J., Schoenenberger, A., Muecke, M. B., Liang, N., Moritz, R. L., Sun, Z., Deutsch, E. W., Gotthardt, M., Mudge, J. M., Prensner, J. R., … Hubner, N. (2023b). Evolutionary origins and interactomes of human, young microproteins and small peptides translated from short open reading frames. *Molecular Cell*, *83*(6), 994-1011.e18. https://doi.org/10.1016/j.molcel.2023.01.023

Santiago, C., & Bashaw, G. J. (2014). Transcription factors and effectors that regulate neuronal morphology. *Development*, *141*(24), 4667–4680. https://doi.org/10.1242/dev.110817

Sarropoulos, I., Marin, R., Cardoso-Moreira, M., & Kaessmann, H. (2019). Developmental dynamics of lncRNAs across mammalian organs and species. *Nature*, *571*(7766), 510–514. https://doi.org/10.1038/s41586-019-1341-x

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, *9*(7), 676–682. https://doi.org/10.1038/nmeth.2019

Schuller, A. P., & Green, R. (2018). Roadblocks and resolutions in eukaryotic translation. *Nature Reviews Molecular Cell Biology*, *19*(8), 526–541. https://doi.org/10.1038/s41580-018-0011-4

Sehnal, D., Bittrich, S., Deshpande, M., Svobodová, R., Berka, K., Bazgier, V., Velankar, S., Burley, S. K., Koča, J., & Rose, A. S. (2021). Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Research*, *49*(W1), W431–W437. https://doi.org/10.1093/nar/gkab314

Senís, E., Esgleas, M., Najas, S., Jiménez-Sábado, V., Bertani, C., Giménez-Alejandre, M., Escriche, A., Ruiz-Orera, J., Hergueta-Redondo, M., Jiménez, M., Giralt, A., Nuciforo, P., Albà, M. M., Peinado, H., Toro, D. del, Hove-Madsen, L., Götz, M., & Abad, M. (2021). TUNAR lncRNA Encodes a Microprotein that Regulates Neural Differentiation and Neurite Formation by Modulating Calcium Dynamics. *Frontiers in Cell and Developmental Biology*, *9*. https://doi.org/10.3389/fcell.2021.747667

Shi, C., Zhang, L., & Qin, C. (2017). Long non-coding RNAs in brain development, synaptic biology, and Alzheimer's disease. *Brain Research Bulletin*, *132*, 160–169. https://doi.org/10.1016/j.brainresbull.2017.03.010

Shibata, M., Pattabiraman, K., Lorente-Galdos, B., Andrijevic, D., Kim, S.-K., Kaur, N., Muchnik, S. K., Xing, X., Santpere, G., Sousa, A. M. M., & Sestan, N. (2021). Regulation of prefrontal patterning and connectivity by retinoic acid. *Nature*, *598*(7881), 483–488. https://doi.org/10.1038/s41586-021-03953-x

Simone, R., Javad, F., Emmett, W., Wilkins, O. G., Almeida, F. L., Barahona-Torres, N., Zareba-Paslawska, J., Ehteramyan, M., Zuccotti, P., Modelska, A., Siva, K., Virdi, G. S., Mitchell, J. S., Harley, J., Kay, V. A., Hondhamuni, G., Trabzuni, D., Ryten, M., Wray, S., … de Silva, R. (2021). MIR-NATs repress MAPT translation and aid proteostasis in neurodegeneration. *Nature*, *594*(7861), 117–123. https://doi.org/10.1038/s41586-021-03556-6

Sjöstedt, E., Zhong, W., Fagerberg, L., Karlsson, M., Mitsios, N., Adori, C., Oksvold, P., Edfors, F., Limiszewska, A., Hikmet, F., Huang, J., Du, Y., Lin, L., Dong, Z., Yang, L., Liu, X., Jiang, H., Xu, X., Wang, J., … Mulder, J. (2020). An atlas of the protein-coding genes in the human, pig, and mouse brain. *Science*, *367*(6482). https://doi.org/10.1126/science.aay5947

Slavoff, S. A., Mitchell, A. J., Schwaid, A. G., Cabili, M. N., Ma, J., Levin, J. Z., Karger, A. D., Budnik, B. A., Rinn, J. L., & Saghatelian, A. (2013). Peptidomic discovery of short open reading frame–encoded peptides in human cells. *Nature Chemical Biology*, *9*(1), 59–64. https://doi.org/10.1038/nchembio.1120

Sosińska, P., Mikuła-Pietrasik, J., & Książek, K. (2015). The double-edged sword of long non-coding RNA: The role of human brain-specific BC200 RNA in translational control, neurodegenerative diseases, and cancer. *Mutation Research/Reviews in Mutation Research*, *766*, 58–67. https://doi.org/10.1016/j.mrrev.2015.08.002

Sparber, P., Filatova, A., Khantemirova, M., & Skoblov, M. (2019). The role of long non-coding RNAs in the pathogenesis of hereditary diseases. *BMC Medical Genomics*, *12*(S2), 42. https://doi.org/10.1186/s12920-019-0487-6

Spencer, H. L., Sanders, R., Boulberdaa, M., Meloni, M., Cochrane, A., Spiroski, A.-M., Mountford, J., Emanueli, C., Caporali, A., Brittan, M., Rodor, J., & Baker, A. H. (2020). The LINC00961 transcript and its encoded micropeptide, small regulatory polypeptide of amino acid response, regulate endothelial cell function. *Cardiovascular Research*, *116*(12), 1981–1994. https://doi.org/10.1093/cvr/cvaa008

Statello, L., Guo, C.-J., Chen, L.-L., & Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. *Nature Reviews Molecular Cell Biology*, *22*(2), 96–118. https://doi.org/10.1038/s41580-020-00315-9

Stein, C. S., Jadiya, P., Zhang, X., McLendon, J. M., Abouassaly, G. M., Witmer, N. H., Anderson, E. J., Elrod, J. W., & Boudreau, R. L. (2018). Mitoregulin: A lncRNA-Encoded Microprotein that Supports Mitochondrial Supercomplexes and Respiratory Efficiency. *Cell Reports*, *23*(13), 3710-3720.e8. https://doi.org/10.1016/j.celrep.2018.06.002

Stiles, J., & Jernigan, T. L. (2010). The Basics of Brain Development. *Neuropsychology Review*, *20*(4), 327–348. https://doi.org/10.1007/s11065-010-9148-4

Stothard, P. (2000). The Sequence Manipulation Suite: JavaScript Programs for Analyzing and Formatting Protein and DNA Sequences. *BioTechniques*, *28*(6), 1102–1104. https://doi.org/10.2144/00286ir01

Strother, L., Miles, G. B., Holiday, A. R., Cheng, Y., & Doherty, G. H. (2021). Long-term culture of SH-SY5Y neuroblastoma cells in the absence of neurotrophins: A novel model of neuronal ageing. *Journal of Neuroscience Methods*, *362*, 109301. https://doi.org/10.1016/j.jneumeth.2021.109301

Szafron, L. M., Balcerak, A., Grzybowska, E. A., Pienkowska-Grela, B., Felisiak-Golabek, A., Podgorska, A., Kulesza, M., Nowak, N., Pomorski, P., Wysocki, J., Rubel, T., Dansonka-Mieszkowska, A., Konopka, B., Lukasik, M., & Kupryjanczyk, J. (2015). The Novel Gene CRNDE Encodes a Nuclear Peptide (CRNDEP) Which Is Overexpressed in Highly Proliferating Tissues. *PLOS ONE*, *10*(5), e0127475. https://doi.org/10.1371/journal.pone.0127475

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N. T., Pyysalo, S., Bork, P., Jensen, L. J., & von Mering, C. (2023). The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research*, *51*(D1), D638–D646. https://doi.org/10.1093/nar/gkac1000

Taft, R. J., Pheasant, M., & Mattick, J. S. (2007). The relationship between non-protein-coding DNA and eukaryotic complexity. *BioEssays*, *29*(3), 288–299. https://doi.org/10.1002/bies.20544

Takeichi, M., Nimura, K., Mori, M., Nakagami, H., & Kaneda, Y. (2013). The Transcription Factors Tbx18 and Wt1 Control the Epicardial Epithelial-Mesenchymal Transition through Bi-Directional Regulation of Slug in Murine Primary Epicardial Cells. *PLoS ONE*, *8*(2), e57829. https://doi.org/10.1371/journal.pone.0057829

Tamariz, E., & Varela-EchavarrÃa, A. (2015). The discovery of the growth cone and its influence on the study of axon guidance. *Frontiers in Neuroanatomy*, *9*. https://doi.org/10.3389/fnana.2015.00051

Tao, Y., & Zhang, S.-C. (2016). Neural Subtype Specification from Human Pluripotent Stem Cells. *Cell Stem Cell*, *19*(5), 573–586. https://doi.org/10.1016/j.stem.2016.10.015

Tatemoto, K., Carlquist, M., & Mutt, V. (1982). Neuropeptide Y—a novel brain peptide with structural similarities to peptide YY and pancreatic polypeptide. *Nature*, *296*(5858), 659–660. https://doi.org/10.1038/296659a0

Tharakan, R., & Sawa, A. (2021). Minireview: Novel Micropeptide Discovery by Proteomics and Deep Sequencing Methods. *Frontiers in Genetics*, *12*. https://doi.org/10.3389/fgene.2021.651485

Therianos, S., Leuzinger, S., Hirth, F., Goodman, C. S., & Reichert, H. (1995). Embryonic development of the Drosophila brain: formation of commissural and descending pathways. *Development*, *121*(11), 3849–3860. https://doi.org/10.1242/dev.121.11.3849

Toraih, E. A., El-Wazir, A., Hussein, M. H., Khashana, M. S., Matter, A., Fawzy, M. S., & Hosny, S. (2019). Expression of long intergenic non-coding RNA, regulator of reprogramming, and its prognostic value in patients with glioblastoma. *The International Journal of Biological Markers*, *34*(1), 69–79. https://doi.org/10.1177/1724600818814459

Toth, M., Grimsby, J., Buzsaki, G., & Donovan, G. P. (1995). Epileptic seizures caused by inactivation of a novel gene, jerky, related to centromere binding protein–B in transgenic mice. *Nature Genetics*, *11*(1), 71–75. https://doi.org/10.1038/ng0995-71

Tsagakis, I., Douka, K., Birds, I., & Aspden, J. L. (2020). Long non-coding RNAs in development and disease: conservation to mechanisms. *The Journal of Pathology*, *250*(5), 480–495. https://doi.org/10.1002/path.5405

Turnbull, C., Scott, R. H., Thomas, E., Jones, L., Murugaesu, N., Pretty, F. B., Halai, D., Baple, E., Craig, C., Hamblin, A., Henderson, S., Patch, C., O'Neill, A., Devereau, A., Smith, K., Martin, A. R., Sosinsky, A., McDonagh, E. M., Sultana, R., … Caulfield, M. J. (2018). The 100 000 Genomes Project: bringing whole genome sequencing to the NHS. *BMJ*, k1687. https://doi.org/10.1136/bmj.k1687

Tzani I, Ivanov IP, Andreev DE, et al. Systematic analysis of the PTEN 5' leader identifies a major AUU initiated proteoform. Open Biol. 2016;6(5):150203. doi:10.1098/rsob.150203

Ulitsky, I., Shkumatava, A., Jan, C. H., Sive, H., & Bartel, D. P. (2011). Conserved Function of lincRNAs in Vertebrate Embryonic Development despite Rapid Sequence Evolution. *Cell*, *147*(7), 1537–1550. https://doi.org/10.1016/j.cell.2011.11.055

Valente, A. J., Maddalena, L. A., Robb, E. L., Moradi, F., & Stuart, J. A. (2017). A simple ImageJ macro tool for analyzing mitochondrial network morphology in mammalian cell culture. *Acta Histochemica*, *119*(3), 315–326. https://doi.org/10.1016/j.acthis.2017.03.001

van Heesch, S., Witte, F., Schneider-Lunitz, V., Schulz, J. F., Adami, E., Faber, A. B., Kirchner, M., Maatz, H., Blachut, S., Sandmann, C.-L., Kanda, M., Worth, C. L., Schafer, S., Calviello, L., Merriott, R., Patone, G., Hummel, O., Wyler, E., Obermayer, B., … Hubner, N. (2019). The Translational Landscape of the Human Heart. *Cell*, *178*(1), 242-260.e29. https://doi.org/10.1016/j.cell.2019.05.010

Vangoor, V. R., Gomes-Duarte, A., & Pasterkamp, R. J. (2021). Long non-coding RNAs in motor neuron development and disease. *Journal of Neurochemistry*, *156*(6), 777–801. https://doi.org/10.1111/jnc.15198

Velasco, S., Kedaigle, A. J., Simmons, S. K., Nash, A., Rocha, M., Quadrato, G., Paulsen, B., Nguyen, L., Adiconis, X., Regev, A., Levin, J. Z., & Arlotta, P. (2019). Individual brain organoids reproducibly form cell diversity of the human cerebral cortex. *Nature*, *570*(7762), 523–527. https://doi.org/10.1038/s41586-019-1289-x

Villanueva, A. A., Puvogel, S., Lois, P., Muñoz-Palma, E., Ramírez Orellana, M., Lubieniecki, F., Casco Claro, F., Gallegos, I., García-Castro, J., Sanchez-Gomez, P., Torres, V. A., & Palma, V. (2019). The Netrin-4/Laminin γ1/Neogenin-1 complex mediates migration in SK-N-SH neuroblastoma cells. *Cell Adhesion & Migration*, *13*(1), 33–40. https://doi.org/10.1080/19336918.2018.1506652

Volders, P.-J., Anckaert, J., Verheggen, K., Nuytens, J., Martens, L., Mestdagh, P., & Vandesompele, J. (2019). LNCipedia 5: towards a reference set of human long non-coding RNAs. *Nucleic Acids Research*, *47*(D1), D135–D139. https://doi.org/10.1093/nar/gky1031

Wang, T., Cao, L., Dong, X., Wu, F., De, W., Huang, L., & Wan, Q. (2020). LINC01116 promotes tumor proliferation and neutrophil recruitment via DDX5-mediated regulation of IL-1β in glioma cell. *Cell Death & Disease*, *11*(5), 302. https://doi.org/10.1038/s41419-020-2506-0

Wang, W., Zhao, F., Ma, X., Perry, G., & Zhu, X. (2020). Mitochondria dysfunction in the pathogenesis of Alzheimer's disease: recent advances. *Molecular Neurodegeneration*, *15*(1), 30. https://doi.org/10.1186/s13024-020-00376-6

Wang, Y., Xu, Z., Jiang, J., Xu, C., Kang, J., Xiao, L., Wu, M., Xiong, J., Guo, X., & Liu, H. (2013). Endogenous miRNA Sponge lincRNA-RoR Regulates Oct4, Nanog, and Sox2 in Human Embryonic Stem Cell Self-Renewal. *Developmental Cell*, *25*(1), 69–80. https://doi.org/10.1016/j.devcel.2013.03.002

Weijts, B. G. M. W., Bakker, W. J., Cornelissen, P. W. A., Liang, K.-H., Schaftenaar, F. H., Westendorp, B., de Wolf, C. A. C. M. T., Paciejewska, M., Scheele, C. L. G. J., Kent, L., Leone, G., Schulte-Merker, S., & de Bruin, A. (2012). E2F7 and E2F8 promote angiogenesis through transcriptional activation of VEGFA in cooperation with HIF1. *The EMBO Journal*, *31*(19), 3871–3884. https://doi.org/10.1038/emboj.2012.231

Wilson, B. A., Foy, S. G., Neme, R., & Masel, J. (2017). Young genes are highly disordered as predicted by the preadaptation hypothesis of de novo gene birth. *Nature Ecology & Evolution*, *1*(6), 0146. https://doi.org/10.1038/s41559-017-0146

Wright, D. J., Hall, N. A. L., Irish, N., Man, A. L., Glynn, W., Mould, A., Angeles, A. D. L., Angiolini, E., Swarbreck, D., Gharbi, K., Tunbridge, E. M., & Haerty, W. (2022). Long read sequencing reveals novel isoforms and insights into splicing regulation during cell state changes. *BMC Genomics*, *23*(1), 42. https://doi.org/10.1186/s12864-021-08261-2

Wu, J., Yao, J., Jia, S., Yao, X., Shao, J., Cao, W., Ma, S., Yao, X., & Li, H. (2023). A cuproptosis-related lncRNA signature for predicting prognosis and immune response in hepatocellular carcinoma. *Heliyon*, *9*(9), e19352. https://doi.org/10.1016/j.heliyon.2023.e19352

Wu, Y.-Y., & Kuo, H.-C. (2020). Functional roles and networks of non-coding RNAs in the pathogenesis of neurodegenerative diseases. *Journal of Biomedical Science*, *27*(1), 49. https://doi.org/10.1186/s12929-020-00636-z

Xie, Y., Lin, H., Wei, W., Kong, Y., Fang, Q., Chen, E., Liu, J., & Li, M. (2022). LINC00839 promotes malignancy of liver cancer via binding FMNL2 under hypoxia. *Scientific Reports*, *12*(1), 18757. https://doi.org/10.1038/s41598-022-16972-z

Yan, Y., Yang, D., Zarnowska, E. D., Du, Z., Werbel, B., Valliere, C., Pearce, R. A., Thomson, J. A., & Zhang, S. (2005). Directed Differentiation of Dopaminergic Neuronal Subtypes from Human Embryonic Stem Cells. *STEM CELLS*, *23*(6), 781–790. https://doi.org/10.1634/stemcells.2004-0365

Ye, J., Zhu, J., Chen, H., Qian, J., Zhang, L., Wan, Z., Chen, F., Sun, S., Li, W., & Luo, C. (2020a). A novel lncRNA-LINC01116 regulates tumorigenesis of glioma by targeting VEGFA. *International Journal of Cancer*, *146*(1), 248–261. https://doi.org/10.1002/ijc.32483

Ye, J., Zhu, J., Chen, H., Qian, J., Zhang, L., Wan, Z., Chen, F., Sun, S., Li, W., & Luo, C. (2020b). A novel lncRNA-LINC01116 regulates tumorigenesis of glioma by targeting VEGFA. *International Journal of Cancer*, *146*(1), 248–261. https://doi.org/10.1002/ijc.32483

Yeasmin, F., Yada, T., & Akimitsu, N. (2018). Micropeptides Encoded in Transcripts Previously Identified as Long Noncoding RNAs: A New Chapter in Transcriptomics and Proteomics. *Frontiers in Genetics*, *9*. https://doi.org/10.3389/fgene.2018.00144

Yen, Z. C., Meyer, I. M., Karalic, S., & Brown, C. J. (2007). A cross-species comparison of X-chromosome inactivation in Eutheria. *Genomics*, *90*(4), 453–463. https://doi.org/10.1016/j.ygeno.2007.07.002

Yin, J., Ding, F., Cheng, Z., Ge, X., Li, Y., Zeng, A., Zhang, J., Yan, W., Shi, Z., Qian, X., You, Y., Ding, Z., Ji, J., & Wang, X. (2023). METTL3-mediated m6A modification of LINC00839 maintains glioma stem cells and radiation resistance by activating Wnt/β-catenin signaling. *Cell Death & Disease*, *14*(7), 417. https://doi.org/10.1038/s41419-023-05933-7

Yoon, J.-H., Abdelmohsen, K., Srikantan, S., Yang, X., Martindale, J. L., De, S., Huarte, M., Zhan, M., Becker, K. G., & Gorospe, M. (2012). LincRNA-p21 Suppresses Target mRNA Translation. *Molecular Cell*, *47*(4), 648–655. https://doi.org/10.1016/j.molcel.2012.06.027

Zang, Y., Chaudhari, K., & Bashaw, G. J. (2021). *New insights into the molecular mechanisms of axon guidance receptor regulation and signaling* (pp. 147–196). https://doi.org/10.1016/bs.ctdb.2020.11.008

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

Zeng, C., & Hamada, M. (2018). Identifying sequence features that drive ribosomal association for lncRNA. *BMC Genomics*, *19*(S10), 906. https://doi.org/10.1186/s12864-018-5275-8

Zhang, C., Zhou, B., Gu, F., Liu, H., Wu, H., Yao, F., Zheng, H., Fu, H., Chong, W., Cai, S., Huang, M., Ma, X., Guo, Z., Li, T., Deng, W., Zheng, M., Ji, Q., Zhao, Y., Ma, Y., … Guo, C. (2022). Micropeptide PACMP inhibition elicits synthetic lethal effects by decreasing CtIP and poly(ADP-ribosyl)ation. *Molecular Cell*, *82*(7), 1297-1312.e8. https://doi.org/10.1016/j.molcel.2022.01.020

Zhang, N., Shen, H., Huang, S., Wang, F., Liu, H., Xie, F., Jiang, L., & Chen, X. (2021). LncRNA FGD5-AS1 functions as an oncogene to upregulate GTPBP4 expression by sponging miR-873-5p in hepatocellular carcinoma. *European Journal of Histochemistry*, *65*(4). https://doi.org/10.4081/ejh.2021.3300

Zhang, Q., Wei, J., Li, N., & Liu, B. (2022). LINC00839 Promotes Neuroblastoma Progression by Sponging miR-454-3p to Up-Regulate NEUROD1. *Neurochemical Research*, *47*(8), 2278–2293. https://doi.org/10.1007/s11064-022-03613-0

Zhang, R., Li, J., Badescu, D., Karaplis, A., Ragoussis, J., & Kremer, R. (2023). PTHrP Regulates Fatty Acid Metabolism via Novel lncRNA in Breast Cancer Initiation and Progression Models. *Cancers*, *15*(15), 3763. https://doi.org/10.3390/cancers15153763

Zhang, X., Mao, L., Li, L., He, Z., Wang, N., & Song, Y. (2019). Long noncoding RNA GIHCG functions as an oncogene and serves as a serum diagnostic biomarker for cervical cancer. *Journal of Cancer*, *10*(3), 672–681. https://doi.org/10.7150/jca.28525

Zhang, X., Wang, W., Zhu, W., Dong, J., Cheng, Y., Yin, Z., & Shen, F. (2019). Mechanisms and Functions of Long Non-Coding RNAs at Multiple Regulatory Levels. *International Journal of Molecular Sciences*, *20*(22), 5573. https://doi.org/10.3390/ijms20225573

Zhao, J. B., Xue, J. F., Zhang, W. Z., Ren, Y. L., & Yan, D. M. (2020). Long Noncoding RNA FGD5-AS1 Promotes Glioma Cell Proliferation, Migration and Invasion by Regulating wnt/β-Catenin Pathway. *Cancer Management and Research*, *Volume 12*, 6187–6193. https://doi.org/10.2147/CMAR.S250284

Zhao, M.-Y., Wang, G.-Q., Wang, N.-N., Yu, Q.-Y., Liu, R.-L., & Shi, W.-Q. (2019). The long-non-coding RNA NEAT1 is a novel target for Alzheimer's disease progression via miR-124/BACE1 axis. *Neurological Research*, *41*(6), 489–497. https://doi.org/10.1080/01616412.2018.1548747

Zhao, Y., Liu, H., Zhang, Q., & Zhang, Y. (2020). The functions of long non-coding RNAs in neural stem cell proliferation and differentiation. *Cell & Bioscience*, *10*(1), 74. https://doi.org/10.1186/s13578-020-00435-x

Zhao, Y., Liu, Y., Lin, L., Huang, Q., He, W., Zhang, S., Dong, S., Wen, Z., Rao, J., Liao, W., & Shi, M. (2018). The lncRNA MACC1-AS1 promotes gastric cancer cell metabolic plasticity via AMPK/Lin28 mediated mRNA stability of MACC1. *Molecular Cancer*, *17*(1), 69. https://doi.org/10.1186/s12943-018-0820-2

Zheng, X., Zheng, D., Zhang, C., Guo, H., Zhang, Y., Xue, X., Shi, Z., Zhang, X., Zeng, X., Wu, Y., & Gao, W. (2023). A cuproptosis-related lncRNA signature predicts the prognosis and immune cell status in head and neck squamous cell carcinoma. *Frontiers in Oncology*, *13*. https://doi.org/10.3389/fonc.2023.1055717

Zhu, S., Wang, J.-Z., Chen, D., He, Y.-T., Meng, N., Chen, M., Lu, R.-X., Chen, X.-H., Zhang, X.-L., & Yan, G.-R. (2020). An oncopeptide regulates m6A recognition by the m6A reader IGF2BP1 and tumorigenesis. *Nature Communications*, *11*(1), 1685. https://doi.org/10.1038/s41467-020-15403-9

Zhu, S.-Y., Zou, H.-C., Gao, M.-M., Chen, Y.-X., Xu, M., & Qin, X.-H. (2022). LncRNA GIHCG promoted the proliferation and migration of renal cell carcinoma through regulating miR-499a-5p/XIAP axis. *Translational Oncology*, *20*, 101356. https://doi.org/10.1016/j.tranon.2022.101356

Zuccarini, M., Giuliani, P., Ziberi, S., Carluccio, M., Iorio, P. Di, Caciagli, F., & Ciccarelli, R. (2018). The Role of Wnt Signal in Glioblastoma Development and Progression: A Possible New Pharmacological Target for the Therapy of This Tumor. *Genes*, *9*(2), 105. https://doi.org/10.3390/genes9020105