Transcriptional Mechanisms Involved in Long-term Potentiation of Hippocampal Neurons

Author:
Bliim, Nicola

Publication Date:
2017

DOI:
https://doi.org/10.26190/unsworks/3274

License:
https://creativecommons.org/licenses/by-nc-nd/3.0/au/
Link to license to see what you are allowed to do with this resource.

Downloaded from http://hdl.handle.net/1959.4/58628 in https://unsworks.unsw.edu.au on 2024-04-27

Transcriptional Mechanisms Involved in Long-term Potentiation of Hippocampal Neurons

by

Nicola Bliim

z3379852

Supervisor

Dr. Michael Janitz

A thesis submitted as partial fulfillment

of the requirement for the degree of

Master of Philosophy (BABS)

School of Biotechnology and Biomolecular Sciences

UNSW Sydney

2017

PLEASE TYPE

THE UNIVERSITY OF NEW SOUTH WALES
Thesis/Dissertation Sheet

Surname or Family name: **Bliim**

First name: **Nicola** Other name/s:

Abbreviation for degree as given in the University calendar: **MPhil**

School: **School of Biotechnology and Biomolecular Sciences** Faculty: **Science**

Title: **Transcriptional Mechanisms Involved in Long-term Potentiation of Hippocampal Neurons**

Abstract 350 words maximum: (PLEASE TYPE)

Long-term potentiation (LTP), the persistent strengthening of synaptic connections following high frequency stimulation, is a form of synaptic plasticity proposed to underlie memory formation and consolidation. Despite decades of extensive study our understanding of the molecular mechanisms underpinning LTP remains incomplete. Whilst many of the protein components involved in the mechanisms of LTP have been extensively described the long non-protein-coding RNAs (lncRNAs) remain largely unexplored. Expression of lncRNAs is particularly enriched in the mammalian brain where they potentially impact LTP through regulation of epigentetic processes, transcription, mRNA splicing and translation. Characterisation of all the proteins and lncRNAs involved in LTP may elucidate the molecular mechanisms underlying memory formation and consolidation as well as provide greater insight into perturbation of these processes in developmental and neurodegenerative diseases. This study aimed to comprehensively analyse the transcriptome of primary hippocampal neurons, from neonatal mice, undergoing LTP induction in order to identify and quantify the protein-coding genes and ncRNAs expressed during LTP induction. Furthermore, this project aimed to investigate the transcriptomic changes that result from inhibition of LTP through disruption of synaptic adhesion. Analysis of transcriptome sequencing data led to identification of 64 differentially expressed genes, including four unannotated noncoding lincRNAs, across four distinct LTP conditions. Among those genes four distinct expression patterns could be identified. Gene Ontology (GO) analysis identified numerous enriched GO terms including those associated with intracellular signalling, transcriptional and translational regulation, as well as numerous clusters associated with immune response. The novel unannotated transcripts identified in this study were characterised as putative long intervening non-coding RNAs (lincRNAs), two of which demonstrated potential micropeptide expression. Future studies will determine the role and function of these putative lincRNAs in the induction of LTP. Meta-analysis comparing the results of the present study with those of a recent study on LTP induction in rat hippocampal neurons found no common differentially expressed genes.

Declaration relating to disposition of project thesis/dissertation

I hereby grant to the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or in part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all property rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstracts International (this is applicable to doctoral theses only).

…… …… ……..……

Signature Witness Signature ……..……..

Date

The University recognises that there may be exceptional circumstances requiring restrictions on copying or conditions on use. Requests for restriction for a period of up to 2 years must be made in writing. Requests for a longer period of restriction may be considered in exceptional circumstances and require the approval of the Dean of Graduate Research.

FOR OFFICE USE ONLY Date of completion of requirements for Award:

Copyright statement

I hereby grant the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation. I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstract International (this is applicable to doctoral theses only). I have either used no substantial portions of copyright material in my thesis or I have obtained permission to use copyright material; where permission has not been granted I have applied/will apply for a partial restriction of the digital copy of my thesis or dissertation.

Signed ……

Dated……..

Authenticity statement

I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.

Signed ……

Dated……..

Originality statement

I, Nicola Bliim, hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution. Except where due acknowledgement is made in the thesis. Any contribution made to the research by others with whom I have worked with at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

Signed ……

Dated……..

Abstract

Long-term potentiation (LTP), the persistent strengthening of synaptic connections following high frequency stimulation, is a form of synaptic plasticity proposed to underlie memory formation and consolidation. Despite decades of extensive study our understanding of the molecular mechanisms underpinning LTP remains incomplete. Whilst many of the protein components involved in the mechanisms of LTP have been extensively described the long non-protein-coding RNAs (lncRNAs) remain largely unexplored. Expression of lncRNAs is particularly enriched in the mammalian brain where they potentially impact LTP through regulation of epigentetic processes, transcription, mRNA splicing and translation. Characterisation of all the proteins and lncRNAs involved in LTP may elucidate the molecular mechanisms underlying memory formation and consolidation as well as provide greater insight into perturbation of these processes in developmental and neurodegenerative diseases. This study aimed to comprehensively analyse the transcriptome of primary hippocampal neurons, from neonatal mice, undergoing LTP induction in order to identify and quantify the protein-coding genes and ncRNAs expressed during LTP induction. Furthermore, this project aimed to investigate the transcriptomic changes that result from inhibition of LTP through disruption of synaptic adhesion. Analysis of transcriptome sequencing data led to identification of 64 differentially expressed genes, including four unannotated noncoding lincRNAs, across four distinct LTP conditions. Among those genes four distinct expression patterns could be identified. Gene Ontology (GO) analysis identified numerous enriched GO terms including those associated with intracellular signalling, transcriptional and translational regulation, as well as numerous clusters associated with immune response. The novel unannotated transcripts identified in this study were characterised as putative long intervening non-coding RNAs (lincRNAs), two of which demonstrated potential micropeptide expression. Future studies will determine the role and function of these putative lincRNAs in the induction of LTP. Meta-analysis comparing the results of the present study with those of a recent study on LTP induction in rat hippocampal neurons found no common differentially expressed genes.

Acknowledgements

I would like to thank all those who assisted and supported me through the course of this project. It has been an incredibly challenging process and without your support none of this would have been possible.

Thank you to my supervisor, Dr Michael Janitz, for giving me the opportunity to undertake this project. Thank you for your patience, enthusiasm, insight and guidance every step of the way.

Thank you to my co-supervisor Dr Vladimir Sytnyk as well as Dr Iryna Leshchysn'ka for their expert knowledge and advice.

Thank you to the other members of the Janitz lab for creating friendly work environment, for offering encouragement and advice. My particular thanks to James Mills and Bei Jun Chen for sparing their time to teach me how to use the bioinformatics software and helping me with technical problems. Thank you to Konii Takenaka for helping create a friendly and supportive lab environment.

Finally, I would like to thank my friends and family who have put up with me being a stressed lunatic throughout the course of this project. Thank you to my parents for supporting me and making it possible for me to continue studying.

Table of Contents

ORIGINALITY STATEMENT…….2

ABSTRACT……...3

ACKNOWLEDGEMENTS……...4

TABLE OF CONTENTS……...5

LIST OF FIGURES……...7

LIST OF TABLES…….9

ABBREVIATIONS……..10

1. INTRODUCTION……...12
    1.1 Long-term potentiation…….12
    1.2 Non-coding RNAs……14
    1.3 Neural cell adhesion molecule 2 (NCAM2)…….……17
    1.4 RNA Sequencing……...19
    1.5 Aims……..21
2. MATERIALS AND METHODS……22
    2.1 Tissue culture, NCAM2-ED treatment and *in vitro* induction of LTP…….22
    2.2 RNA extraction and RNA-Seq…….22
    2.3 Differential expression testing……..23
    2.4 Gene ontology enrichment testing……24
    2.5 Characterisation of unannotated lncRNAs: Determination of protein-coding potential and RNA secondary structure……24
    2.6 Meta-analysis of rat RNA-Seq data……..……25
3. RESULTS……27
    3.1 General RNA-Seq metrics……27
    3.2 Expression pattern of significantly differentially expressed genes……..29
    3.3 Gene ontology analysis of significantly differentially expressed genes……...33
    3.4 Characterisation of lincRNA expression……..42
        3.4.1 Annotated lincRNAs: Neat1, Malat1 and Miat...……42
        3.4.2 XLOC_013906 transcript……43
        3.4.3 XLOC_006440 transcript……46

3.4.4 XLOC_006256 transcript……50
3.5 Meta-analysis rat hippocampal transcriptome data.……57
4. DISCUSSION……..58
4.1 Expression pattern of significantly differentially expressed protein-coding genes……58
4.1.1 Expression pattern 1……58
4.1.2 Expression pattern 2……59
4.1.3 Expression pattern 3……60
4.1.4 Expression pattern 4……61
4.1.5 Fos Expression……61
4.2 Gene ontology analysis of significantly differentially expressed genes……..62
4.3 Characterisation of lincRNA expression……..64
4.3.1 Annotated lincRNAs……64
4.3.2 Unannotated lincRNAs……68
4.4 Concordance between rat and mouse transcriptome data..……..69
4.5 Future directions……...71
REFERENCES……72
COPYRIGHT STATEMENT……..81
AUTHENTICITY STATEMENT……81

List of Figures

Figure 1. Gene Expression Pattern 1……...29

Figure 2. Gene Expression Pattern 2……...30

Figure 3. Gene Expression Pattern 3……...31

Figure 4. Gene Expression Pattern 4……...32

Figure 5. Fos Expression Pattern…….32

Figure 6. Gene ontology enrichment map of genes identified as differentially expressed in Control vs. LTP comparison…….34

Figure 7. Gene ontology enrichment map of genes identified as differentially expressed in Control vs. NCAM2-ED comparison……...36

Figure 8. Gene ontology enrichment map of genes identified as differentially expressed in LTP vs. NCAM2-ED comparison……38

Figure 9. Gene ontology enrichment map of genes identified as differentially expressed in NCAM2-ED vs. LTP+NCAM2-ED comparison…….40

Figure 10. Schematic of the MALAT1 and Neat1 loci……..42

Figure 11. Comparative expression patterns of the Neat1, Malat1 and Miat transcripts...43

Figure 12. Schematic representation of the XLOC_013906 locus……45

Figure 13. Expression patterns of XLOC_013906 transcripts……...46

Figure 14. Expression patterns of Kcna4 isoforms……46

Figure 15. Location of predicted ORFs within the XLOC_013906 sequence……...47

Figure 16. Ribosome binding profile of XLOC_013906 locus in mouse hippocampal tissue…….47

Figure 17. Schematic representation of the XLOC_006440 locus……48

Figure 18. Expression pattern of XLOC_006440 transcript……..48

Figure 19. Location of predicted ORFs within the XLOC_006440 sequence……...49

Figure 20. Predicted RNA secondary structure of XLOC_006440……...50

Figure 21. Ribosome binding profile of XLOC_006440 locus in mouse hippocampal tissue…….50

Figure 22. Schematic representation of the XLOC_006256 locus……51

Figure 23. Cumulative XLOC_006256 gene expression pattern and expression patterns of XLOC_006256 isoform 1 and isoform 2……..52

Figure 24. Location of predicted ORFs within the XLOC_006256 isoform 1 and 2……52

Figure 25. Predicted RNA secondary structure of XLOC_006256 isoform 1……...54

Figure 26. Predicted RNA secondary structure of XLOC_006256 isoform 2……...55

Figure 27. Ribosome binding profile of XLOC_006256 locus in mouse hippocampal tissue…….56

List of Tables

Table 1. Cell treatment groups……...22
Table 2. Number of genes and isoforms detected in each sample……..27
Table 3. Number of significantly differentially expressed genes for each pairwise analysis of treatment groups…….28
Table 4. Number of upregulated and downregulated significantly differentially expressed genes……28
Table 5. Top 10 GO terms enriched in the Control vs. LTP comparison……...35
Table 6. Top 10 GO terms enriched in the Control vs. NCAM2-ED comparison…….37
Table 7. Top 10 GO terms enriched in the LTP vs. LTP+NCAM2-ED comparison….39
Table 8. Top 10 GO terms enriched in the NCAM2-ED vs. LTP+NCAM2-ED comparison……41
Table 9. Number of unannotated transcripts detected with expression >1 FPKM in each treatment group…….44
Table 10. Characteristics of significantly differentially expressed unannotated transcripts……..44
Table 11. Predicted ORFs for XLOC_006440 sequence……..49
Table 12. Predicted ORFs for XLOC_006256 isoforms 1 and 2……..53
Table 13. Comparison of mouse and rat expression levels for eight genes that were significantly differentially expressed in the mouse Control vs. LTP comparison……57
Table S1. Top 100 expressed genes based on p-value
Table S2. Full list of significantly differentially expressed genes for each pairwise comparison
Table S3. Full non-redundant list of significantly differentially expressed genes
Table S4. List of GO enrichment clusters
Table S5. Full list of predicted ORFs for XLOC_013906 sequence
Table S6. Full list of significantly differentially expressed genes in rat hippocampus

Abbreviations

Aβ - amyloid-β protein

aa – amino acids

AD – Alzheimer's disease

bp – base pairs

CAMs – cell adhesion molecules

circRNA – circular RNA

DAVID – the database for annotation, visualization and integrated discovery

DE – differentially expressed

E-LTP – induction or early phase long-term potentiation

FPKM – fragments per kilobase of transcript per million mapped reads

GO – Gene ontology

Grp – Gastrin releasing peptide

GWIPS-viz - genome wide information on protein synthesis browser

IEG – immediate early genes

IGV – interactive genomics viewer

I-LTP – intermediate phase long-term potentiation

L-LTP – maintenance or late phase long-term potentiation

LRG – late response genes

lincRNA – long intervening non-coding RNA

lncRNA – long non-coding RNA

LTP – long-term potentiation

MFE - minimum free energy

miRNA - microRNAs

NCAM2 – neural cell adhesion molecule 2

NCAM2-ED – the extracellular domain of neural cell adhesion molecule 2

ncRNA – non-coding RNA

Neat1 - nuclear enriched abundant transcript 1

Ntsr2 – Neurotensin receptor subtype 2

nt – nucleotide

ORF – open reading frame

Rxrg – Retinoid X receptor gamma

Ribo-Seq – ribosome profiling

RNA-Seq – RNA sequencing

SR - serine/arginine-rich

ssRNA – single stranded RNA

UCSC – University of California, Santa Cruz

1. Introduction

The human brain is one of the most complex entities in the known universe, perhaps second only to the intricate societies and cultures to which it gives rise (Chaisson 2002). It is estimated that the average human brain has approximately 80-100 billion neurons, each neuron sharing synaptic connections with thousands of other neurons, culminating in networks of approximately 100 trillion synapses (Azevedo *et al*. 2009; Markram 2012; Noctor *et al*. 2007; Toga *et al*. 2012). These elaborate networks of interconnected neurons make up the brain structures that constitute our minds, personalities and memories. Thoughts are transmitted throughout these networks as electrochemical signals, moving from one neuron to the next via synaptic connections. These networks are not static, constantly changing and adapting in response to the environment, transforming intangible thoughts and experiences into enduring physical memory traces or engrams (Hebb 2005; Malenka and Nicoll 1999; Rubin *et al*. 2014). The processing power of the human brain remains unrivalled by anything known to humankind, conferring upon us as a species the ability to reason, to learn and to demystify the workings of the world around us. Yet the inner working of the brain itself remain elusive.

1.1. Long-term potentiation

Each time something new is learnt or a memory is recalled synaptic connections are formed and existing synapses are reinforced or weakened, altering the pathway of signals throughout neuronal networks. The term synaptic plasticity is used to describe the strengthening or weakening of synaptic connections between neurons, altering neuronal networks in response to neural activity (Abbott and Nelson 2000; Costa-Mattioli *et al*. 2009).

Long-term potentiation (LTP) refers to the persistent strengthening of synaptic connections following high levels of excitatory synaptic stimulation. First described in the mammalian hippocampus, a region of the brain associated with the encoding and recall of episodic memory, LTP is the most extensively studied mechanism of synaptic plasticity as it is proposed to be the basis of memory formation and retention (Bliss and Collingridge 1993; Bliss and Lømo 1973; English and Sweatt 1996; Fortin *et al*. 2002; Martin *et al*. 2000; Neves *et al*. 2008). Long-term potentiation has been observed in hippocampal neurons following learning events and inhibition of LTP hindered learning, providing compelling evidence for its role in memory formation (Gruart *et al*. 2006; Whitlock *et al*. 2006).

Alteration of synaptic connections during LTP requires coordinated remodelling of both the presynaptic and postsynaptic neurons to increase the connectivity of the synapse and intensifying the transmission of electrochemical signals between the two neurons (Kandel 2012). The intricate molecular mechanisms of LTP have been the subject of over four decades of neuroscientific research amassing an expansive body of literature and numerous extensive reviews (Baudry *et al*. 2015; Bliim *et al*. 2016; Herring and Nicoll 2016; McHail and Dumas 2015; Sweatt 2016). Induction of LTP leads to dynamic regulation of a range of molecular pathways including those involved in response to extracellular stimuli, intercellular communication, intracellular signal cascades, transcriptional regulation, neutrite growth and cytoskeletal organisation to affect enduring changes in synaptic connectivity (Kelleher *et al*. 2004; Park *et al*. 2006).

High levels of neuronal stimulation activate complex molecular signalling cascades in the presynaptic and postsynaptic neurons, prompting the modification of molecules present at the synapse as well as alteration of RNA transcription and protein synthesis. These protein synthesis independent and dependent processes of LTP appear to occur in distinct temporal phases. The early or induction phase of LTP (E-LTP), lasting approximately one hour, is characterised by the rapid modification of pre-existing synaptic proteins to transiently enhance synaptic transmission (Costa-Mattioli *et al*. 2009; Kelleher *et al*. 2004; Park *et al*. 2006). Whilst independent of protein synthesis, the E-LTP phase is an essential prerequisite for subsequent more stable protein synthesis-dependent phases. The intermediate phase of LTP (I-LTP), persisting for approximately eight hours, requires the localised editing and translation of pre-existing mRNAs at the synapse (Bliss and Collingridge 1993; Matthies *et al*. 1990; Mercer *et al*. 2008a; Raymond 2007; Stough *et al*. 2006). The late or maintenance phase of LTP (L-LTP), enduring longer than eight hours, is dependent upon novel RNA transcription and protein synthesis to persistently structurally remodel synapses (Costa-Mattioli *et al*.. 2009; Kelleher *et al*.. 2004; Park *et al*.. 2006). During the L-LTP phase, synthesis of all protein components involved in transcriptional regulation, neutrite outgrowth, composition of the dendritic spine and synaptic terminals occurs to facilitate the growth and remodelling of synapses (Kelleher *et al*. 2004; Park *et al*. 2006).

Numerous studies have identified proteins and genes involved in LTP, categorising them as immediate early genes (IEG) or late response genes (LRG) depending on the temporal phase in which they are expressed (Hermey *et al*. 2013). Following LTP induction IEG are rapidly

and transiently expressed with many acting as transcription factors that regulate the expressison of LRG (Hermey *et al*. 2013). Corresponding with the early and intermediate phases of LTP, IEG are often translated from pre-existing mRNAs at the synapse (Hermey *et al*. 2013). The LRG, tends to be expressed during the L-LTP phase and it requires synthesis of RNA and proteins to affect lasting changes in synaptic structure (Hermey *et al*. 2013; Hong *et al*. 2004). At present IEG have been most extensively studied whilst the LRG with which they interact and the molecular mechanisms by which LRGs incite LTP remain largely unexplored.

1.2. Non-coding RNAs

The protein-coding genes involved in the expression of LTP potentaion have been extensively investigated, however the non-protein-coding (non-coding RNA or ncRNA) components of the genome remain largely uncharacterised. The ~20,000 protein-coding genes in humans account for only ~1.2% of the over ~3 billion base pairs (bp) that make up the human genome (Harrow *et al*. 2012; Lander *et al*. 2001). The remaining ~98.8% of the genome that does not encode proteins was originally disregarded as 'junk DNA' – evolutionary residue not considered to serve biological functions (Mattick 2011; Mills *et al*. 2015). Numerous studies have however shown that the mammalian genome is pervaisvely transcribed with as much as ~80-90% of the human genome actively transcribed (Bertone *et al*. 2004; Birney *et al*. 2007; Consortium 2012; Ip and Nakagawa 2012; Johnson *et al*. 2005; Kapranov *et al*. 2005). This supposed 'junk' DNA is transcribed into ncRNAs in a spatio-temporally regulated manner with highly tissue-, cell- and subsellular-specific expression patterns suggesting that these transcripts are biologically functional rather than just 'transcriptional noise' (Kadakkuzha *et al*. 2015; Mattick 2011; Mercer *et al*. 2008b; Taft *et al*. 2007). Although protein-coding transcirpts tend to be expressed in higher quantities, the transcriptome is predominantly comprised of non-coding RNAs (ncRNAs) (Birney *et al*. 2007; Carninci *et al*. 2005).

It was initially anticipated that the number of protein-coding genes would correlate with organism complexity however the advent of genomic sequencing has revealed that the number of protein-coding genes is relatively conserved throughout the metazoan kingdom (Mattick 2011; Taft *et al*. 2007). Higher vertebrates, such mice and humans, have ~20,000-25,000 protein-coding genes (Chinwalla *et al*. 2002; Harrow *et al*. 2012; Lander *et al*. 2001)

Simpler invertebrates such as the nematode *C. elegans* have ~19000 genes (Consortium 1998). In contrast the ~37, 000 of protein-coding genes in plants such as rice (*Oryza sativa L.*) far exceeds that of higher vertebrates demonstrating that gene number does not correlate with cognitive and developmental complexity (Mattick 2011; Sequencing ProjectInternational Rice 2005; Taft *et al*. 2007). The proportion of ncRNA sequences, on the other hand, appears to scale with organismal complexity, with higher organisms exhibiting the highest levels of ncRNAs (Taft *et al*. 2007). Protein-coding genes account for ~35-97% of bacterial genomes, ~10-75% of protist genomes, ~1-27% of metazoan genomes and only ~1-2% of primate genomes (Barry *et al*. 2014).

Non-coding RNAs play important roles in transcriptional regulation and epigenetic mechanisms forming a cryptic level of gene expression regulation (Lee 2012; Mattick 2011; Mills *et al*. 2015; Qu and Adelson 2012b). Expansion of the ncRNA repertoire appears to have been necessary for the emergence of higher organismal complexity and cognition, coordinating development and molecular function through sophisticated spatiotemoral regulation of protein-coding gene expression (Barry *et al*. 2014; Mattick 2001, 2007; Mattick 2011; Taft *et al*. 2007). These ncRNAs perform vital roles in numerous biological processes including transcription, translation and pre-mRNA processing, regulating gene expression through ncRNA-DNA, ncRNA-RNA and ncRNA-protein interactions (Qu and Adelson 2012b; Zhao *et al*. 2016). Non-coding RNA expression is particularly enriched in the mammalian central nervous system and potentially facilitated the development of the immense complexity of the human brain (Briggs *et al*. 2015; Mehler and Mattick 2006; Mercer *et al*. 2008a; Mercer *et al*. 2008b). Differentially expressed across brain regions associated with memory, ncRNAs are implicated in regulation of synaptogenesis, neuronal development and function including the formation of long-term memory (Briggs *et al*. 2015; Mehler and Mattick 2006; Mercer *et al*. 2008a; Mercer *et al*. 2008b).

Non-coding RNAs may be subdivided into various categories delineated by a number of factors including length, structure and genomic locus (Mercer *et al*. 2008a). Primarily distinguished by size, ncRNAs range from short transcripts as little as 10 nucleotides (nt) to long ncRNA (lncRNA) transcripts over 200 nt (Briggs *et al*. 2015; Qu and Adelson 2012a, b; Rinn and Chang 2012). Short ncRNA subcategories have been most extensively studied and

include micro RNAs (miRNAs), small nucleolar RNAs (snoRNAs), short interfering RNAs (siRNAs) and piwi-RNA (piRNA) (Barry *et al*. 2014; Goodrich and Kugel 2006). Short ncRNA regulatory mechanisms have been observed in bacteria, unicellular eukaryotes and metazoans well-described roles in the regulation of gene expression (Barry *et al*. 2014; Goodrich and Kugel 2006; Mercer *et al*. 2008a; Ryan *et al*. 2015).

Whilst small ncRNAs represent an ancient regulatory mechanism of gene expression, lncRNAs have only been identified in metazoan species and are far less extensively characterised (Barry *et al*. 2014). First characterised in the early 1990s, lncRNAs are classified as non-protein-coding transcripts over 200 nt in length (Brannan *et al*. 1990; Brown *et al*. 1991). Approximately 92% of annotated lncRNAs produced no detectable peptides (Derrien *et al*. 2012). Categories of long ncRNAs include circular RNAs (circRNA), intronic long ncRNAs, antisense long ncRNAs and intervening long ncRNAs (lincRNAs) (Rinn and Chang 2012). Long ncRNAs are dynamically expressed, independent of other functional genomic elements. Like protein-coding genes, they exhibit promoter regions, intron-exon boundaries and alternative splicing patterns (Barry *et al*. 2014; Briggs *et al*. 2015). They however tend to be less polyadenylated, localise to the nucleus and are expressed in a more tissue specific manner (Barry *et al*. 2014; Briggs *et al*. 2015). They may be single exon or multi-exon transcripts undergoing alternative splicing (Briggs *et al*. 2015). Relative to non-transcribed regions of the genome the primary sequence of lncRNAs are evolutionarily conserved, which suggests a biological function, however when compared with the sequences of protein-coding and small ncRNAs genes lncRNAs appear less conserved (Guttman *et al*. 2009; Ponjavic *et al*. 2007; Qu and Adelson 2012b; Zhao *et al*. 2015). The primary sequence of lncRNAs is less constrained than that of protein coding trancripts, imbuing greater adaptive potential while higher conservation of secondary strucutre preserves biological functions of transcripts (Pang *et al*. 2006; Qu and Adelson 2012b; Smith *et al*. 2013). Interestingly, the lncRNA genomic loci, exonic sequences and promoter regions are more highly conserved than the primary sequence (Guttman *et al*. 2009; Ponjavic *et al*. 2007; Zhao *et al*. 2015). Approximately 60-81% of annotated lncRNAs appear to be primate-specific, having evolved within the last 25 million years and exhibiting low sequence conservation (Derrien et al. 2012; Lin et al. 2016; Necsulea et al. 2014; Washietl et al. 2014). Meanwhile,

more ancient lncRNAs, which arose over 90 million years ago, exhibit higher sequence conservation (Necsulea et al. 2014).

The NONCODE database (http://www.bioinfo.org/noncode/) is the most comprehensive public archive of ncRNA data, amassing 527,336 lncRNA transcripts across 16 species (Zhao *et al*. 2015). Humans and mice are the most comprehensively annotated with 167,150 and 130,558 lncRNA transcripts, respectively (Zhao *et al*. 2015). Over 68% of the human transcriptome has been classifed as lncRNAs and while the number of identified ncRNAs has increased exponentially, relatively few have been functionally characterised (Iyer *et al*. 2015; Scadden 2009). Long ncRNAs are highly differentially expressed across brain regions, tissues and cell types to the extent that lncRNA expression patterns are more informative about cell type than protein-coding gene expresssion (Mercer *et al*. 2008b; Molyneaux *et al*. 2015). Approximately 40% of annotated lncRNAs exhibit brain-specific expression patterns potentially indicating a role in neural development and function (Derrien *et al*. 2012; Harrow 2016; Zhao *et al*. 2016). Expression of lncRNAs is dynamically regulated in response to neuronal activity potentially indicating a role in mechanisms of synaptic plasticity (Barry *et al*. 2014; Kim *et al*. 2010; Lipovich *et al*. 2012). Together long ncRNAs appear to help modulate every stage of gene expression (Änkö and Neugebauer 2010).

Long intervening ncRNAs (lincRNAs), a lncRNA subcategory, are of particular interest in the regulation of learning and memory formation as they are most abundant in the brains of higher vertebrates (Mills *et al*. 2015). LincRNAs are transcripts over 200 nt in length that exhibit tissue specific expression patterns, lack protein coding capacity and are transcribed from genomic loci in the intervening regions between protein-coding gene loci (Mills *et al*. 2015). Approximately one third of lincRNAs is primate-specific, with the richest source being the human brain, where they are believed to serve as part of a highly complex mechanism of synaptic plasticity through regulation of gene expression (Briggs et al. 2015; Derrien et al. 2012; Mills et al. 2015).

1.3. Neural cell adhesion molecules 2 (NCAM2)

Cell adhesion molecules (CAMs) are membrane bound glycoproteins present on the surface of cells (Benson and Huntley 2012; Chua *et al*. 2010; Edelman and Crossin 1991; Leshchyns'ka *et al*. 2015; Rønn *et al*. 2000; Sytnyk *et al*. 2002; Winther *et al*. 2012). CAMs

facilitate cell-cell adhesion through homophilic and heterophilic *trans*-interactions with other cells and components of the extracellular matrix (Benson and Huntley 2012; Chua *et al*. 2010; Edelman and Crossin 1991; Leshchyns'ka *et al*. 2015; Rønn *et al*. 2000; Sytnyk *et al*. 2002; Winther *et al*. 2012). Binding of CAMs also initiates intracellular signalling pathways which stimulate metabolic and structural changes in the cell (Rønn *et al*. 2000). In neurons CAMs form physical connections between the pre- and post-synaptic membranes, stabilising the synapses (Benson and Huntley 2012; Chua *et al*. 2010; Edelman and Crossin 1991; Leshchyns'ka *et al*. 2015; Sytnyk *et al*. 2002; Winther *et al*. 2012).

Neural cell adhesion molecule 2 (NCAM2) is a member of the immunoglobulin superfamily of CAMs that is enriched at synapses of the hippocampus in both human and mice (Leshchyns'ka *et al*. 2015; Winther *et al*. 2012). The extracellular domains of the NCAM2 protein take part in homophilic *trans*-interactions with proteins of the membranes of other neurons, forming physical connections between the pre- and postsynaptic membranes (Winther *et al*. 2012; Yoshihara *et al*. 1997).

Besides its involvement in cell-cell adhesion NCAM2 triggers intracellular signalling pathways associated with regulation of neurite outgrowth, synapse creation, axonal guidance and formation of dendritic bundles (Hamlin *et al*. 2004; Ichinohe *et al*. 2003; Sheng *et al*. 2015; Winther *et al*. 2012). Thus, expression of NCAM2 is crucial in neuronal development as well as in function and maintenance of mature cells (Sheng *et al*. 2015; Walz *et al*. 2006; Winther *et al*. 2012). Depletion or alteration of NCAM2 has been associated with developmental and neurodegenerative disorders including autism and Alzheimer's disease (AD) (Kulahin and Walmod 2010; Petit et al. 2015).

Alzheimer's disease (AD) is the most common and devastating neurodegenerative disease in humans (Alzheimer's 2009; Brookmeyer *et al*. 2007; Han *et al*. 2010). AD is the leading cause of dementia affecting millions of people worldwide and projected to affect millions more as improved living conditions and medical advancements have led to increased life expectancy (Alzheimer's 2009; Han *et al*. 2010). AD is characterised by chronic progressive cognitive impairment with symptoms including gradual decline in episodic memory, behavioural changes, impaired judgement, loss of social skills, difficulty in communicating, disorientation and emotional instability (Alzheimer's 2009; Han *et al*. 2010; Karttunen *et al*. 2011; Mucke 2009).

One of the earliest and most telling signs of AD onset is synapse loss (Coleman and Yao 2003; Scheff and Price 2003). The pathology of AD is complex, however the accumulation of amyloid-β (Aβ) protein plaques appears to be a key event in the pathogenesis (Hardy and Selkoe 2002; Masters *et al*. 1985). The Aβ precursor protein that accumulates in AD has been observed to bind to NCAM2 at the cell membrane of cultured hippocampal neurons, cleaving the extracellular portion of the NCAM2 molecule and resulting in an accumulation of soluble extracellular NCAM2 fragments (NCAM2-ED) as well as reducing the number of functional NCAM2 at the synapse (Leshchyns'ka *et al*. 2015). In AD hippocampal tissues synaptic NCAM2 levels are reduced whilst soluble NCAM2-ED levels are significantly elevated (Leshchyns'ka *et al*. 2015). Incubation of cultured hippocampal neurons with NCAM2-ED has been demonstrated to induce glutamatergic synapse disassembly (Leshchyns'ka *et al*. 2015). These interactions may represent a possible mechanism of synaptic loss in AD and a better understanding of this pathway could lead to novel treatment strategies for this disease (Leshchyns'ka *et al*. 2015). Synaptic disassembly may result from the sequestering of functional NCAM2 molecules by NCAM2-ED and hindering of normal synaptic adhesion or from intricate intracellular cascades induced by NCAM2 binding. In this study, NCAM2-ED was used to simulate the AD phenotype and induce synapse disassembly in order to examine its effect on neuronal transcriptomes.

1.4. RNA Sequencing

The molecular processes underlying LTP are immensely convoluted. Thus despite decades of focued study our knowledge of the transcriptomics and molecular mechanisms underpinning LTP remains incomplete (Park *et al*. 2006). Current availability of high throughput sequencing techniques such as RNA sequencing (RNA-Seq) opened new perspectives for genome-wide identification of genes involved in LTP. RNA-Seq provides transcriptome coverage to a single base pair (bp) resolution facilitating the discovery and characterization of novel ncRNAs and previously unidentified isoforms of annotated transcripts (Wang *et al*. 2009).

The current study utilised RNA-Seq analysis and primary hippocampal neurons cultures from neonate mice to determine the identity and abundance of protein-coding and non-codingRNA transcripts expressed during LTP in order to attain a better understanding of the molecular mechanisms that underlie this form of synaptic plasticity (Kelleher *et al*. 2004). The effect of

disruption of cellular adhesion on the transcriptome of cells undergoing LTP and treated with recombinant NCAM2 protein, has also been evaluated.

1.5. Aims

The primary objective of this project was to perform a comprehensive analysis of expression profiles of the mRNA and ncRNA transcripts involved in the induction of LTP at synapses of primary hippocampal neurons. Furthermore, the study aimed to investigate changes in gene expression patterns resulting from LTP inhibition through disruption of synaptic adhesion. Specific aims comprised:

Aim 1: To analyse the transcriptome of hippocampal neurons undergoing LTP in order to identify the protein-coding genes and non-coding RNAs expressed during LTP.

Aim 2: To identify changes to the transcriptome following inhibition of LTP via disruption of the synaptic adhesion.

2. Materials and Methods

2.1 Tissue culture, NCAM2-ED treatment and *in vitro* induction of LTP

Hippocampal neurons were isolated from one day old C57Bl mice and maintained for 14 days in a CO_2 incubator as described by Andreyeva *et al*. (2010). The neuronal cultures were grown on glass coverslips coated with poly-$_D$-lysine (100 μg ml^{-1}, Sigma) using a Neurobasal A medium supplemented with 2% B-27, Glutamax and 2 ng ml^{-1} bFGF-2 (all reagents from ThermoFisher) (Andreyeva *et al*. 2010; Leshchyns'ka *et al*. 2015).

The cultured neurons were divided into four treatment groups with three biological samples in each treatment group (Table 1). Recombinant NCAM2-ED was obtained as described by Leshchyns'ka *et al*. (2015). One hour before LTP induction, two groups were mock-treated with cell culture medium whilst the remaining two were treated with 2.5μg/ml of recombinant NCAM2-ED diluted with cell culture medium and incubated for 20 mins in a CO_2 incubator (Table 1).

LTP was induced in two of the treatment groups by treating neurons with 200 μM glycine for 3 min in a Mg^{2+}-free solution containing 150 mM NaCl, 2 mM CaCl2, 4 mM KCl, 10 mM N-2-hydroxyethylpiperazine-N′-2-ethanesulfonic acid, 10 mM glucose at pH 7.4 (Lu *et al*. 2001; Puchkov *et al*. 2011). Successful induction of LTP was demonstrated biochemically by increased AMPA receptors levels in the plasma membrane resulting from synaptic insertion of AMPA receptors (Sytnyk and Leshchysn'ka, unpublished data)

Table 1. Cell treatment groups.

Treatment Group	NCAM2-ED Treatment	LTP Induction
Control	Mock-treated	Untreated
LTP	Mock-treated	LTP
NCAM2-ED	NCAM2-ED	Untreated
LTP+NCAM2-ED	NCAM2-ED	LTP

2.2 RNA extraction and RNA-Seq

Following LTP induction all treatment groups were incubated at 37°C and 5% CO_2. Thirty minutes after the induction of LTP all neuronal cultures were washed with Neurobasal A medium (ThermoFisher), centrifuged and the cell pellets subjected to total RNA isolation using an RNeasy Mini kit (Qiagen). Genomic DNA traces were removed using a RNase-free DNase treatment. The quality of the obtained total RNA was evaluated using Agilent 2100

Bioanalyzer RNA Nano Chip with RNA integrity number (RIN) values ranging from 6.0 to 7.0. Total RNA was selected for poly(A)+ fraction and prepared for RNA-Seq according to manufacturer's guidelines using the Illumina TruSeq Stranded mRNA Library Preparation Kit. RNA sequencing was performed using 100bp paired-end Illumina HiSeq2000 sequencing.

2.3 Mapping of RNA-Seq results and differential expression testing

Entire bioinformatic analysis was performed using next-generation sequencing genomic analysis tools available on the Galaxy web-based platform (Afgan *et al*. 2016). Specifically, the RNA-Seq output was processed using the Tuxedo protocol described by Trapnell *et al*. (2012). The FastQC tool was used to test the quality of RNA-Seq reads and detect any sequencing biases. The Trimmomatic tool was applied to trim the sequence reads for quality improvement (Bolger *et al*. 2014). Following trimming the read quality was assessed again using FastQC and was considered to be acceptable if the mean per sequence quality score was over 20.

The Bowtie2 is a high-throughput short read alignment tool that aligns reads to a reference genome however it cannot align reads with gaps such as those spanning introns (Trapnell *et al*. 2009; Trapnell *et al*. 2012). Forward and reverse reads were aligned to the mouse reference (UCSC mm10) using the Bowtie2 tool with a minimum fragment length of zero and a maximum fragment length of 400. The Bowtie2 outputs were run through the 'Insertion Size Metrics' tool in order to obtain the mean fragment length (identified as mean insert size in the output) and the standard deviation. The mean inner distance (the size of the gap present between the 3' end of the forward and reverse reads) was then calculated using the formula:

$$Mean\ Inner\ Distance = total\ fragment\ length - 2 \times read\ length$$

The mean inner distance and standard deviation results were then used as input data for the TopHat gaped-read mapper tool to identify splice junctions and to align reads to the mm10 reference genome (Trapnell *et al*. 2012). The quality of the TopHat alignment was assessed by running the 'accepted hits' output file through the Flagstat tool.

The Cuffinks tool was used to compile the assembled reads from the 'accepted hits' TopHat output file into transcripts with the UCSC mm10 mouse reference annotation as well as

providing estimates of each transcript's abundance in fragments per kilobase of transcript per million mapped reads (FPKM). The 'assembled transcripts' Cufflinks output files were merged into a single, non-redundant set of transcripts using the Cuffmerge tool. Using the Cuffdiff tool, differential expression testing was performed with the Cuffmerge transcript file and the TopHat 'accepted hits' files as an input inorder to detect statistically significant differences in transcript expression and splicing patterns between the treatment groups. The Cuffdiff tool utilises a beta negative binomial distribution model for statistical analysis (Trapnell et al. 2010). The threshold for statistical significance was highly rigorous with transcripts considered significantly differentially expressed if they obtained a p-value 0.00005 or less. The Cuffdiff results also generated a corrected p-value, the q-value, which was deemed statistically significant if q-value<0.05.

2.4 Gene Ontology enrichment analysis

Protein-coding differentially expressed genes, identified by Cuffdiff, were subjected to gene ontology (GO) enrichment analysis using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (http://cbl-gorilla.cs.technion.ac.il/). The DAVID tool identifies the molecular function, cellular localisation and the biological processes each transcript is involved in within the cell as well as identifies enriched GO clusters of differentially expressed genes (Huang *et al*. 2008; Huang *et al*. 2009). Only annotated transcripts can be utilised by DAVID thus unannotated transcripts and isoforms were removed from the differentially expression results prior to the analysis. The number of statistically significantly differentially expressed genes was not sufficient to perform DAVID analysis due to the highly stringent p-value of 0.00005. To enable DAVID analysis, the top 100 transcripts with the lowest p-values ($p<0.05$) were included as transcripts of potential biological interest (Table S1). The results of the DAVID GO enrichment analysis with p-values$\leq$0.05 were visualised using the Cytoscape Enrichment Map plugin (Merico *et al*. 2011).

2.5 Characterisation of unannotated lncRNAs: Determination of protein-coding potential and RNA secondary structure.

Three of the transcripts identified as being significantly differentially expressed were unannotated. Unannotated transcripts were considered putative lincRNAs if they (i) were

found to be longer than 200 nucleotides, (ii) were located between protein-coding loci with no overlap and (iii) bore putative open reading frames (ORF) less than 100 amino acids (aa) in length (Hangauer *et al*. 2013).

To characterise the transcripts, they were visualised using the Interactive Genomics Viewer (IGV) (http://www.broadinstitute.org/igv/) and categorised as intervening if they were located on the genome in-between protein-coding loci and did not overlap with annotated transcribed regions on either the sense or antisense strand. The sequence of unannotated transcripts was obtained from IGV.

Putative ORF were detected using the NCBI ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/). Predicted ORFs were only accepted if they had a distinct start codon and either a stop codon or an ORF that ran off the end of the transcript.

The minimum free energy (MFE) secondary structure of the RNA transcripts was predicted using the RNAfold web server (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) (Gruber *et al*. 2008; Zuker and Stiegler 1981). Transcripts over 7500 nucleotides long cannot be modelled using RNAfold and hence were not considered for this analysis. The ORFs were considered potentially transcriptionally active if they were flanked by single stranded RNA structures (Wan *et al*. 2014).

To determine whether predicted ORFs were translationally active, the Genome Wide Information on Protein Synthesis (GWIPS-viz) browser (http://gwips.ucc.ie/cgi-bin/hgGateway) was used to determine whether ribosomal interactions were detected at the sequence of the predicted ORFs. The GWIPS-viz browser is an online repository for Ribo-Seq data(Michel et al. 2013). The Ribo-Seq data set used for this analysis was derived from Cho *et al*. (2015) study of transcriptional and translational regulation during memory formation in the mouse hippocampus.

2.6 Meta-analysis of rat RNA-Seq data

In a recently published study Maag *et al*. (2015) utilised RNA-Seq to generate transcriptome profiles of LTP in the hippocampus of adult rats. The study employed a time course analysis of the transcriptome at 30 mins, 2 h and 5 h after the induction of LTP. To compare the results of this study against mouse gene expression profiles generated in the current study, the RNA-Seq transcriptome data for the control and 30 min post LTP groups were downloaded

from ArrayExpress (https://www.ebi.ac.uk/arrayexpress/), an online repository of genomic data, (accession number E-MTAB-3375). The 30 min post LTP group was chosen as the mouse RNA was harvested 30 mins after LTP induction making it the most appropriate time point for comparison. The raw rat RNA-Seq data was analysed using the methods described in Section 2.2 and 2.3 using the UCSC Rn6 rat reference genome.

3. Results

3.1 General RNA-Seq metrics

Bioinformatic analysis detected an average of 27405 genes with distinct 47686 isoforms per sample across all investigated conditions (Table 2). Differential expression testing yielded six pairwise comparisons of the four conditions. Of most biological interest were the comparisons between Control vs. LTP, NCAM2-ED vs. LTP+NCAM2-ED, Control vs. NCAM2-ED and LTP vs. LTP+NCAM2-ED. The remaining two comparisons comprised Control vs. LTP+NCAM2-ED and LTP vs. NCAM2-ED. In total 64 differentially expressed genes were identified across the six pairwise comparisons (Table 3 and Table S2). Of the 64 differentially expressed genes detected, 54 encode protein-coding genes whist the remaining nine non-protein-coding genes encode three lincRNAs, two miRNAs and four previously unannotated transcripts (Table S3).

Table 2. Number of genes and isoforms detected in each sample.

Treatment Group	Replicate	Genes	Isoforms
Control	1	27339	47667
	2	27274	47370
	3	27212	47659
	mean	27275	47565
LTP	1	27349	47568
	2	27039	47120
	3	27621	47973
	mean	27336	47554
NCAM2-ED	1	27361	47647
	2	27183	47419
	3	28742	49117
	mean	27762	48061
LTP+NCAM2-ED	1	27471	47796
	2	27204	47303
	3	27060	47598
	mean	27245	47565

The genes identified as significantly differentially expressed between the control and LTP groups were all downregulated in the presence of LTP (Table 4). Conversely, amongst neurons that were treated with NCAM2-ED, 30 of the 35 genes were upregulated in response to LTP induction, whereas treatment of resting neurons with NCAM2-ED led to a downregulation of 20 of the 24 genes. In the LTP vs. LTP+NCAM2-ED pairwise

comparison, the presence of NCAM2-ED resulted in upregulation of 9 out of 10 significantly differentially expressed genes (Table 4).

The NCAM2-ED molecule does not impair induction LTP induction however it does alter the dynamics of LTP induction (Sytnyk and Leshchysn'ka, unpublished data). Changes to LTP-dependent gene expression in the presence of NCAM2-ED most likely reflect these synaptic abnormalities.

Table 3. Number of significantly differentially expressed annotated and unannotated transcripts for each pairwise analysis of treatment groups.

Pair-wise Comparison	Significantly DE Transcripts *	Significantly DE Unannotated Transcripts *
Control vs. LTP	11	1
Control vs. NCAM2-ED	24	3
LTP v LTP+NCAM2-ED	10	0
NCAM2-ED vs. LTP+NCAM2-ED	35	2
Control vs. LTP+NCAM2-ED	4	0
LTP vs. NCAM2-ED	24	2
Total	64	4

*The totals in the table refer to the number of significantly differentially expressed (DE) genes detected across all pairwise comparisons without redundancy.

Table 4. Number of upregulated and downregulated significantly differentially expressed genes.

Pair-wise Comparison	Upregulated*	Downregulated*
Control vs. LTP	0	11
Control vs. NCAM2-ED	4	20
LTP vs. LTP+NCAM2-ED	9	1
NCAM2-ED vs. LTP+NCAM2-ED	30	5
Control vs. LTP+NCAM2-ED	3	1
LTP vs. NCAM2-ED	18	6

*In each pairwise comparison genes were considered upregulated if they exhibited higher FPKM values in the latter treatment group and downregulated if they exhibited lower FPKM values in the latter treatment group.

3.2 Expression pattern of significantly differentially expressed genes

The FPKM results of 64 significantly differentially expressed genes were plotted to identify expression patterns for individual pairwise comparisons and to determine the effect of each treatment on gene expression. Although FKPM levels varied greatly between genes most exhibited one of four expression patterns. Four genes were chosen at random and plotted to illustrate each of the four observed expression patterns. The Cuffdiff differential expression test p-value results are shown to indicate significant differential expression between treatment groups, supporting the efficacy of observed expression patterns.

Expression pattern 1 was the most common with 31 significantly differentially expressed genes characterised by downregulation of gene expression following induction of LTP or treatment with NCAM2-ED. In contrast induction of LTP in the presence of NCAM2-ED led to upregulation of gene expression (Fig. 1).

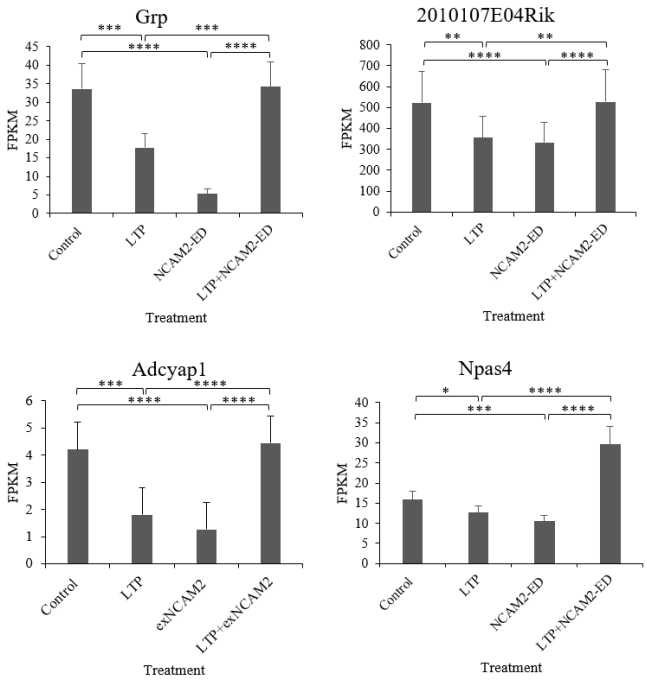

Figure 1. Expression pattern 1: Gene expression is down regulated in response to LTP and the application of NCAM2-ED however when LTP is induced in the presence of NCAM2-ED expression is upregulated. P-values≤0.00005 are indicated by ****, ≤0.0005 by ***, ≤0.005 by ** and ≤0.05 by *. Error bars are ±SD.

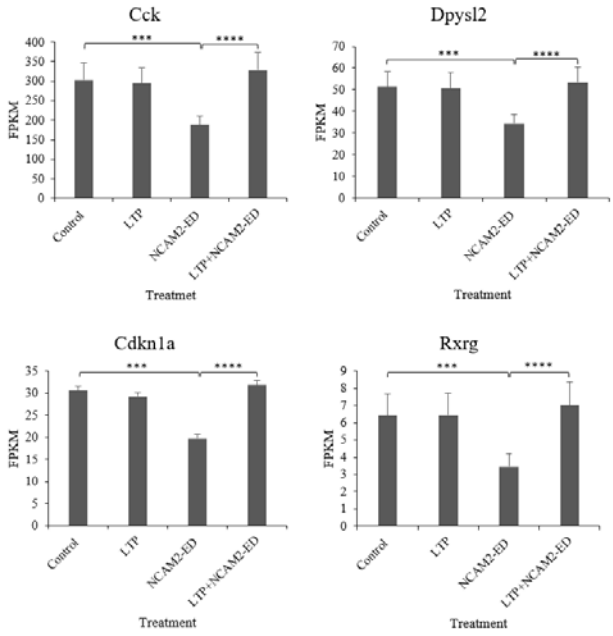

Figure 2. Expression pattern 2: Gene expression is not significantly altered in response to LTP however application of NCAM2 led to a significant downregulation of transcript levels. Further, induction of LTP in the presence of NCAM2 resulted in upregulation bringing gene expression levels to those observed in the absence of NCAM2. P-values≤0.00005 are indicated by **** and ≤0.0005 by ***. Error bars are ±SD.

Expression of the eight genes represented by expression pattern 2 was not significantly altered in response to LTP however application of NCAM2 lead to a significant downregulation of expression (Fig. 2). Induction of LTP did not significantly alter gene expression. In the presence of NCAM2-ED however induction of LTP resulted in upregulation of the gene expression, restoring transcript levels to those observed in the absence of NCAM2-ED.

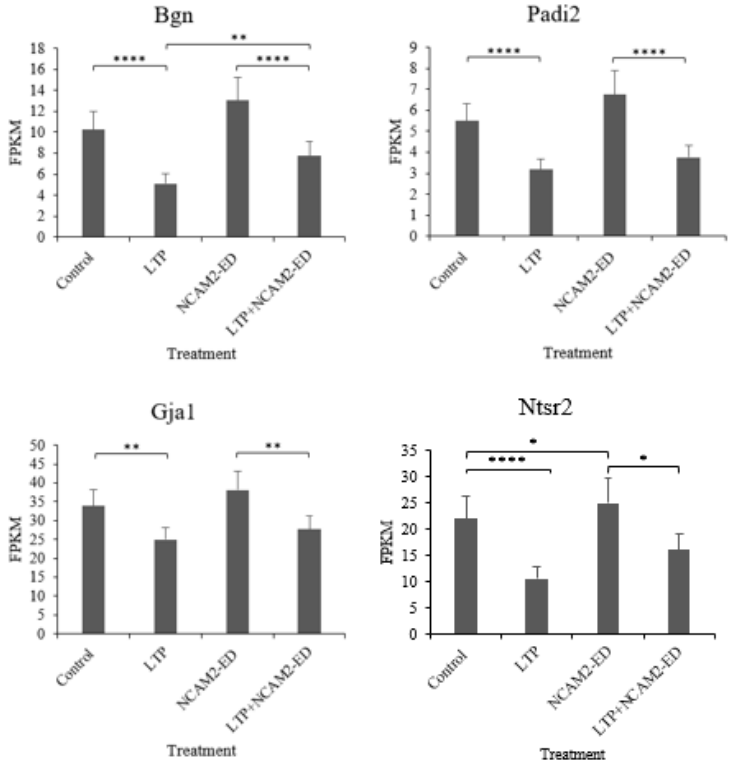

Figure 3. Expression pattern 3: Induction of LTP results in downregulation of gene expression. P-values≤0.00005 are indicated by ****, ≤0.0005 by *** and ≤0.005 by **. Error bars are ±SD.

Expression pattern 3 was represented by 18 genes and it is characterised by downregulation in response to induction of LTP (Fig. 3). The presence of NCAM2-ED has had some effect on gene expression but in all cases the induction of LTP led to downregulation of gene expression.

The five genes encapsulated in expression pattern 4 showed upregulation of gene expression in the presence of NCAM2-ED (Fig. 4). Four of the five genes exhibited slight downregulation or no change in response to induction of LTP. Finally, the expression of the Fos proto-oncogene, AP-1 transcription factor subunit (Fos) gene did not match any of the four patterns, with induction of LTP and NCAM2-ED treatment resulting in upregulation (Fig. 5).

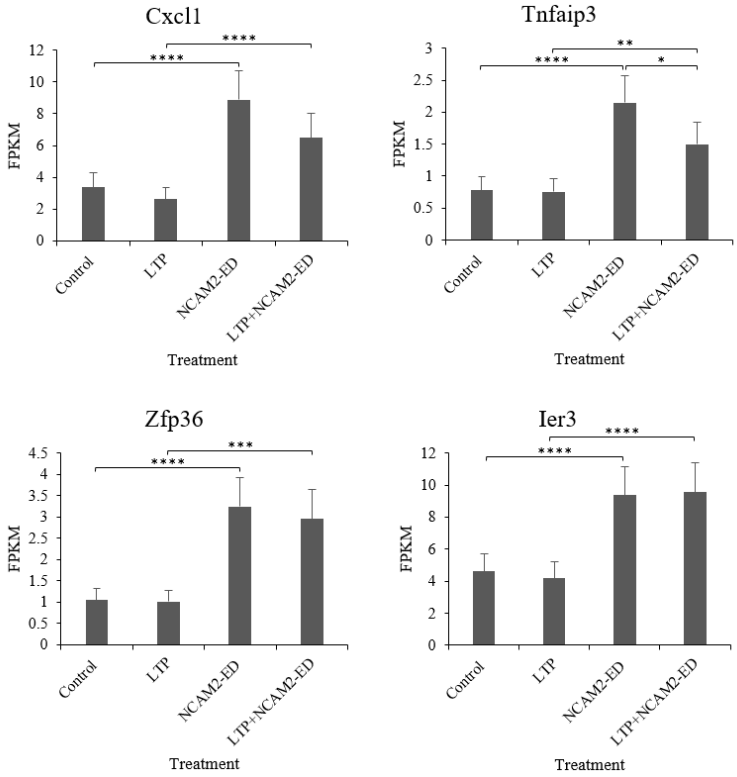

Figure 4. Expression pattern 4: Gene expression is upregulated in response to NCAM2-ED application. P-values≤0.00005 are indicated by ****, ≤0.0005 by ***, ≤0.005 by ** and ≤0.05 by *. Error bars are ±SD.

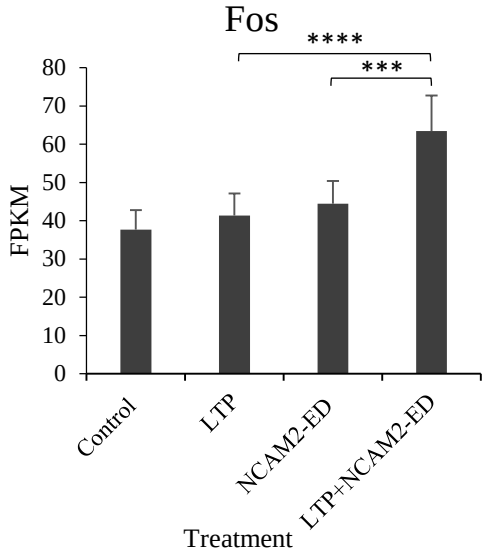

Figure 5. Fos expression pattern: The Fos gene was upregulated in response to induction of LTP and the application of NCAM2-ED also lead to upregulation of expression. P-values≤0.00005 are indicated by **** and ≤0.0005 by ***. Error bars are ±SD.

3.3 Gene ontology enrichment analysis of significantly differentially expressed genes

Gene ontology (GO) enrichment analysis was performed using the top 100 genes with the lowest p-values for each pairwise comparison (Table S1). The full DAVID results for each comparison can be found in supplementary table 4. The results of each DAVID analysis were visualised as GO enrichment maps using the Cytoscape program (Figs. 6-9).

Gene ontology enrichment analysis for Control vs. LTP comparison yielded 81 GO terms (listed in Table S4) of which 49 exhibited p-values$\leq$0.05 thus were incorporated into an enrichment map (Fig.6). Amongst the top GO terms identified for the Control vs. LTP comparison were wound healing and pathways related to immune function, inflammatory response and the complement cascade (Fig. 6 and Table 5). Other GO terms comprised extracellular matrix structure and function, regulation gene expression, intracellular signalling pathways such as the JAK-STAT cascade and cellular response to various cell surface molecules (Fig. 6, Table S4 and Table 5). Several unexpected GO terms such as 'response to gamma radiation', 'heart development' and 'embryo implantation' were also identified (Fig.6 and Table S4).

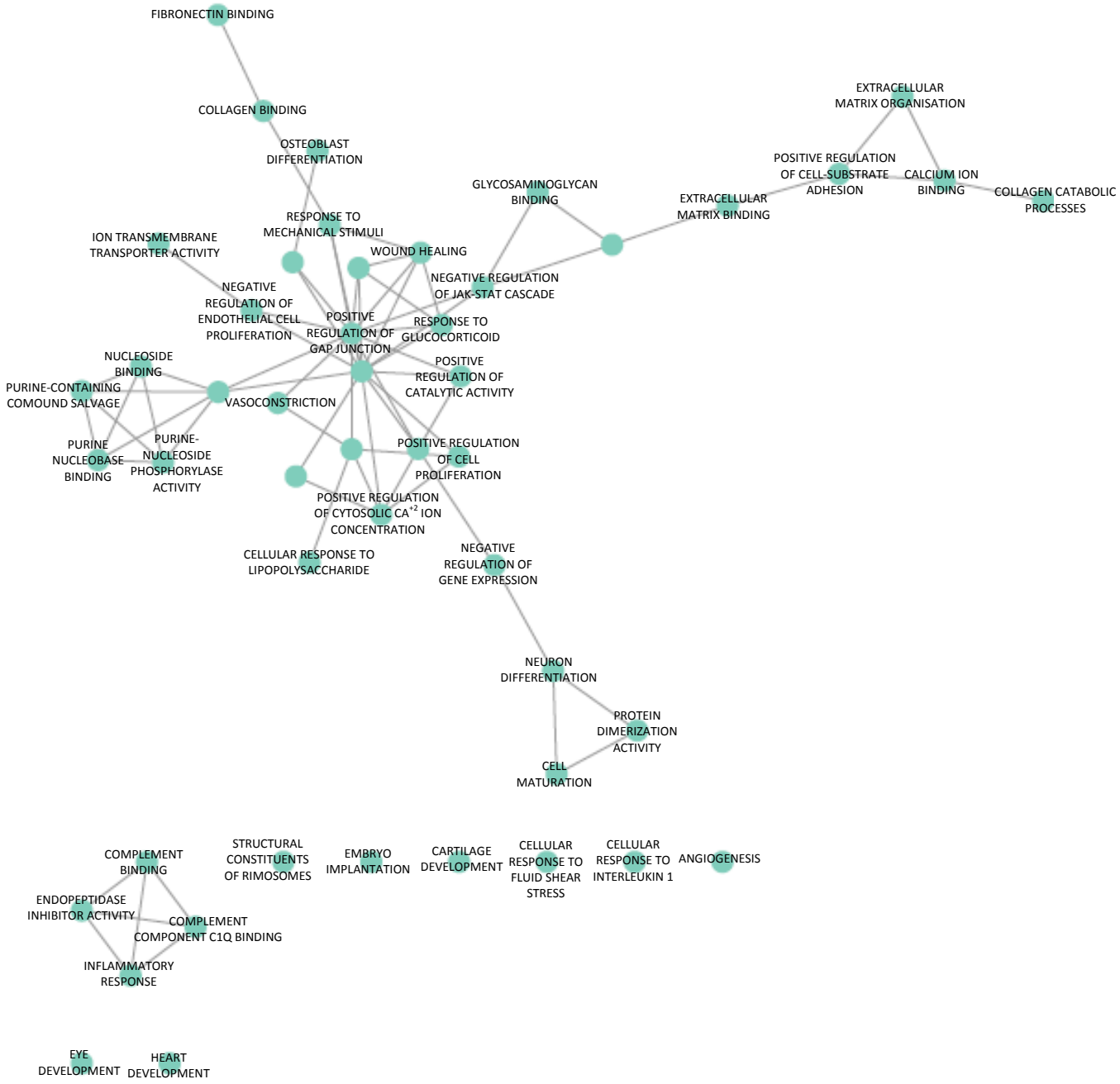

Figure 6. Gene ontology enrichment map of genes identified as differentially expressed in Control vs. LTP comparison. The blue circles represent GO term clusters and the grey lines linking the GO term circles represent genes shared between the clusters. Clusters that are not linked to any other clusters do not share genes with any other GO terms.

Table 5. Top 10 GO terms enriched in the Control vs. LTP comparison.

No.	GO Term		No. of genes	Genes	P-value
1	GO:0042060	wound healing	6	CAV1, IGF2, DCN, SDC4, PLAU, FN1	1.55E-04
2	GO:0030574	collagen catabolic process	4	MMP10, CTSK, MMP3, MMP13	2.29E-04
3	GO:0050840	extracellular matrix binding	4	BGN, SMOC1, DCN, SPP1	3.44E-04
4	GO:0071222	cellular response to lipopolysaccharide	7	EDNRB, CXCL5, AXL, ENTPD2, PLAU, B2M, FN1	8.24E-04
5	GO:0071498	cellular response to fluid shear stress	3	MMP13, PLAU, SPP1	0.002857
6	GO:0042310	vasoconstriction	3	EDNRB, CAV1, AGT	0.004032
7	GO:0009612	response to mechanical stimulus	4	CAV1, CHIL1, DCN, MMP13	0.004649
8	GO:0005539	glycosaminoglycan binding	3	SMOC2, BGN, DCN	0.004724
9	GO:0008284	positive regulation of cell proliferation	9	EDNRB, HES5, AGT, CST3, MYEOV2, MFGE8, PLAU, FN1, ADCYAP1	0.005939
10	GO:0007204	positive regulation of cytosolic calcium ion concentration	5	EDNRB, AGT, GJA1, CIB2, ADCYAP1	0.00706

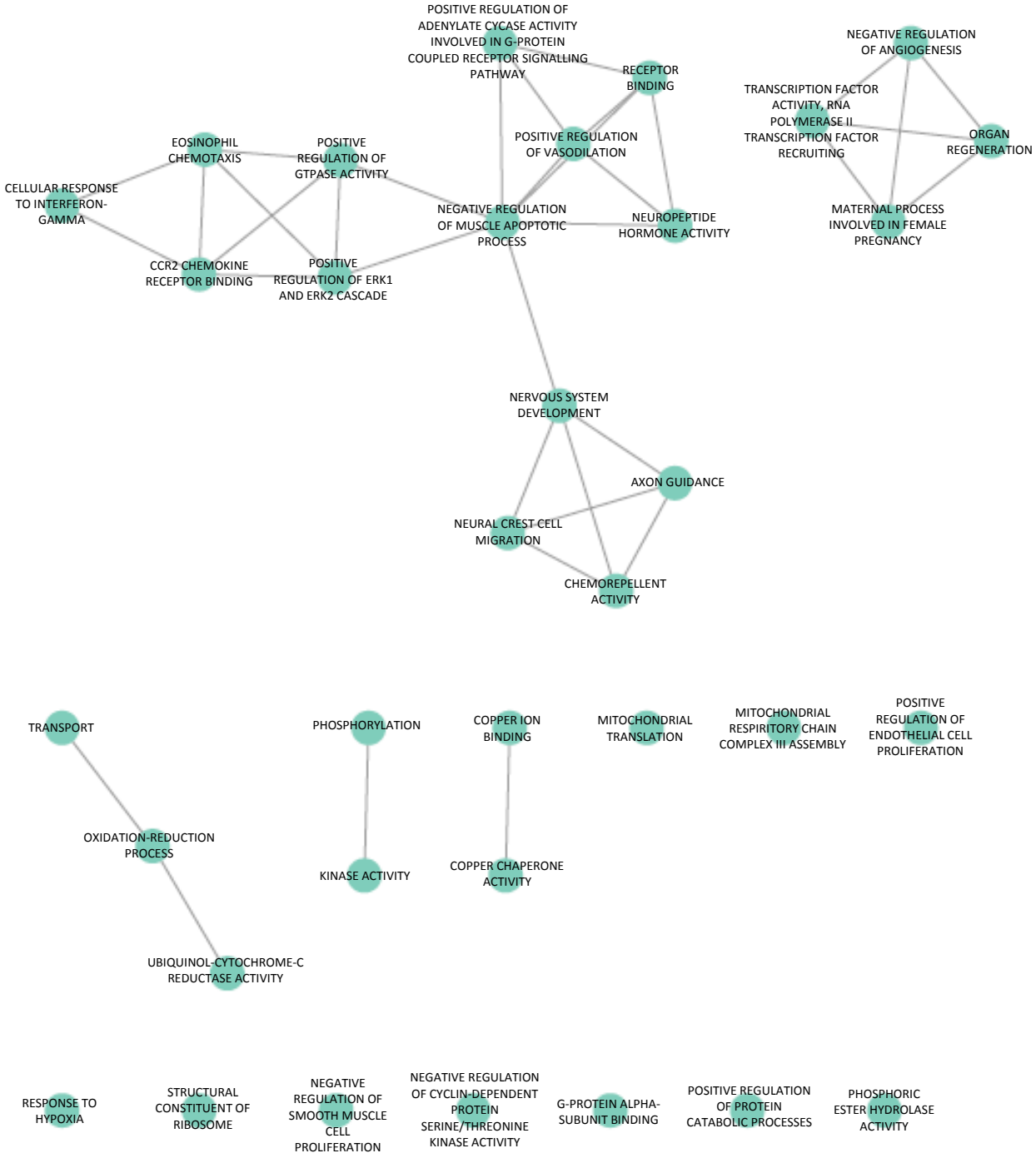

Figure 7. Gene ontology enrichment map of genes identified as differentially expressed in Control vs. NCAM2-ED comparison.

Gene ontology enrichment analysis of genes differentially expressed in Control vs. NCAM2-ED comparisons yielded 57 GO terms, listed in Table S3, of which 35 exhibited p-values≤0.05 and hence have been incorporated into an enrichment map, presented in Figure 8. Enriched GO terms comprised intracellular signalling cascades such as the ERK1/ERK2 cascade, neuropeptide hormone activity, axon guidance, regulation of transcription and translation (Fig. 7 and Table 6). As with the Control vs. LTP results, several unusual GO terms were identified including 'maternal process involved in female pregnancy', 'organ regeneration' and 'negative regulation of muscle apoptotic process' (Fig. 7 and Table S3).

Table 6. Top 10 GO terms enriched in the Control vs. NCAM2-ED comparison.

No.	GO Term		No. of genes	Genes	P-value
1	GO:0008121	ubiquinol-cytochrome-c reductase activity	3	UQCR10, UQCR11, UQCRQ	4.60E-04
2	GO:0045909	positive regulation of vasodilation	4	VIP, PTPRM, NPPC, ADCYAP1	7.98E-04
3	GO:0010579	positive regulation of adenylate cyclase activity involved in G-protein coupled receptor signaling pathway	3	VIP, CRHR1, ADCYAP1	8.02E-04
4	GO:0045736	negative regulation of cyclin-dependent protein serine/threonine kinase activity	3	CDKN1A, TNFAIP3, LATS2	0.004941
5	GO:0045499	chemorepellent activity	3	SEMA5A, SEMA6A, EFNA5	0.007233
6	GO:0005184	neuropeptide hormone activity	3	GRP, NPPC, ADCYAP1	0.008315
7	GO:0060135	maternal process involved in female pregnancy	3	LIF, PAM, CCL2	0.008498
8	GO:0003735	structural constituent of ribosome	6	RPS26, RPL41, NDUFA7, RPL26, MRPL57, RPS24	0.00937
9	GO:0007399	nervous system development	7	SEMA5A, SEMA6A, ERBB4, EFNA5, DPYSL2, HDAC9, ADCYAP1	0.009445
10	GO:0016531	copper chaperone activity	2	ATOX1, COX17	0.00947

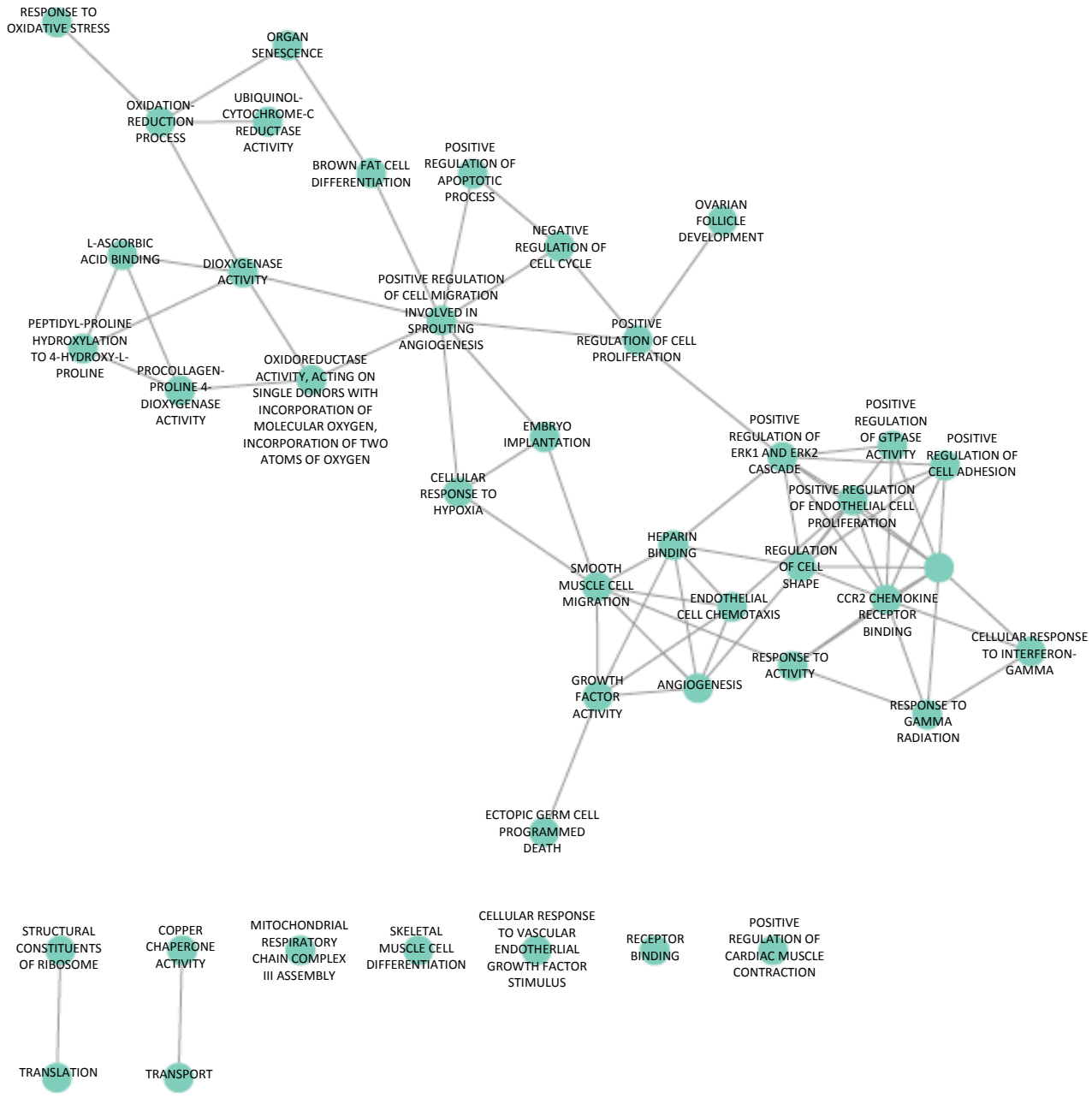

Figure 8. Gene ontology enrichment map of genes identified as differentially expressed in LTP vs. LTP+NCAM2-ED comparison.

Gene ontology enrichment analysis of differentially expressed genes in the LTP vs. LTP+NCAM2-ED comparison yielded 71 GO terms clusters (Table S3) of which 42 had p-values≤0.05. GO terms included gene clusters involved in translation, intracellular signalling cascades such as the ERK1/ERK2 cascade, cell adhesion and regulation of cell shape (Fig.8 and Table 7). Other GO terms not manifestly related to neuronal physiology

were also identified including ‘skeletal muscle cell differentiation’, ‘ovarian follicle development’ and ‘response to gamma radiation’ (Fig.8 and Table S3).

Table 7. Top 10 GO terms enriched in the LTP vs. LTP+NCAM2-ED comparison.

No.	GO Term		No. of genes	Genes	P-value
1	GO:0008201	heparin binding	7	NOV, CCL2, CTGF, VEGFA, LIPG, CCL7, FN1	7.68E-05
2	GO:0001525	Angiogenesis	8	NOV, CCL2, PTGS2, CTGF, PLXDC1, VEGFA, PLAU, FN1	1.38E-04
3	GO:0071456	cellular response to hypoxia	6	PTGS2, VEGFA, FAM162A, NDRG1, ERO1L, PLAU	1.41E-04
4	GO:0003735	structural constituent of ribosome	8	RPL30, MRPL13, RPL41, RPS29, MRPL14, NDUFA7, RPL26, SLC25A45	2.60E-04
5	GO:0018401	peptidyl-proline hydroxylation to 4-hydroxy-L-proline	3	P4HA2, P4HA1, EGLN3	3.22E-04
6	GO:0006979	response to oxidative stress	6	NDUFB4, PTGS2, ATOX1, NDUFA6, MSRB1, TRPM2	3.90E-04
7	GO:0008083	growth factor activity	6	NOV, CTGF, VEGFA, FGF11, NENF, KITL	5.65E-04
8	GO:0035767	endothelial cell chemotaxis	3	NOV, VEGFA, NR4A1	9.54E-04
9	GO:0043065	positive regulation of apoptotic process	8	PTGS2, RPS29, DUSP1, PRKDC, NR4A1, FAM162A, GADD45B, LATS2	0.001
10	GO:0055114	oxidation-reduction process	11	UQCR10, UQCR11, PTGS2, DIO2, P4HA2, P4HA1, NDUFA6, NDUFA7, EGLN3, MSRB1, ERO1L	0.001

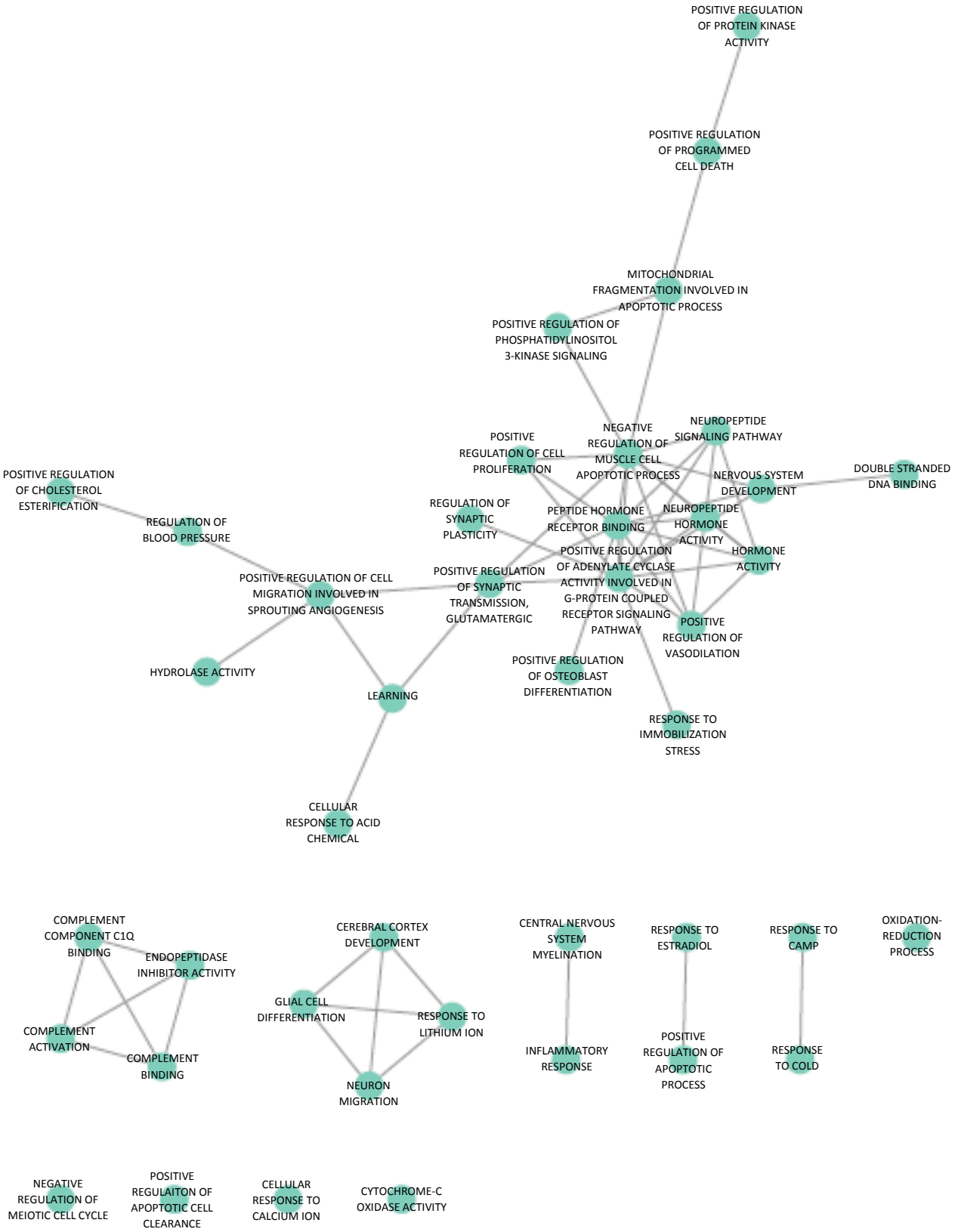

Figure 9. Gene ontology enrichment map of genes identified as differentially expressed in NCAM2-ED vs. LTP+NCAM2-ED comparison.

Gene ontology enrichment resulting from the NCAM2-ED vs. LTP+NCAM2-ED comparative analysis yielded 69 gene clusters (Table S3). Amongst the 43 significantly enriched clusters were biological processes such as intracellular signalling, protein kinase activity, glutaminergic synaptic transmission, regulation of synaptic plasticity, learning, inflammatory response and the complement cascade (Fig.9 and Table 8).

Table 8. Top 10 GO terms enriched in NCAM2-ED vs. LTP+NCAM2-ED comparison.

No.	GO Term		No. of genes	Genes	P-value
1	GO:0005184	neuropeptide hormone activity	6	GRP, CCK, NPY, NPPC, VGF, ADCYAP1	1.92E-07
2	GO:0010001	glial cell differentiation	4	ASCL1, PLP1, MIAT, RELN	1.33E-04
3	GO:0009409	response to cold	4	FOS, AGT, ACOT11, VGF	0.001071
4	GO:0005179	hormone activity	5	CCK, NPY, AGT, NPPC, ADCYAP1	0.001792
5	GO:0006956	complement activation	3	C4A, C4B, C3	0.002615
6	GO:0045860	positive regulation of protein kinase activity	4	CDKN1A, RELN, ADCYAP1, CYR61	0.003602
7	GO:0043065	positive regulation of apoptotic process	7	CCK, PTGS2, DUSP1, ERBB4, BNIP3, LATS2, CYR61	0.005647
8	GO:0007218	neuropeptide signaling pathway	4	GRP, CRHR1, NPY, ADCYAP1	0.005832
9	GO:0051968	positive regulation of synaptic transmission, glutamatergic	3	PTGS2, RELN, ADCYAP1	0.005867
10	GO:0006954	inflammatory response	7	PLP1, C4A, PTGS2, C4B, C3, HDAC9, BMPR1B	0.006667

3.4 Characterisation of differentially expressed lincRNAs

3.4.1 Annotated lincRNAs: Neat1, Malat1 and Miat

Three annotated lincRNAs, Neat1, Malat1 (Neat2) and Miat (Gomafu), were identified amongst the significantly differentially expressed results across different comparisons. The Neat1 and Malat1 loci are located on mouse chromosome 19 (Fig. 10a) whilst Miat is located on chromosome 5 (Fig. 10b).

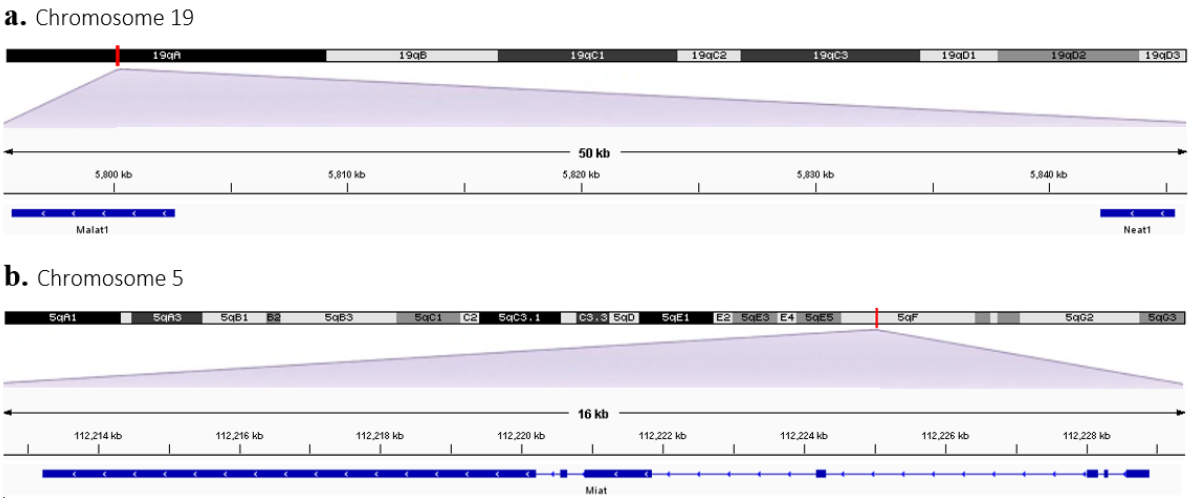

Figure 10. Schematic representations of the Malat1 and Neat1 loci on the long arm of chromosome 19 (a) and the Miat locus on the long arm of chromosome 5 (b). Figure adapted from UCSC Genome Browser and IGV displays.

Expression of Malat1 and Miat conform with expression pattern 1, with NCAM2-ED treatment of resting neurons leading to a significant downregulation and subsequent LTP induction resulting in a significant upregulation of the lincRNAs expression (Fig. 11a and b). On the other hand, expression of Neat1 appeared to be more consistent with expression pattern 3 where induction of LTP prompted a significant gene downregulation (Fig. 11c).

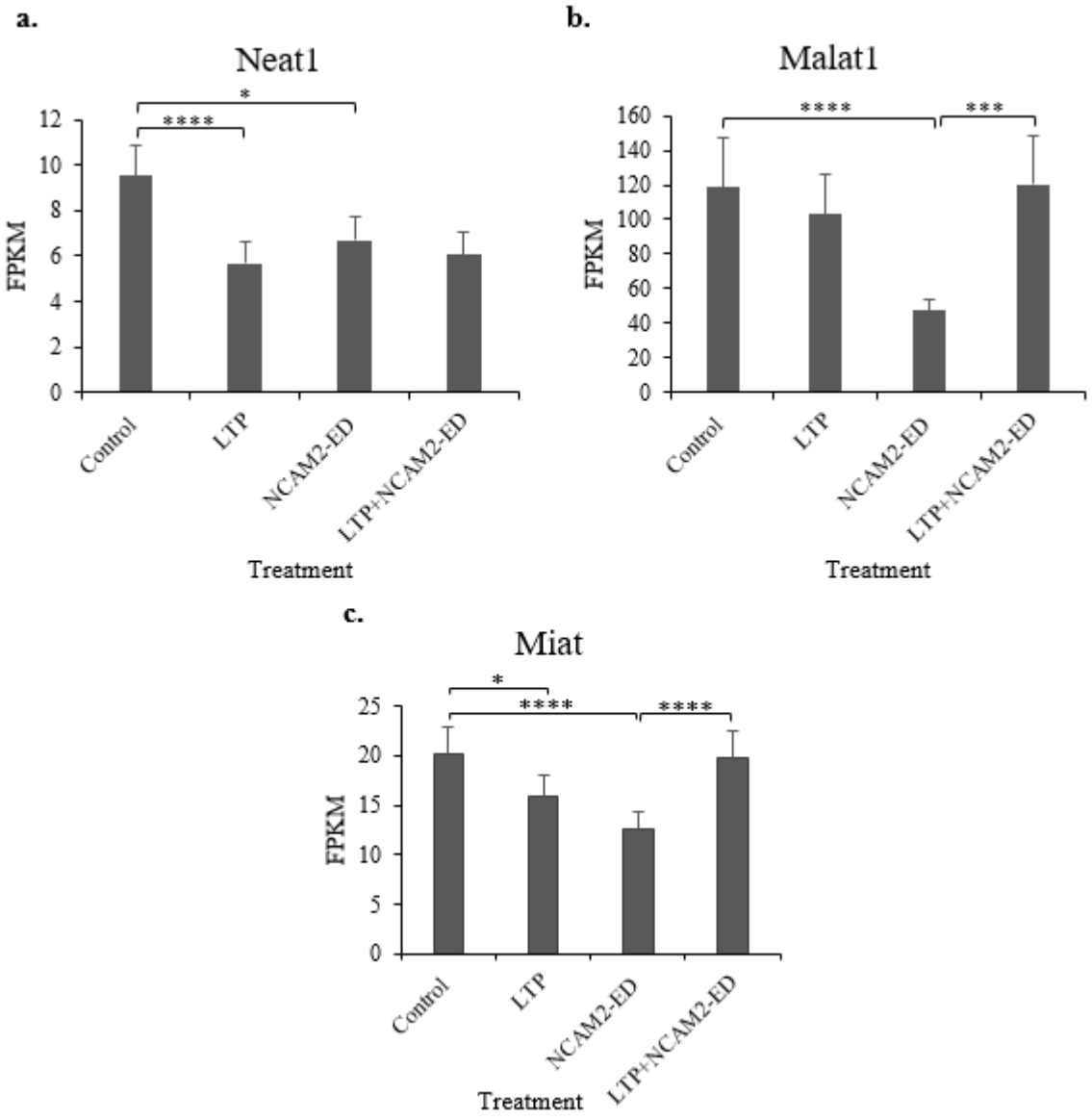

Figure 11. Comparative expression patterns of the Neat1 (a), Malat1 (b) and Miat (c) transcripts. P-values≤0.00005 are indicated by ****, ≤0.0005 by *** and ≤0.05 by *. Error bars are ±SD.

Bioinformatic analysis detected, as expressed, between 15256 and 18933 previously unannotated transcripts with an average of 16187 unannotated transcripts per sample across all conditions (Table 9). Notably only four unannotated transcripts, presented in Table 10 and described in detail in the following sections, were identified as significantly differentially expressed between the four treatment groups.

Table 9. Number of unannotated transcripts detected with expression >1 FPKM in each treatment group.

Treatment Group	Replicate	Unannotated transcripts
Control	1	16100
	2	15814
	3	15734
	mean	15883
LTP	1	15919
	2	15256
	3	16682
	mean	15952
NCAM2-ED	1	16005
	2	15644
	3	18933
	mean	16861
LTP+NCAM2-ED	1	16340
	2	15804
	3	16016
	mean	16053

Table 10. Characteristics of significantly differentially expressed unannotated transcripts.

Gene ID	**Locus**	**Size** (nt)	**Control** (FPKM)	**LTP** (FPKM)	**NCAM2-ED** (FPKM)	**LTP+ NCAM2-ED** (FPKM)
XLOC_013906	chr2:107317543-107327074	9531	1.52943	1.20061	0.400154	1.58344
XLOC_006440	chr13:111593258-111593654	396	276.561	171.549	174.343	249.481
XLOC_006256 isoform 1	chr13:84025314-84065015	754	5.88515	5.52162	1.49188	6.09361
XLOC_006256 isoform 2	chr13:84025314-84056420	603	23.6672	12.3951	14.4354	22.7182

3.4.2 XLOC_013906 transcript

The XLOC_013906 gene locus encompasses a single 9531 bp long exon located on the long arm of chromosome 2 in the intervening region between protein-coding genes Kcna4

(potassium voltage-gated channel subfamily A member 4) and Mettl15 (Methyltransferase Like 15) (Fig. 12). The XLOC_013906 locus is 19040 bp downstream of the 4844 bp long Kcna4 gene. Of note, the current analysis identified two previously unannotated isoforms of Kcna4 which are 23823 nt and 23525 nt long, respectively, reducing the gap between XLOC_013906 and Kcna4 to only 56bp, as shown in Fig. 12 . The annotated Kcna4 transcript, designated here as isoform 1, was the most abundantly expressed isoform (Fig.14). The Kcna4 gene encodes the fourth member of the shaker-related voltage-gated potassium channel family and the proximity of XLOC_013906 may indicate potential *cis* interactions between the two genes (Fig. 12).

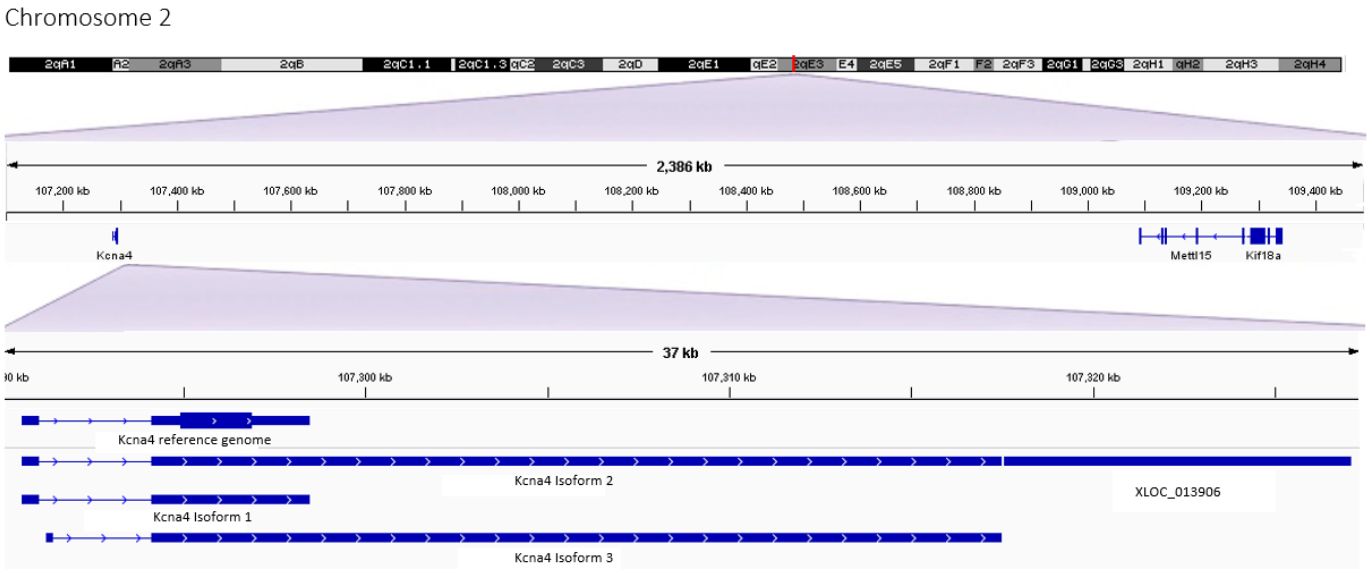

Figure 12. Schematic representation of the XLOC_013906 locus. The XLOC_013906 locus maps to the long arm of the mouse chromosome 2, downstream of the protein-coding Kcna4 and upstream of Mettl15. Figure is adapted from UCSC Genome Browser and IGV displays.

Expression of XLOC_013906 is consistent with expression pattern 1 with significant downregulation in the presence of NCAM2-ED and upregulation following induction of LTP in the presence of NCAM2-ED (Fig.13). No significant differential expression was detected for the Kcna4 gene nor any of the three individual isoforms (Fig. 14). The expression pattern of the Kcna4 gene does not appear to correlate closely with the expression pattern of XLOC_013906 (Figs.13 and 14). The expression of Kcna4 isoform 2, like XLOC_013906, conforms with expression pattern 1 however it remained at low expression levels (<1.5 FPKM) (Figs.13 and 14).

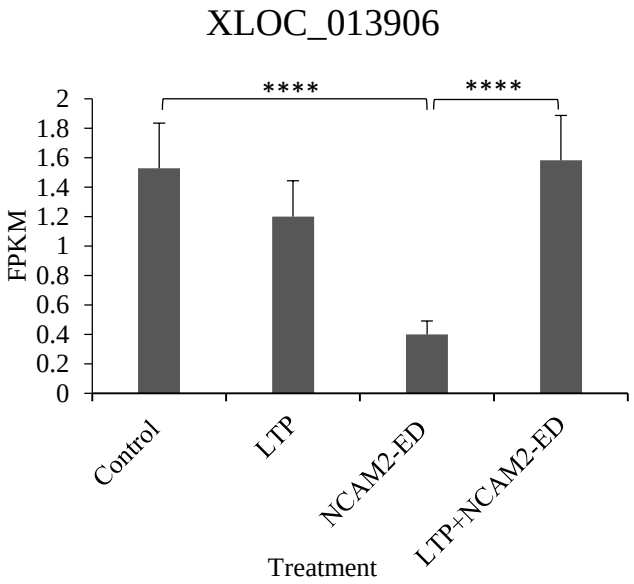

Figure 13. Expression patterns of XLOC_013906 transcripts. P-value ≤0.00005 indicated by ****. Error bars are ±SD.

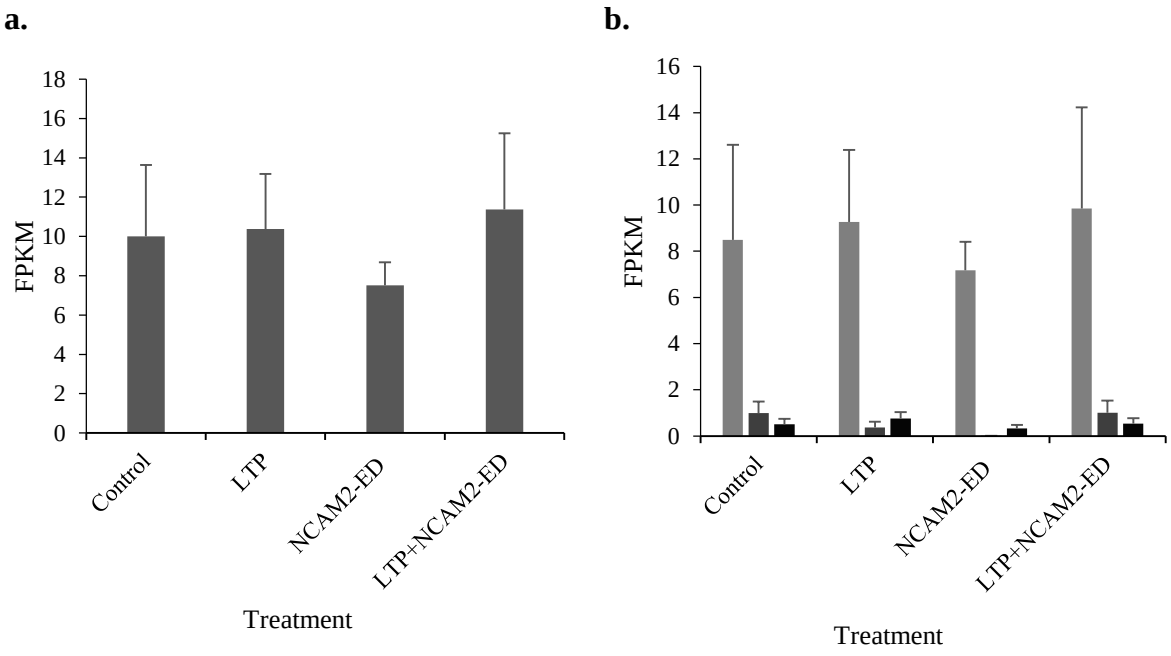

Figure 14. Expression pattern of Kcna4 (a) and individual Kcna4 isoforms (b). Isoform 1 is represented in light grey, isoform 2 in dark grey and isoform 3 in black. Error bars are ±SD.

The XLOC_013906 transcript was excluded from RNA structural prediction as, at 9531 nt long, it exceeded the maximum size limit for the RNAfold tool. The ORF finder predicted 61 potential ORFs (Fig.15 and Table S5). None of the putative ORFs was over 100 aa long thus

XLOC_013906 may still be considered a putative lincRNA (Fig. 15 and Table S5). Ribosome profiling (Ribo-Seq) sequence reads mapping to the XLOC_013906 locus suggest interactions with ribosomes potentially indicating micro-peptide coding regions within the transcript (Fig. 16). The highest peak of Ribo-Seq expression, as shown in Figure 16, corresponds to putative ORF 3 which spans 78 nucleotides of the transcript from nucleotide position 4972-5049 (Table S5). Active translation of this putative reading frame would yield a micro peptide of 25aa in length (Table S5).

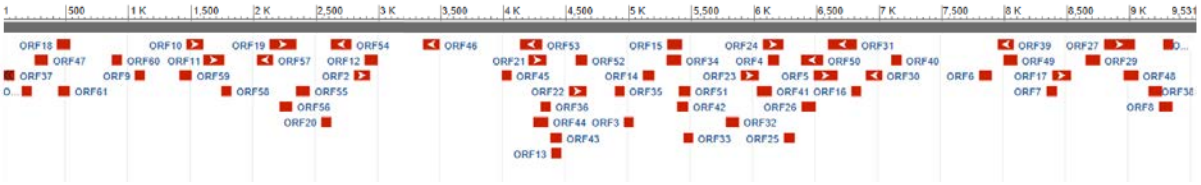

Figure 15. Location of predicted ORFs within the XLOC_013906 sequence. The grey bar represents the 9531nt long sequence (5'→ 3') whilst putative ORFs are depicted in red with the direction indicated by arrows. Figure generated using the NCBI ORF Finder.

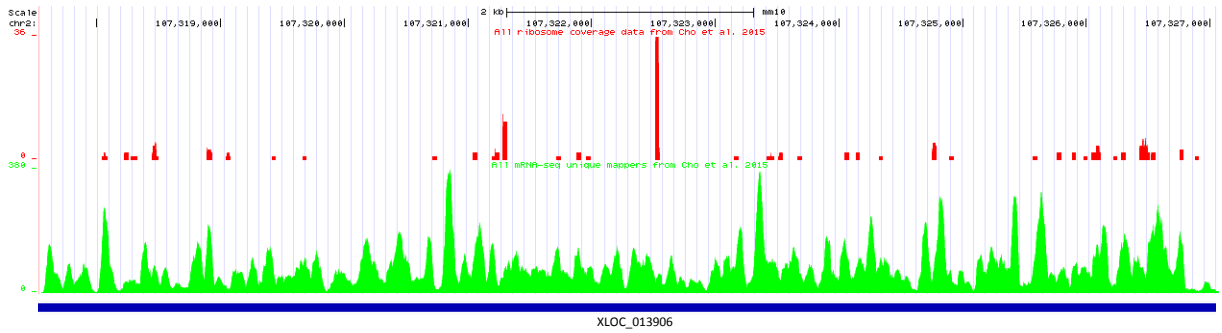

Figure 16. Ribosome binding profile of XLOC_013906 locus in mouse hippocampal tissue, obtained from the GWIPS browser. The Ribo-Seq expression data is depicted in red, RNA-Seq mRNA expression levels are portrayed in green and the XLOC_013906 transcript locus is depicted in blue.

3.4.3 XLOC_006440 transcript

The XLOC_006440 transcript is 390nt long and is transcribed from a single exon located on chromosome 13 in the intervening region between protein-coding genes Gpbp1 (GC-rich promoter binding protein 1) and Mier3 (mesoderm induction early response family member 3) (Fig.17). Following expression pattern 1, induction of LTP and treatment with NCAM2-

ED, resulted in a significant downregulation of the XLOC_006440 transcript (Fig.18). Induction of LTP combined with NCAM2-ED treatment led to upregulation of XLOC_006440 expression, which although not statistically significant, might be biologically relevant (Fig.18).

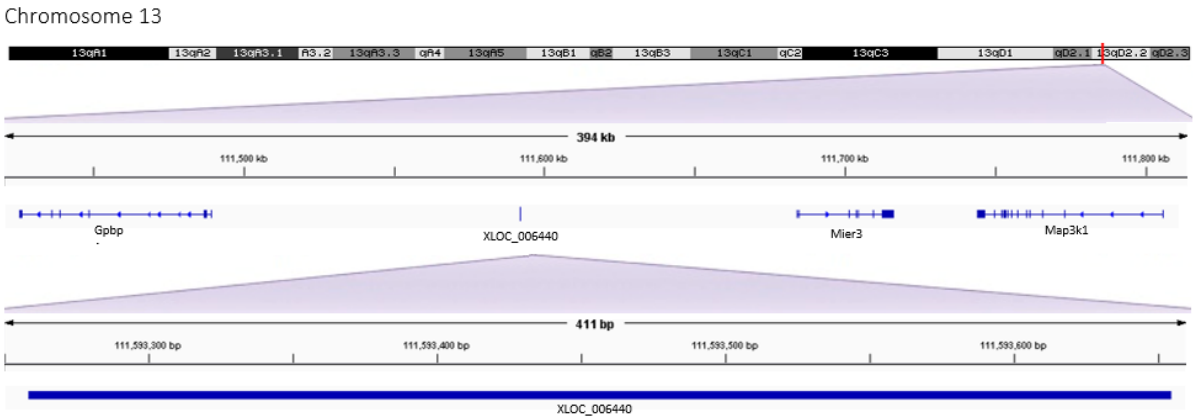

Figure 17. Schematic representation of the XLOC_006440 locus. The XLOC_006440 locus maps to the long arm of the mouse chromosome 13, in the intervening region downstream of the Gpbp and upstream of Mier3. Figure adapted from UCSC Genome Browser and IGV displays.

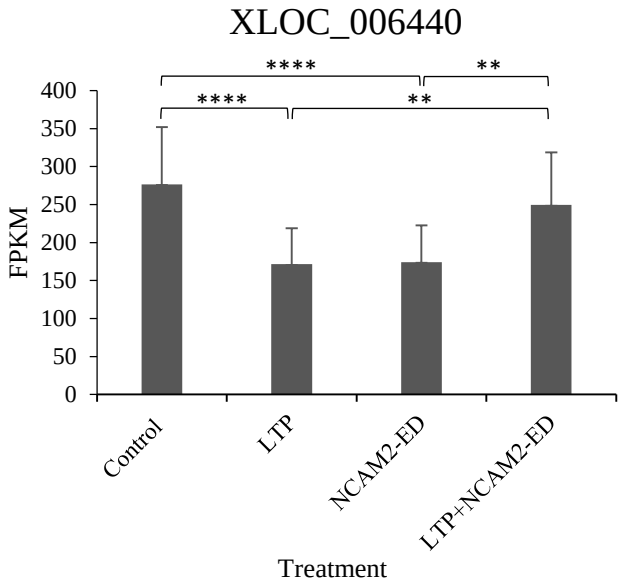

Figure 18. Expression pattern of XLOC_006440 transcript. P-value ≤0.00005 indicated by **** and ≤0.005 by **. Error bars are ±SD.

Two putative open reading frames were predicted by the ORF finder tool however neither was over 100aa long (Fig. 19 and Table 11). When imposed onto the RNA secondary structure the predicted ORFs were not flanked by single stranded RNA (ssRNA) structures as would be expected if the regions were translated (Fig. 20). Moreover, Ribo-Seq data of the XLOC_006440 region indicates that whilst the RNA transcripts are present, no ribosomal interactions were detected, thus suggesting a lack of translation (Fig.21).

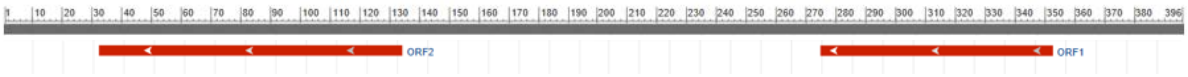

Figure 19. Location of predicted ORFs within the XLOC_006440 sequence. The grey bar represents the 396nt long sequence (5'→ 3') whilst putative open reading frames are depicted in red with the direction indicated by arrows. Figure generated using the NCBI ORF Finder.

Table 11. Predicted ORFs for XLOC_006440 sequence.

Predicted ORF	Start*	Stop*	length (nt|aa)
ORF1	353	276	78 | 25
ORF2	134	33	102 | 33

*Start designates the location of the first nucleotide in the start codon (ATG) on the XLOC_006440 gene whilst Stop denotes the last nucleotide of the stop codon. The length of each ORF is given in nucleotides (nt) as well as amino acids (aa).

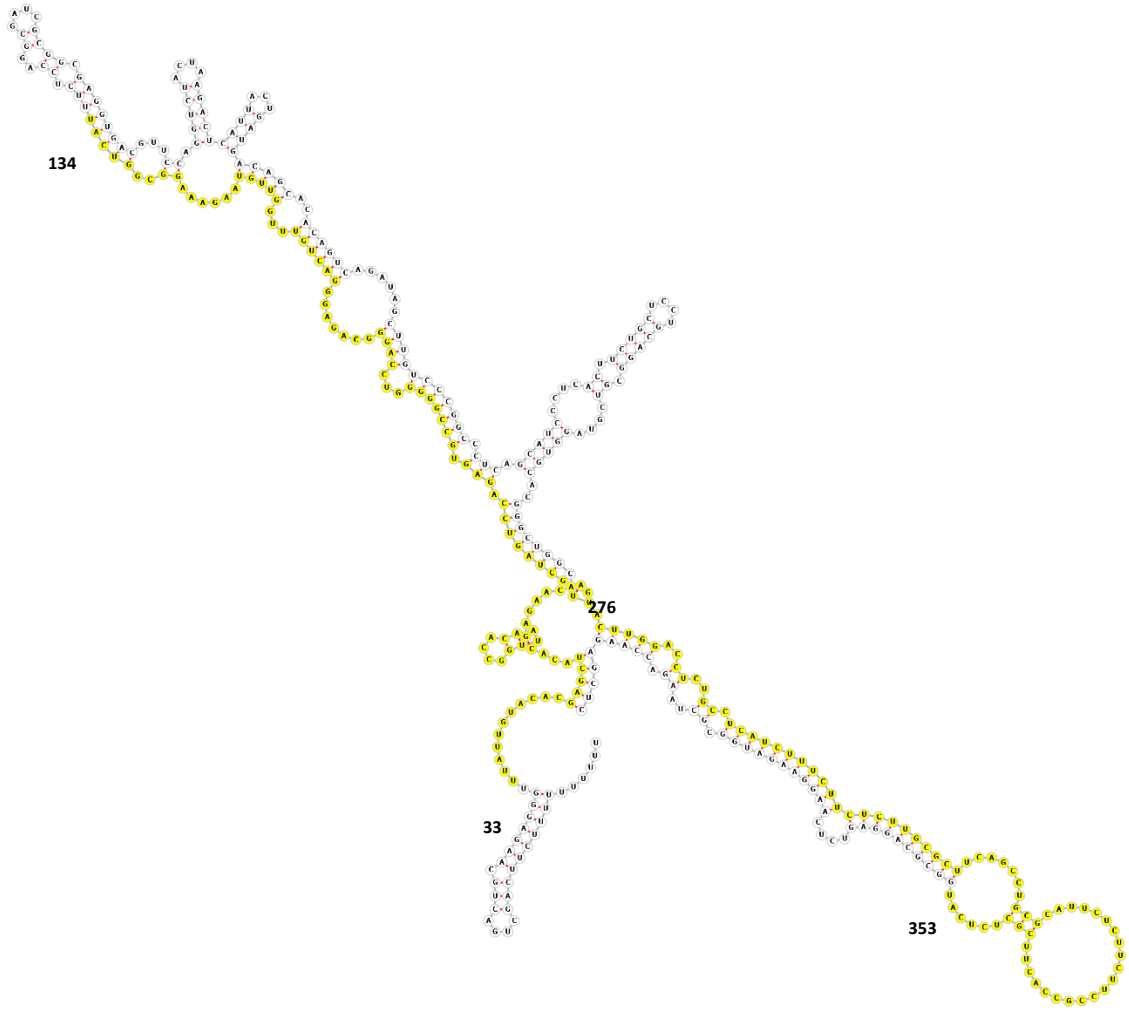

Figure 20. Predicted RNA secondary structure of XLOC_006440. Predicted ORFs are indicated in yellow with the numbers delineating the first and last nucleotide of each ORF. The ORFs are located at nucleotides positions 33-134 and 276-353.

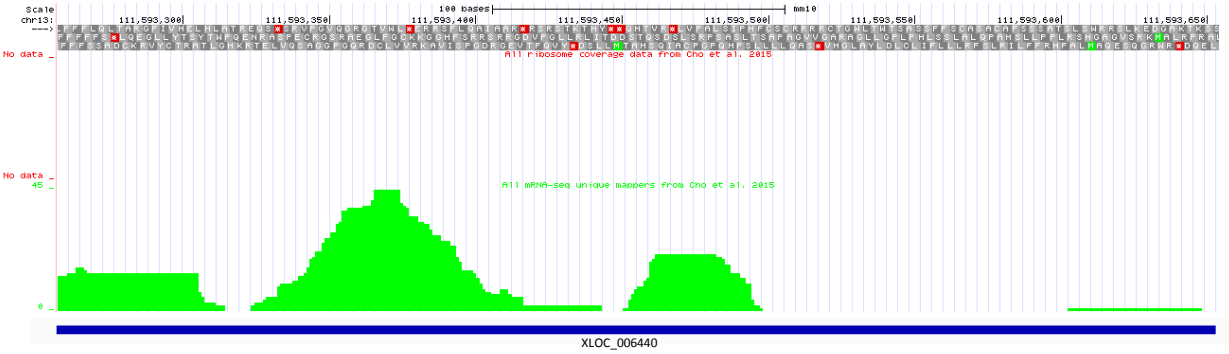

Figure 21. Ribosome binding profile of XLOC_006440 locus in mouse hippocampal tissue, obtained from the GWIPS browser. The Ribo-Seq expression data is depicted in red, mRNA expression levels are portrayed in green and the XLOC_006440 transcript locus is shown in blue.

3.4.4 XLOC_006256 transcript

The XLOC_006256 gene locus, expressing two isoforms, is located on the long arm of chromosome 13, in the intervening region between the non-coding 2810049E08Rik (RIKEN cDNA 2810049E08) gene and the potein-coding gene Tmem161b (transmembrane protein 161B) (Fig. 22). Isoform 1 of the XLOC_006256 transcript is 754nt long whilst isoform 2 comprises 603nt. The XLOC_006256 gene locus consists of three exons with exon 1 common to both isoforms whilst the second exon is unique for each isoform (Fig. 22).

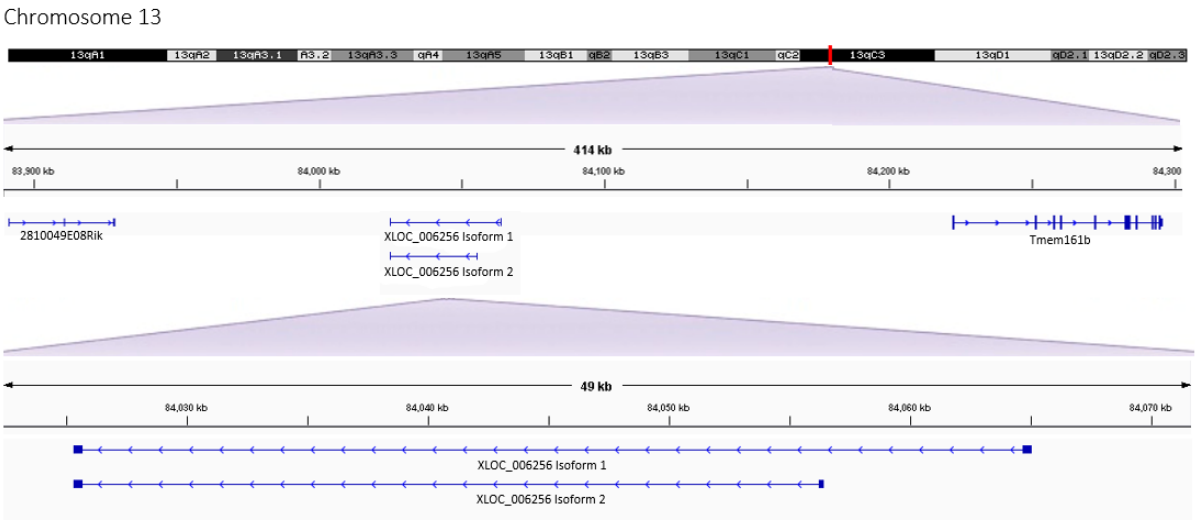

Figure 22. Schematic representation of the XLOC_006256 locus. The XLOC_006256 gene locus maps to the long arm of the mouse chromosome 13, in the intervening region downstream of the annotated nonprotein-coding gene 2810049E08Rik and upstream of protein-coding gene Tmem161b. The XLOC_006256 gene is expressed as two distinct isofroms. Figure adapted from UCSC Genome Browser and IGV displays.

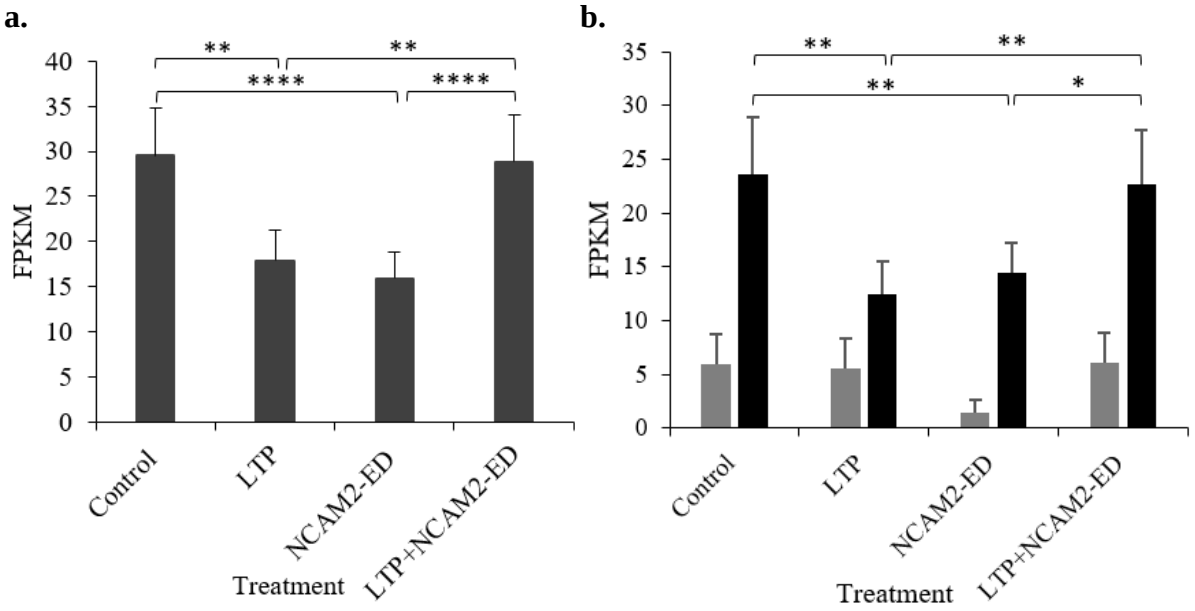

Figure 23. Cumulative XLOC_006256 gene expression pattern (a) and expression patterns of XLOC_006256 Isoform 1 (light grey) and Isoform 2 (black) (b). P-values ≤0.00005 indicated by ****, p≤0.0005 by ***, p-values≤0.005 by ** and p-values≤0.05 by *. Error bars are ±SD.

Expression profile of the XLOC_006256 gene conforms to expression pattern 1 with NCAM2-ED treatment leading to a significant downregulation of the transcript (Fig.23a). LTP induction in the presence of NCAM2-ED results in significant upregulation of the gene (Fig. 23a). Expression of isoform 1 aligns to expression pattern 2 and yielded lower FKPM levels than isoform 2 with the latter more consistent with expression pattern 1 (Fig. 23b).

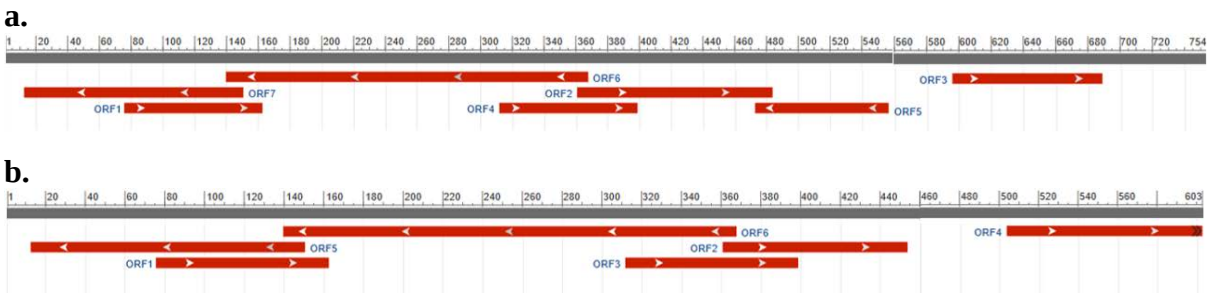

Figure 24. Location of predicted ORFs within XLOC_006256 isoform 1 (a) and isoform 2 (b) sequences. The grey bars represent the 754 and 603 nt long isoform sequences (5'→ 3') whilst putative ORFs are depicted in red with the direction indicated by arrows. The graphic representation of predicted ORFs was generated using the NCBI ORF Finder.

Table 12. Predicted ORFs for XLOC_006256 isoforms 1 and 2.

XLOC_05256 Isoform 1			
Predicted ORF	**Start***	**Stop***	**length (nt|aa)**
ORF1	76	162	87 | 28
ORF2	361	483	123 | 40
ORF3	596	688	93 | 30
ORF4	312	398	87 | 28
ORF5	556	473	84 | 27
ORF6	367	140	228 | 75
ORF7	150	13	138 | 45
XLOC_006256 Isoform 2			
ORF1	76	162	87 | 28
ORF2	361	453	93 | 30
ORF3	312	398	87 | 28
ORF4	504	>602	99 | 32
ORF5	150	13	138 | 45
ORF6	367	140	228 | 75

*Start designates the location of the first nucleotide in the start codon (ATG) on the XLOC_006256 gene; Stop denotes the last nucleotide of the stop codon. The length of each ORF is given in nucleotides (nt) and amino acids (aa).

The ORF finder tool predicted seven ORFs for XLOC_006256 isoform 1 and six for isoform 2 (Fig. 24 and Table 12). None of the predicted ORFs were longer than 100aa. RNA secondary structure predictions identified multiple ORFs flanked by ssRNA structures within both isoforms (Fig. 25 and 26). Furthermore, Ribo-Seq data identified multiple sites of ribosomal binding corresponding with ORFs in each of the XLOC_006256 exons (Fig. 27). Ribosomal binding was identified at loci corresponding to ORF2, ORF4, ORF5 and ORF6 of XLOC_006256 isoform 1, each of which was flanked by ssRNA structures (Figs. 25 and 27). Isoform 2 of XLOC_006256 exhibited ribosomal binding activity corresponding to ORF3 and ORF6 (Figs. 26 and 27).

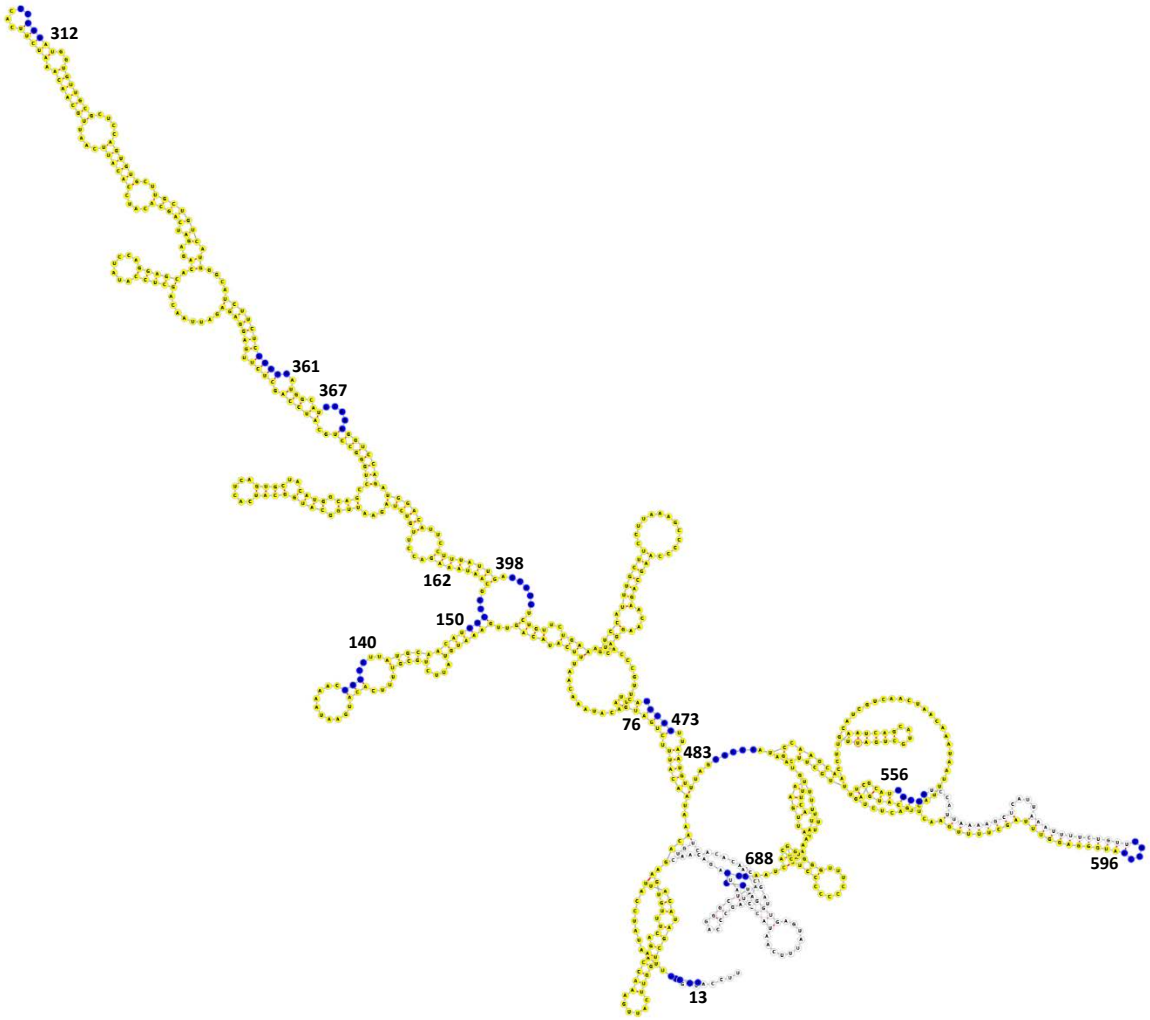

Figure 25. Predicted RNA secondary structure of XLOC_006256 isoform 1. Predicted ORFs are indicated in yellow whilst the regions marked in blue designate ssRNA structures which may facilitate ribosome binding. The ORFs are located at nucleotides positions 76-162, 361-483, 596-688, 312-398, 556-473, 367-140 and 150-13, respectively.

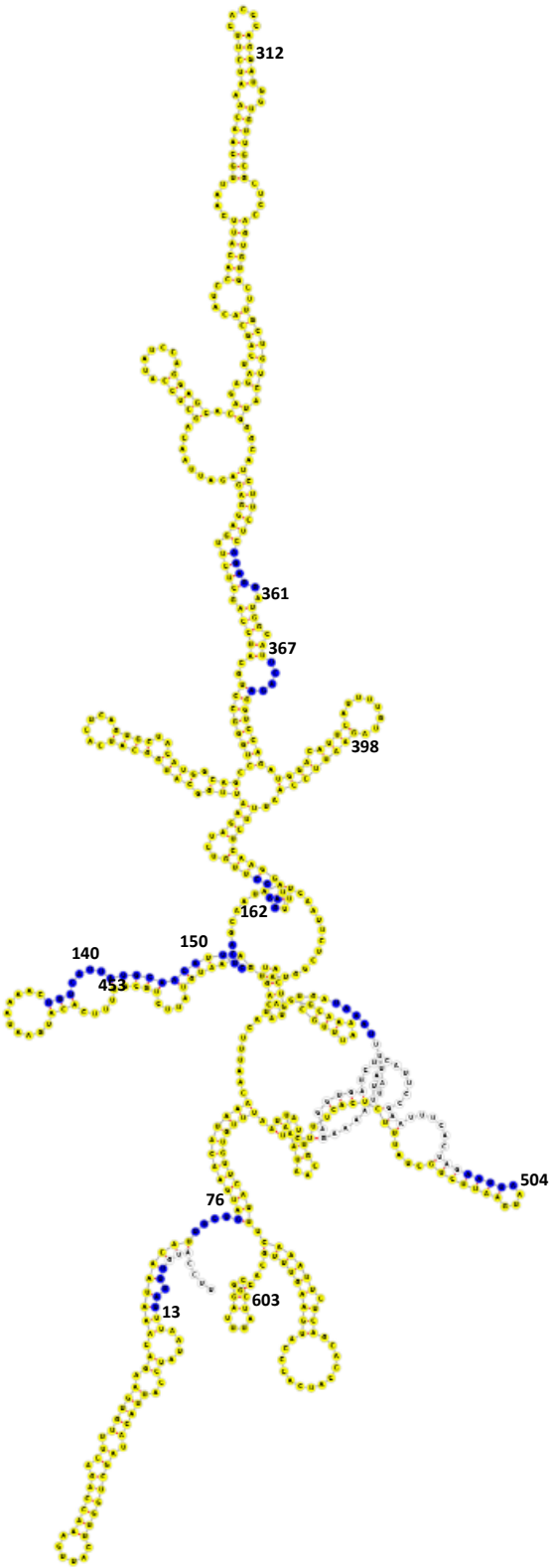

Figure 26. Predicted RNA secondary structure of XLOC_006256 isoform 2. Predicted ORFs are indicated in yellow whilst the regions marked in blue designate ssRNA structures which may facilitate ribosome binding. The ORFs are located at nucleotides positions 76-162, 361-453, 312-398, 504->602, 150-13 and 367-140, respectively.

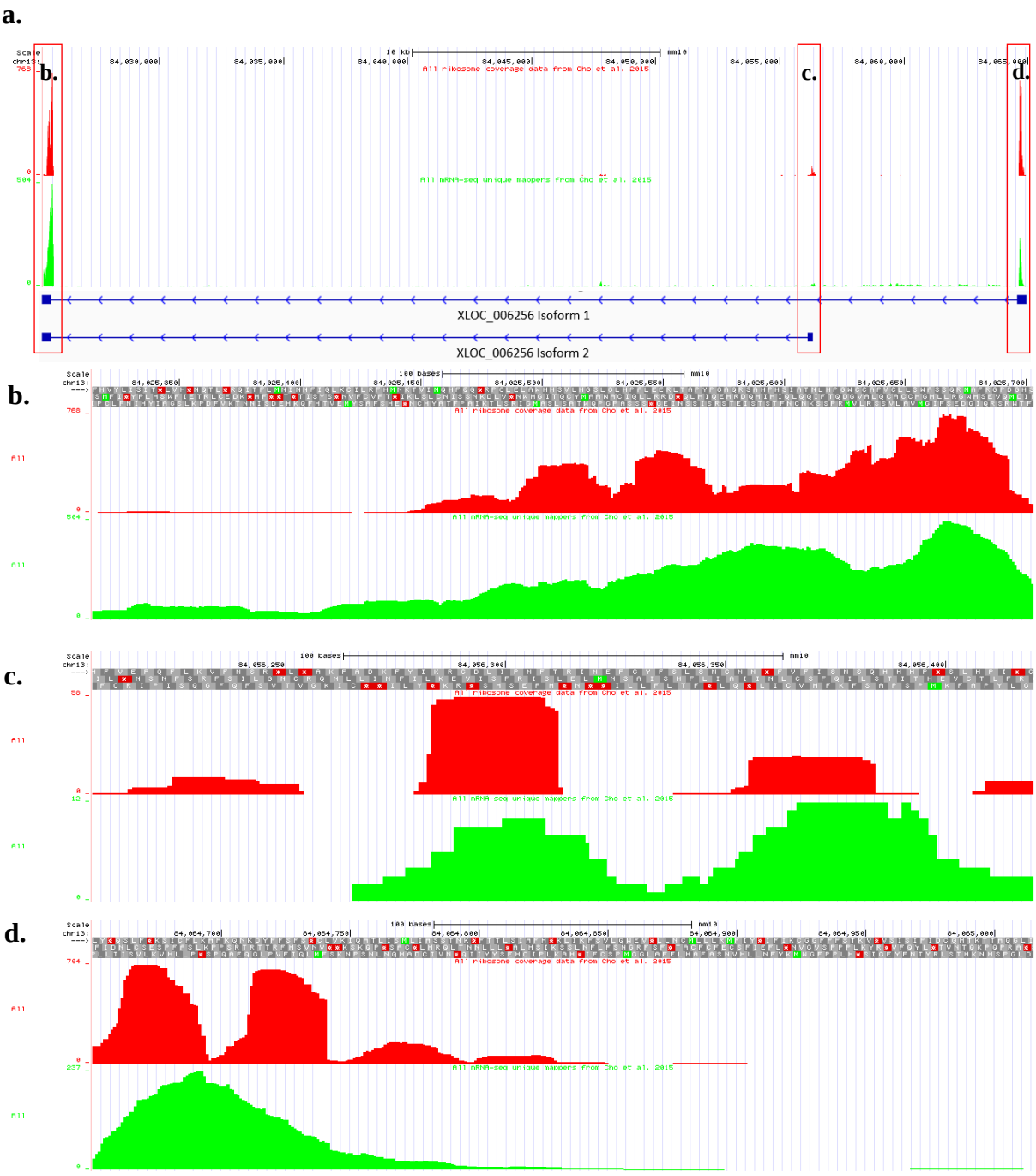

Figure 27. Ribosome binding profile of XLOC_006256 locus in mouse hippocampal tissue (a) with closer views of common exon 1 (b), exon 2 (c) and exon 3 (d). The Ribo-Seq expression data is depicted in red, mRNA expression levels are portrayed in green and the XLOC_006256 transcript locus is depicted in blue.

In summary, all four significantly differentially expressed transcripts, listed in Table 10, are (i) over 200 nt long, (ii) located in the intervening regions between annotated protein-coding genes and (iii) exhibited putative ORFs less than 100 aa long. Therefore, the transcripts can

be considered as putative lincRNAs (Hangauer *et al*. 2013). Ribosomal binding was detected for XLOC_013906 and XLOC_006256 transcripts indicating their potential for expression of micropeptides (Fig. 16 and 27).

3.5 Meta-analysis of rat hippocampal transcriptome data

Bioinformatic analysis detected, as expressed, 29425 genes and 66736 isoforms in combined data sets representing naïve hippocampal rat tissue and hippocampal rat tissue 30 mins after induction of LTP. Comparative analysis of these two RNA-Seq data sets identified 557 genes and 353 isoforms as significantly differentially expressed ($p<0.00145$ and $p<0.0005$ respectively) (Table S6). Of the 557 differentially expressed genes, 153 genes lack annotation to the rat reference genome, thus may potentially represent novel lncRNAs.

Comparison of the rat differentially expressed genes with the genes differentially expressed in the mouse Control vs. LTP pairwise comparison identified eight mouse genes overlapping with the rat data set (Table 13). In contrast, the Neat1 and unannotated XLOC_006440 transcripts were not identified as expressed in the rat data set. Of the eight remaining mouse genes, none was found to be significantly differentially expressed in the rat data set as shown in Table 13. Furthermore, most rat gene expression levels increased in response to LTP, although not significantly, which contrasted with decreased expression of mouse orthologs in the same condition.

Table 13. Comparison of mouse and rat expression levels for eight genes that were significantly differentially expressed in the mouse Control vs. LTP comparison.

Gene ID	Mouse*		Rat*		p-value**
	Control	LTP	Control	LTP	
Ppm1h	2235.1	20.8388	25.4134	24.7499	0.74855
Ntsr2	21.9495	10.5885	86.733	90.8759	0.672
Cldn10	10.2569	4.89505	18.9142	19.2541	0.84935
Adcyap1	4.20938	1.80734	6.87105	10.1178	0.00805
Rpl22l1	26.9999	13.5049	17.6743	20.2528	0.3084
Id3	48.5083	30.1312	14.5908	16.1118	0.34625
Padi2	5.49414	3.17117	7.32114	7.36753	0.94095
Lcat	8.99151	3.99782	5.44879	5.97247	0.4458
Bgn	10.2591	5.10267	1.22348	1.23423	0.95555

*in FPKM; ** ‘p-value’ refers to the differential expression analysis results of the rat data sets.

4. Discussion

Mechanisms of synaptic plasticity, specifically LTP, are believed to underlie the formation and maintenance of memory. By comparing the transcriptomes of the resting control neurons and neurons in which LTP has been induced it may be possible to identify changes in transcript expression, elucidating the underlying mechanisms of this crucial process. This project aimed to comprehensively analyse the transcriptomes of hippocampal neurons in order to identify which transcripts are differentially expressed following the induction of LTP and treatment with NCAM2-ED. Hippocampal neurons were treated with the NCAM2-ED molecule to disrupt synaptic adhesion and trigger synapse disassembly, simulating the AD phenotype (Leshchyns'ka *et al*. 2015). Changes in the transcriptomes of neurons treated with NCAM2-ED may provide insight into the mechanisms underlying the pathology of neurodegenerative disorders such as AD.

4.1 Expression pattern of significantly differentially expressed protein-coding genes

4.1.1 Expression pattern 1

Characterised by downregulation of gene expression following induction of LTP or treatment with NCAM2-ED, expression pattern 1 represents 31 significantly differentially expressed genes. In each of these genes, induction of LTP in the presence of NCAM2-ED led to upregulation of gene expression (Fig. 1).

In this study, expression of Grp gene appeared consistent with expression pattern 1. Gastrin releasing peptide (Grp) is a neuropeptide that interacts with G-protein coupled Grp receptors on the surface of neurons to mediate intracellular signalling involved in LTP and memory (Roesler and Schwartsmann 2012). Activation of Grp receptors by Grp propagates intracellular signal cascades that trigger the PKC, ERK and MAPK signalling pathways which have clearly established roles in the mechanisms of LTP (Chen and Kroog 2004; Hellmich et al. 1999; Roesler and Schwartsmann 2012). The Grp receptor interacts with numerous growth factors and IEGs including Fos (Chatzistamou et al. 2000; Roesler and Schwartsmann 2012). Following neuronal activity, presynaptic neurons excrete Grp which binds to Grp receptors on the postsynaptic neurons triggering intracellular signalling cascades that facilitate remodelling of the synapse during LTP (Lee et al. 1999; Shumyatsky et al. 2002). The Grp receptors are expressed in high densities in the neuronal cell bodies and

dendrites of the mammalian hippocampus indicating a critical role for Grp and its receptor in manifestation of synaptic transmission and plasticity (Kamichi et al. 2005; Wolf and Moody 1985). In the CA1 neurons of the hippocampus treatment with bombesin, the amphibian homolog of Grp, resulted in enhanced memory formation whereas treatment with Grp receptor antagonists or inhibitors of PKC and MAPK impaired memory consolidation (Roesler et al. 2006; Roesler and Schwartsmann 2012). Whilst low doses of bombesin lead to enhanced memory consolidation, higher doses impaired consolidation of memory (Roesler and Schwartsmann 2012). A significant downregulation of Grp was observed following treatment of neurons with NCAM2-ED. In contrast, induction of LTP in the presence of NCAM2-ED led to significant upregulation of the gene. Although not statistically significant, Grp was downregulated following induction of LTP which is consistent with previous observations that lower levels of bombesin/Grp facilitate enhanced memory consolidation.

Ppm1h (protein phosphatase, Mg2+/Mn2+ Dependent 1H), also known as NERPP (Neurite Extension-related Protein Phosophatase Related to PP2C), is a cytoplasmic protein localised to neurites where it plays a role in regulating their growth (Labes *et al*. 1998). Expressed most highly in a subset of neurons, including those of the hippocampus, Ppm1h is part of an intracellular signalling pathway that impedes transcription of genes involved in neurite growth in response to myelin-associated inhibition factors (Labes et al. 1998). Downregulation of Ppm1h results in reduced sensitivity to myelin-associated inhibitors of neuronal outgrowth and a significant increase in neurite outgrowth (Labes et al. 1998). In the present study Ppm1h appears consistent with expression pattern 1, exhibiting a significant downregulation in response LTP induction facilitating expression of proteins associated with neurite outgrowth. In the presence of NCAM2-ED, Ppm1h expression is significantly downregulated almost ~100 fold. On the other hand, induction of LTP in the presence of NCAM2-ED resulted in significant upregulation of the gene, rivalling the levels observed in resting neurons.

4.1.2 Expression pattern 2

In expression pattern 2, LTP induction did not significantly alter gene expression however application of NCAM2 led to significant downregulation of expression (Fig. 2). Induction of LTP in the presence of NCAM2-ED however resulted in upregulation of the gene expression, restoring FPKM levels to those observed in resting neurons in the absence of NCAM2-ED.

Eight genes conform to expression pattern 2, downregulated in response to NCAM2-ED treatment, potentially indicating a role for these genes in AD pathology. Further experimentation is necessary to establish the cause and implications of the observed expression pattern.

Retinoid X receptor gamma (Rxrg or Rxrγ) conforms with expression pattern 2, demonstrating a significant downregulation in response to NCAM2-ED treatment, alleviated with the induction of LTP. A component of the retinoid signalling pathway, Rxrg plays a crucial role in hippocampal synaptic plasticity, mediating transcriptional events necessary for synaptic remodelling (Nomoto et al. 2012). By binding to retinoic acid receptors or other retenoid X receptors, Rxrg forms complexes that act as transcription factors, regulating expression of target genes (Chambon 1996; Mangelsdorf et al. 1995). The exact role of Rxrg in mechanisms of synaptic plasticity is unclear however a study of Rxrg deficient mice found that no change was observed in LTP expression (Chiang et al. 1998). Consistent with these previous finding LTP induction had no significant effect on Rxrg expression. Treatment with NCAM2-ED however led to a significant downregulation in Rxrg expression which was alleviated with induction of LTP.

4.1.3 Expression pattern 3

The 18 genes representing expression pattern 3 are characterised by downregulation in response to induction of LTP (Fig. 3). Whilst the presence of NCAM2-ED had some effect on gene expression the induction of LTP consistently led to downregulation of gene expression. Downregulation of genes expression in response to LTP induction may indicate alleviation of inhibitory regulation of synaptic plasticity.

Neurotensin receptor subtype 2 (Ntsr2) appears consistent with expression pattern 3, LTP induction leading to a significant downregulation. Ntsr2 is a G-protein coupled receptor for the neuropeptide neurotensin which modulates dopamine transmission in the brain (Mazella et al. 1996; Mazella and Vincent 2006; Sun et al. 2001; Yamauchi et al. 2007). It is expressed diffusely throughout the brain including in the cortex, thalamus, hypothalamus, hippocampus and brainstem (Lépée-Lorgeoux et al. 1999; Maeno et al. 2004; Mazella et al. 1996). Although most highly expressed in astrocytes Ntsr2 is also expressed in neurons and its role in either cell type has yet to be extensively characterised (Yamauchi et al. 2007). Neurotensin

activity is involved in cell proliferation, hypothermia hypotension, analgesia, nociception, memory and learning (Maeno et al. 2004; Yamauchi et al. 2007). The Ntsr2 protein has been implicated in the consolidation of fear learning and memory formation (Yamauchi et al. 2007). Mice deficient in Ntrs2 exhibit altered emotional behaviour and abnormal response to thermal pain during fear conditioning (Furuta et al. 2007; Maeno et al. 2004). Neurotensin has been shown to heighten glutamate release and alter glutamate receptor activity (Yamauchi et al. 2007). Therefore interactions between glutamate, neurotensin and Ntsr2 may play a role in modulating synaptic transmission (Yamauchi et al. 2007). Activation of Ntsr2 initiates signal cascades that elevate intracellular calcium levels (Mazella et al. 2012). A role for Ntsr2 in hippocampal LTP has yet to be established however in the present study it is significantly downregulated in response to LTP induction.

4.1.4 Expression pattern 4

Expression pattern 4 represents genes upregulated in the presence of NCAM2-ED (Fig. 4). No statistically significant change in gene expression was observed in response to LTP induction. The immediate early response 3 (Ier3) gene is consistent with expression pattern 4, potentially indicating a role in AD pathology. This early response gene has been implicated in regulation of cell viability, cell cycle regulation and inhibition of apoptosis (Arlt and Schäfer 2011; Garcia et al. 2002; Wu et al. 1998). Further experimentation is necessary to establish the role of Ier3 in the hippocampus however the observed expression pattern, with upregulation in response to NCAM2-ED, indicated a potential role in AD pathology.

4.1.5 Fos Expression

Whilst Fos expression did not match any of the four observed gene expression patterns, it was consistent with previously published data in which upregulation was observed 30-60 mins following LTP induction (Kovács 1998). In the present study, although not significant, an increase in Fos expression was observed 30 mins after induction of LTP. Furthermore, a significant upregulation was observed when LTP was expressed in the presence of NCAM2-ED. The IEG Fos is a transcription factor expressed during LTP induction that interacts with the AP-1 binding site of its target genes however the transcriptional targets of Fos have yet to be comprehensively characterised (Kovács 1998).

4.2 Gene ontology analysis of significantly differentially expressed genes

Gene ontology enrichment analysis of the top 100 differentially expressed genes for each pairwise comparison identified numerous enriched GO terms involved in neuronal physiology and restructuring including ‘axon guidance’, ‘cell adhesion’, ‘regulation of transcription’, ‘regulation of translation’, and intracellular signalling pathways such as the JAK-STAT and ERK1/ERK2 cascades. Amongst the top GO results were several GO terms not obviously related to neuronal physiology and LTP including ‘response to gamma radiation’, ‘maternal process involved in female pregnancy’ and ‘heart development’. These results likely indicate alternative roles for genes initially annotated in other physiological pathways. The role of proteins may differ depending on the tissue, cell type and subcellular domains in which they are expressed and as such the initial annotation of certain genes may not reflect all their pleiotropic functions (Boulanger 2009). Alternatively, the presence of GO terms relating to pregnancy and embryonic development (fig. 6-9, table S4) may also be due to the use of neonate mouse neurons.

Amongst the top enrichment clusters identified for the Control vs. LTP and NCAM2ED vs. LTP+NCAM2-ED comparisons were GO terms related to immune function, inflammatory response and the complement cascade. In mature brains, the blood-brain barrier limits access of blood circulating immune cells and proteins to the brain, making it immune privileged. It has been suggested that the blood-brain barrier of juvenile animals may not be fully established, potentially allowing blood-borne immune cells and molecules to ‘leak’ into neural tissues however a recent review by Saunders et al. (2014) calls into question the ‘leaky’ nature of the neonatal blood-brain barrier. Hence, the presence of these genes in hippocampal neurons may reflect localised expression rather than contamination from immune competent cells (Boulanger 2009; Veerhuis et al. 2011). Many genes originally identified in the immune system have since been found to be expressed in spatiotemporally regulated manner in healthy neuronal tissue, performing non-immune related roles (Boulanger 2009; Garay and McAllister 2010). Constituents of the immune system with distinct roles in the nervous system include cytokines, components of the complement system and MHC class 1 proteins (Garay and McAllister 2010). Immune related proteins and signalling pathways have been shown to positively regulate synaptic plasticity and remodelling of neural circuits (Donzis and Tronson 2014; Garay and McAllister 2010; Yirmiya and Goshen 2011). The role of immune related genes is well described in their

original context, however their function in the nervous system is just beginning to be characterised (Boulanger 2009; Veerhuis et al. 2011).

Clusters relating to immune function and inflammatory response were amongst the top GO terms identified in this study. Many of inflammatory cytokines are expressed in healthy brain with expression patterns specific for individual cerebral regions (Bauer et al. 2007; Garay and McAllister 2010; Pousset 1994). In the hippocampus, inflammatory cytokine interleukin-1 (IL-1) is expressed within 15 min of LTP induction and elevated IL-1 expression persists long after induction as it is crucial for the maintenance of LTP (Schneider et al. 1998; Yirmiya and Goshen 2011). Sequestration of IL-1 receptors with an interleukin-1 receptor antagonist (IL-1ra) hinders the maintenance phase of LTP however does not alter the induction of LTP (Schneider et al. 1998). Cytokines function in complex signalling networks thus characterisation of the intracellular signalling mechanisms involved in synaptic plasticity requires a network wide view rather than a focus on individual cytokines (Donzis and Tronson 2014; Garay and McAllister 2010).

Amongst the top GO clusters were GO terms relating to components of the complement cascade. The complement system is a major constituent of the innate immune system, critical for discriminating between self and non-self cells, recognising and neutralising pathogens as well as detrimental cellular debris (Boulanger 2009; Ricklin et al. 2010; Stephan et al. 2012). Virtually all of the over 30 components of the complement can be locally expressed in the brain and many components are expressed at low levels in the healthy brain (Cahoy et al. 2008; Lucin and Wyss-Coray 2009; Stevens et al. 2007; Veerhuis et al. 2011). Components of the complement cascade, C1q and C3, have been shown to play a role in activity-dependent synaptic plasticity with knockdown of either protein impairing synaptic remodelling (Boulanger 2009; Garay and McAllister 2010; Stevens et al. 2007). These complement components appear to mark weak or incongruous synapses for elimination, a role analogous to that played by these proteins in the immune system (Stephan et al. 2012). In vitro studies have also shown C1q to influence neurite outgrowth (Benoit and Tenner 2011; Pisalyaput and Tenner 2008). The C1q1 protein has also been shown to interact with the cell-adhesion G protein-coupled receptor BAI3 which is involved in regulating the formation and maintenance of synapses (Bolliger et al. 2011). When the complement cascade is fully activated in the immune system it forms the membrane attack complex (MAC or C5b-9) which lyses the membrane of target cells. However sublytic quantities of the C5b-9 complex

may prompt C^{2+} influx, inducing neuron activation rather than cell death (Cole and Morgan 2003; Veerhuis et al. 2011).

Expression of complement components has been found to be elevated in AD-neurodegenerating neurons and it is believed to play a role in the pathology of the disease (Pisalyaput and Tenner 2008; Stephan et al. 2012; Veerhuis et al. 2011). The expression of C1q may be upregulated as much as 80-fold in AD brains. Further, elevated levels of numerous complement components may be detected in the cerebrospinal fluid of presymptomatic carriers of the AD gene mutation (Ringman et al. 2012; Yasojima et al. 1999). The binding of Aβ protein plaques to C1q has been found to stimulate the classical complement cascade which may in turn propel disease pathology through synapse elimination and neuron loss (Sim et al. 2007; Tacnet-Delorme et al. 2001)

Immune proteins and signalling pathways are crucial for the development of the nervous system, its remodelling and function (Boulanger 2009; Garay and McAllister 2010). Relatively few immune molecules have been extensively studied in the brain and beyond expression patterns little is known of the mechanisms of their involvement in synaptic plasticity (Boulanger 2009; Garay and McAllister 2010).

4.3 Characterisation of lincRNA expression

4.3.1 Annotated lincRNAs

Three of the most extensively characterised annotated lincRNAs, Neat1, Malat1 (Neat2) and Miat (Gomafu), were identified amongst the significantly differentially expressed genes in this study. Expression of Malat1 and Miat conform to expression pattern 1 with NCAM2-ED treatment of resting neurons leading to a significant downregulation and subsequent LTP induction manifested by a significant upregulation of the lincRNAs expression (Fig. 11a and b).

The lincRNA Neat1 (nuclear enriched abundant transcript 1) is an ~3kb long gene located on the long arm of mouse chromosome 19 (Fig. 10a) (Ip and Nakagawa 2012). The Neat1 transcript is an essential structural component involved in the formation of paraspeckles (Clemson *et al*. 2009; Hutchinson *et al*. 2007). Paraspeckles are the nuclear subdomains located on the peripheries of nuclear speckles (SC35 splicing domains) and are believed to be

involved in regulation of gene expression and alternative splicing through coordination of mRNA nuclear retention (Clemson *et al*. 2009; Fox *et al*. 2005). Neat1 plays an important role in regulating gene expression by coordinating the three DNA/RNA binding paraspeckle-associated proteins, PSP1, PSF (PSFQ) and p54 (NonA), which modulate the export of mRNAs from the nucleus to the cytoplasm where they may be translated by ribosomes (Chen and Carmichael 2009; Clemson et al. 2009; Kozlova et al. 2006; Scadden 2009; Xie et al. 2006). The Neat1 transcript is essential for localisation of these nucleic-acid-binding proteins to the paraspeckle and Neat1 depletion results in loss of paraspeckle domains while its overexpression results in an increase in paraspeckle numbers (Chen and Carmichael 2009; Clemson *et al*. 2009).

Upregulation of Neat1 has been observed during neuronal differentiation (Clark and Mattick 2011). Expression of Neat1 results in the sequestration of mRNA to the paraspeckles, suspending translation (Chen and Carmichael 2009). Downregulation of Neat1 leads to mRNA export to cytoplasm ad induction of its translation (Chen and Carmichael 2009). Barry *et al*. (2017) recently reported a transient downregulation of Neat1 following neuronal activity. Consistent with their findings, in this study a significant downregulation of Neat1 was observed in response to induction of LTP (Fig.11). This expression pattern is indicative of nuclear export and translation of mRNAs in response to LTP induction, resulting in dynamic regulation of specific genes. Further studies may elucidate which specific genes are affected by the Neat1 regulatory mechanism. Interestingly, in the current study, Neat1 was downregulated in response to NCAM2-ED treatment, thus ablating the transcript's inhibitory effect on nuclear export and translation (Fig.11).

In the current study expression of Malat1 was consistent with expression pattern 1. The Malat1 (Metastasis-associated lung adenocarcinoma transcript 1) lincRNA transcript, also known as Neat2 (nuclear enriched abundant transcript 2), was first characterised in cancerous cells and has since been observed to be expressed in a range of healthy human tissues (Bernard *et al*. 2010). The Malat1 lincRNA localises to the nuclear speckles (SC35 splicing domains), a discrete nuclear subdomain that houses proteins and RNAs critical to the metabolism, splicing and export of pre-mRNA (Bernard *et al*. 2010; Clemson *et al*. 2009; Hutchinson *et al*. 2007; Lamond and Spector 2003; Tripathi *et al*. 2010). Each cell nucleus contains ~20-30 nuclear speckles which are enriched with SR (serine/arginine-rich) splicing factors proteins, poly(A) RNA and mRNA export factors, components crucial to the

processing of pre-mRNA into mature mRNA (Clemson *et al*. 2009; Hutchinson *et al*. 2007; Lamond and Spector 2003; Tripathi *et al*. 2010). The RNA-binding SR proteins facilitate constitutive and alternative splicing by expediting assembly of the spliceosome by recruiting other splicing factors (Lin and Fu 2006; Long and Caceres 2009). Splicing does not occur at nuclear speckles but rather they are the site for assembly, modification and storage of splicing factors (Tripathi *et al*. 2010). Expressed during transcription events, Malat1 modulates alternative splicing of mRNA by sequestering SR splicing factor proteins (SRSF1, SRSF2 and SRP20) to the nuclear speckles as well as recruiting them to transcription sites (Änkö and Neugebauer 2010; Barry *et al*. 2014; Bernard *et al*. 2010; Ip and Nakagawa 2012; Lamond and Spector 2003; Tripathi *et al*. 2010). In this way Malat1 manages the functional levels of SR factor proteins that facilitate the splicing of pre-mRNAs into functional mRNA transcripts, post-transcriptionally regulating expression of a subset of genes (Bernard *et al*. 2010; Lamond and Spector 2003; Tripathi *et al*. 2010). There is an important role as approximately 95% of multi-exonic genes in humans undergo alternative splicing, a process which is vital for the diversification and regulation of gene function, expanding the repertoire of protein coding genes (Tripathi *et al*. 2010).

Abundantly expressed in neuronal tissues of the adult brain, Malat1 has been implicated in nervous system development (Bernard *et al*. 2010; Briggs *et al*. 2015; Ip and Nakagawa 2012). Malat1 expression is restricted to a subset of neurons including those of the hippocampus (Mercer *et al*. 2008b). Through interactions with components of the nuclear speckle, Malat1 regulates the expression of a subset of genes associated with synaptogenesis and maintenance of synapses (Barry *et al*. 2014; Bernard *et al*. 2010; Briggs *et al*. 2015; Clark and Mattick 2011). Knockdown studies of Malat1 has been found to alter expression levels of genes associated with synapse formation, dendrite development, extracellular matrix, cytoskeletal components, cell motility, nuclear function and organisation (Bernard *et al*. 2010; Tano *et al*. 2010). Depletion of Malat1 transcripts results in diminished recruitment pre-mRNA splicing factors to the paraspeckle, inhibiting correct pre-mRNA splicing and downregulating genes associated with nuclear organisation, synapse formation and maintenance (Bernard *et al*. 2010; Tripathi *et al*. 2010). Inhibition of splicing factor recruitment, due to Malat1 downregulation, results in altered expression patterns of alternatively spliced transcripts (Tripathi *et al*. 2010). In cultured hippocampal neurons Malat1 depletion leads to a reduction of synaptic density and significant downregulation of the levels of cell adhesion molecules Nlgn1 and SynCAM1, which mediate the cell-cell

interactions necessary for LTP (Bernard *et al*. 2010). Conversely over expression of Malat1 prompts an increase in synaptic density (Bernard *et al*. 2010). In this study Malat1 is significantly downregulated in the presence of NCAM2-ED however following induction of LTP the expression levels appear to be restored to those observed in the control group (Fig. 11).

Both Neat1 and Malat1 are evolutionarily conserved across the mammalian lineage (Hutchinson *et al*. 2007). No non-mammalian homologs were identified for either Neat1 or Malat1 suggesting that these lincRNAs are specific for the mammalian lineage (Hutchinson *et al*. 2007). Malat1 exhibits unusually high sequence conservation for a non-coding RNA transcript indicating that it is functional (Bernard *et al*. 2010; Hutchinson *et al*. 2007). The primary sequence of Neat1 is not as conserved as that of Malat1 however two regions within the Neat1 sequence exhibit high levels of conservation across the mammalian lineage (Hutchinson *et al*. 2007). Both Neat1 and Malat1 are single exon transcripts. Whilst the two lincRNA transcripts do not share any significant homology both genes are located within close proximity of one another on the mammalian genome and the RNAs are retained in the nucleus following transcription (Hutchinson *et al*. 2007). Neat1 is located on mouse chromosome 19 less than 50kb from the Malat1 locus (Fig. 10a). The human homolog is located on chromosome 11 less than 70kb from Malat1 (Hutchinson *et al*. 2007; Ip and Nakagawa 2012). Neither Neat1 nor Malat1 possess conserved or significant ORFs (Hutchinson *et al*. 2007). Both Neat1 and Malat1 interact with proteins and RNAs at the nuclear speckle to regulate pre-mRNA processing, splicing and export (Hutchinson *et al*. 2007).

Consistent with expression pattern 1, Miat was shown to be significantly downregulated in response to NCAM2-ED treatment however upregulated when LTP is induced in the presence of NCAM2-ED (Fig. 11). Although not statistically significant, Miat exhibited slight downregulation in response to induction of LTP (Fig. 11). Miat (myocardial infarction associated transcript), also known as GOMAFU and RNCR2 (rential non-coding RNA 2), is a tissue-specific lincRNA highly expressed in a subset of neurons including the cerebral cortex, olfactory bulb and the hippocampal CA1 neurons (Barry *et al*. 2014; Ishii *et al*. 2006; Mercer *et al*. 2008b; Rapicavoli *et al*. 2010; Sone *et al*. 2007; Tsuiji *et al*. 2011). Miat is highly expressed in developing foetal mouse brains with expression persisting in a subset of adult neurons (Ishii *et al*. 2006; Sone *et al*. 2007). Localised to discrete regions of the

nucleus, Miat formed structures could not be assigned to any known nuclear subdomains, potentially representing novel nuclear bodies (Ip and Nakagawa 2012; Sone *et al*. 2007). The Miat transcript acts as a scaffold for splicing factor assembly, sequestering splicing factors (SF1, QKI, SRSF1) in resting neurons and regulating their release in response to neuronal activity (Barry *et al*. 2014; Ip and Nakagawa 2012; Ishii *et al*. 2006; Sone *et al*. 2007).

The Miat gene contains seven exons and alternative splicing results in formation of at least ten different isoforms (Sone *et al*. 2007). The Miat transcript is evolutionarily conserved across mammalian, avian and amphibian species suggesting a crucial function for the transcript within the cell (Rapicavoli *et al*. 2010; Tsuiji *et al*. 2011). In primary mouse neurons depolarisation of the cell in response to neural activity results in a downregulation of the Miat transcript, releasing splicing factors to facilitate splicing activity (Barry *et al*. 2014).

Taken together, Neat1, Malat1 and Miat lincRNAs all appear to regulate gene expression by contributing to the structure or function of nuclear bodies, interacting with proteins to alter the splicing and export patterns of specific mRNAs (Ip and Nakagawa 2012). Together these lincRNA transcripts form part of a network responsible for regulating the maturation and expression of selected pre-mRNA transcripts. These three transcripts may be the first of many lncRNAs with important roles in LTP as the structural and functional characterization of lncRNAs has only recently begun (Bernard *et al*. 2010; Earls *et al*. 2014; Knauss and Sun 2013).

4.3.2 Unannotated lincRNAs

In this study, an average of 16187 previously unidentified transcripts were identified for each treatment of which four transcripts were determined to be significantly differentially expressed. Examination of these transcripts revealed that one of these significant previously unannotated transcript, designated with the ID XLOC_013908, was found to be lowly expressed with levels below one FPKM and significant differential expressed was only detected between the LTP and NCAM2-ED treatment groups, a comparison deemed uninformative as it is influenced by two independent variables, potentially confounding the results. As such ID XLOC_013908 was excluded from further characterisation. The remaining three significant unannotated transcripts were characterised as putative lincRNAs as each was (i) longer than 200 nucleotides, (ii) located between annotated protein-coding

genes with no overlap and (iii) bore putative open reading frames (ORF) no larger than 100 amino acids (aa) in length (Hangauer *et al*. 2013). Two distinct isoforms were identified for XLOC_006256, each of which was characterised independently.

The ribosomal binding profile of the XLOC_006440 transcript revealed no ribosomal interactions, indicating that the transcript is not translated. Ribosomal binding was detected for XLOC_013906 transcript and the two XLOC_006256 isoforms, thus suggesting their potential for expression of micropeptides. Further studies utilising mass spectrometry analysis could confirm the presence of micropeptides translated from the XLOC_013906 and XLOC_006256 transcripts.

Expression of both XLOC_013906 and XLOC_006256 was significantly downregulated in the presence of NCAM2-ED but upregulated following induction of LTP in the presence of NCAM2-ED (Fig.13). These findings may indicate potential roles for these transcripts in the mechanisms of synaptic disassembly induced by NCAM2-ED. The proximity of the XLOC_013906 and Kcna4 loci may indicate *cis* interactions between the genes. Whilst total expression of Kcna4 does not correlate with that of XLOC_013906, the expression pattern of the lowly expressed Kcna4 isoform 2 does correlate. Further experimentation is necessary to establish whether these two genes interact.

A significant downregulation of XLOC_006440 was observed in response to both LTP induction and treatment with NCAM2-ED suggesting a role of the transcript in the induction of LTP as well as in the synaptic disruption caused by NCAM2-ED treatment (Fig.18). Induction of LTP combined with NCAM2-ED treatment led to upregulation of XLOC_006440 expression, which although not statistically significant, might be biologically relevant (Fig.18).

4.4 Concordance between rat and mouse transcriptome data

The advancement of high-throughput sequencing technology and increasing cost-effectiveness has led to a massive increase in the amount of RNA-Seq data available in the public domain. Meta-analysis of this RNA-Seq data can therefore provide excellent research opportunities without incurring any cost.

In the current study, comparative analysis identified eight genes significantly differentially expressed in the mouse Control vs. LTP data set which are also expressed in similar study

using hippocampal rat cells (Table 13). However, none of the rat ortholog genes were identified as significantly differentially expressed. Moreover, in most cases they displayed expression patterns opposite to those observed in the mice. While the mouse genes exhibited downregulation in response to LTP induction, the rat orthologs were as mostly upregulated in the same conditions. It is unlikely that these results reflect different mechanisms of LTP induction between mice and rats as the mechanisms of LTP appear to be relatively conserved (Levenson and Sweatt 2006). A number of other variables may account for the observed disparities in gene expression.

Differential expression testing identified 557 genes in the rat data set compared to the 11 differentially expressed genes identified in the mouse data set. This large discrepancy may be due to the less stringent p-value threshold. The mouse data set had a highly stringent p-value of 0.00005 to maintain statistical power in the context of the six pairwise comparisons. By contrast the rat data had only one pairwise comparison allowing a less severe significance threshold of $p<0.00145$, which might have led to a higher number of differentially expressed genes.

Another contributing factor may be the quality of annotation. The rat reference genome is not as extensively annotated as the mouse genome. For example, the NONCODE database holds data for over 130,000 lncRNA transcripts for mice and only 29,070 lncRNA transcripts for rats (Zhao et al. 2015).

The tissue type and manner in which LTP was induced may also have contributed to the observed discrepancies between the mouse and rat data. The mouse neurons used in this study were primary cultures hippocampal neurons and LTP was induced chemically *in vitro*. In the rat experiment, LTP was induced *in vivo* using high frequency stimulation in live animals then hippocampal neurons, isolated from the dentate gyrus and cornu amonis, were rapidly dissected on ice and flash frozen (Maag *et al*. 2015). Further study may be required to elucidate whether the differences in gene expression profiles are the result of different LTP induction mechanisms or the result of confounding variables.

Most notably, Maag et al. (2015) utilised adult rat brains whilst the present study employed neonatal mouse brains thus differences in gene expression may reflect age related differences in LTP and learning. Developmentally the brain of a one day old mouse is equivalent to that of a human foetus in the third trimester of gestation (Clancy et al. 2007). The synapses of juvenile brains are highly plastic with connections being formed and pruned as the animal

develops and learns (Li and Tsien 2009; Murase et al. 2011). Adult brains are less malleable than those of juveniles, potentially employing different mechanisms of synaptic plasticity which may account for the observed transcriptomic discrepancies."

4.5 Future directions

The development of high-throughput sequencing techniques, including RNA-Seq and Ribo-Seq, has facilitated genome-wide characterisation of genes involved in LTP induction, providing insight into the molecular mechanisms involved in learning and memory formation. Functional characterisation of the novel putative lincRNAs identified in this study, as well as micropeptides they might encode, will be required to validate these findings. For example, mass spectrometry analysis should be utilised to confirm the translation of micropeptides from the XLOC_013906 and XLOC_006256 transcripts.

The 30 min post LTP time point utilised in this study corresponds with E-LTP or the induction phase of LTP. Building upon the results of this study, future work might utilise RNA-Seq to perform a time-course analysis of neurons thus tracking transcriptomic changes throughout the early, intermediate and late stages of LTP. A more expansive time-course analysis with RNA isolated at 30 mins, 1 hr, 2 hrs, 4 hrs, 8 hrs and 16 hrs post LTP induction may elucidate the transcriptomic changes occurring during the transition from E-LTP through I-LTP to L-LTP. The present study utilised primary hippocampal neurons isolated from one day old mouse pups and cultured. Inclusion of adult hippocampal neurons in future studies, in addition to the juvenile neurons, may reveal variations in LTP mechanisms between juvenile and mature brains. Due to budgetary constraints, however such extended experimental design was beyond the scope of the present study."

Additionally, utilisation of publicly available RNA-Seq data may provide excellent opportunities for meta-analytical studies. As well as providing greater insight how memory is formed and maintained such studies may reveal how memory becomes impaired as a result of healthy aging and in the pathology of neurodegenerative diseases such as AD.

References

Abbott LF, Nelson SB (2000) Synaptic plasticity: taming the beast. Nature neuroscience 3:1178-1183

Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Eberhard C (2016) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic acids research:gkw343

Alzheimer's A (2009) 2009 Alzheimer's disease facts and figures. Alzheimer's & Dementia 5:234-270

Andreyeva A, Leshchyns' ka I, Knepper M, Betzel C, Redecke L, Sytnyk V, Schachner M (2010) CHL1 is a selective organizer of the presynaptic machinery chaperoning the SNARE complex. PLoS One 5:e12018

Änkö M-L, Neugebauer KM (2010) Long noncoding RNAs add another layer to pre-mRNA splicing regulation. Molecular cell 39:833-834

Arlt A, Schäfer H (2011) Role of the immediate early response 3 (IER3) gene in cellular stress response, inflammation and tumorigenesis. European Journal of Cell Biology 90:545-552

Azevedo FA, Carvalho LR, Grinberg LT, Farfel JM, Ferretti RE, Leite RE, Lent R, Herculano-Houzel S (2009) Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. Journal of Comparative Neurology 513:532 -541

Barry G, Briggs J, Vanichkina D, Poth E, Beveridge N, Ratnu V, Nayler S, Nones K, Hu J, Bredy T (2014) The long non-coding RNA Gomafu is acutely regulated in response to neuronal activation and involved in schizophrenia-associated alternative splicing. Molecular psychiatry 19:486-494

Barry G, Briggs JA, Do Won Hwang SPN, Fortuna PR, Jonkhout N, Dachet F, Maag JL, Mestdagh P, Singh EM, Avesson L (2017) The long non-coding RNA NEAT1 is responsive to neuronal activity and is associated with hyperexcitability states. Scientific Reports 7

Baudry M, Zhu G, Liu Y, Wang Y, Briz V, Bi X (2015) Multiple cellular cascades participate in long-term potentiation and in hippocampus-dependent learning. Brain research 1621:73-81

Bauer S, Kerr BJ, Patterson PH (2007) The neuropoietic cytokine family in development, plasticity, disease and injury. Nature Reviews Neuroscience 8:221-232

Benoit ME, Tenner AJ (2011) Complement protein C1q-mediated neuroprotection is correlated with regulation of neuronal gene and microRNA expression. Journal of Neuroscience 31:3459-3469

Benson DL, Huntley GW (2012) Synapse adhesion: a dynamic equilibrium conferring stability and flexibility. Current Opinion in Neurobiology 22:397-404

Bernard D, Prasanth KV, Tripathi V, Colasse S, Nakamura T, Xuan Z, Zhang MQ, Sedel F, Jourdren L, Coulpier F (2010) A long nuclear-retained non- coding RNA regulates synaptogenesis by modulating gene expression. The EMBO journal 29:3082-3093

Bertone P, Stolc V, Royce TE, Rozowsky JS, Urban AE, Zhu X, Rinn JL, Tongprasit W, Samanta M, Weissman S (2004) Global identification of human transcribed sequences with genome tiling arrays. Science 306:2242-2246

Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Thurman RE (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 447:799-816

Bliim N, Leshchyns'ka I, Sytnyk V, Janitz M (2016) Transcriptional regulation of long-term potentiation. neurogenetics:1-10

Bliss TV, Collingridge GL (1993) A synaptic model of memory: long-term potentiation in the hippocampus. Nature 361:31-39

Bliss TV, Lømo T (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. The Journal of physiology 232:331-356

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics:btu170

Bolliger MF, Martinelli DC, Südhof TC (2011) The cell-adhesion G protein-coupled receptor BAI3 is a high-affinity receptor for C1q-like proteins. Proceedings of the National Academy of Sciences 108:2534-2539

Boulanger LM (2009) Immune proteins in brain development and synaptic plasticity. Neuron 64:93-109

Brannan CI, Dees EC, Ingram RS, Tilghman SM (1990) The product of the H19 gene may function as an RNA. Molecular and cellular biology 10:28-36
Briggs JA, Wolvetang EJ, Mattick JS, Rinn JL, Barry G (2015) Mechanisms of long non-coding RNAs in mammalian nervous system development, plasticity, disease, and evolution. Neuron 88:861-877
Brookmeyer R, Johnson E, Ziegler-Graham K, Arrighi HM (2007) Forecasting the global burden of Alzheimer's disease. Alzheimer's & Dementia 3:186-191
Brown CJ, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R, Willard HF (1991) A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature 349:38-44
Cahoy JD, Emery B, Kaushal A, Foo LC, Zamanian JL, Christopherson KS, Xing Y, Lubischer JL, Krieg PA, Krupenko SA (2008) A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. Journal of Neuroscience 28:264-278
Carninci P, Kasukawa T, Katayama S, Gough J, Frith M, Maeda N, Oyama R, Ravasi T, Lenhard B, Wells C (2005) The transcriptional landscape of the mammalian genome. Science 309:1559-1563
Chaisson EJ (2002) Cosmic evolution: the rise of complexity in nature. Harvard University Press
Chambon P (1996) A decade of molecular biology of retinoic acid receptors. The FASEB Journal 10:940-954
Chatzistamou I, Schally AV, Sun B, Armatis P, Szepeshazi K (2000) Inhibition of growth of OV-1063 human epithelial ovarian cancers and c-jun and c-fos oncogene expression by bombesin antagonists. British journal of cancer 83:906
Chen L-L, Carmichael GG (2009) Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. Molecular cell 35:467-478
Chen P-W, Kroog GS (2004) Alterations in receptor expression or agonist concentration change the pathways gastrin-releasing peptide receptor uses to regulate extracellular signal-regulated kinase. Molecular pharmacology 66:1625-1634
Chiang M-Y, Misner D, Kempermann G, Schikorski T, Giguère V, Sucov HM, Gage FH, Stevens CF, Evans RM (1998) An essential role for retinoid receptors RARβ and RXRγ in long-term potentiation and depression. Neuron 21:1353-1361
Chinwalla AT, Cook LL, Delehaunty KD, Fewell GA, Fulton LA, Fulton RS, Graves TA, Hillier LW, Mardis ER, McPherson JD (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420:520-562
Cho J, Yu N-K, Choi J-H, Sim S-E, Kang SJ, Kwak C, Lee S-W, Kim J-i, Choi DI, Kim VN (2015) Multiple repressive mechanisms in the hippocampus during memory formation. Science 350:82-87
Chua JJE, Kindler S, Boyken J, Jahn R (2010) The architecture of an excitatory synapse. Journal of Cell Science 123:819-823
Clancy B, Kersh B, Hyde J, Darlington RB, Anand K, Finlay BL (2007) Web-based method for translating neurodevelopment from laboratory species to humans. Neuroinformatics 5:79-94
Clark MB, Mattick JS (2011) Long noncoding RNAs in cell biology. Seminars in cell & developmental biology. Elsevier, pp 366-376
Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, Lawrence JB (2009) An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. Molecular cell 33:717-726
Cole DS, Morgan BP (2003) Beyond lysis: how complement influences cell fate. Clinical Science 104:455-466
Coleman PD, Yao PJ (2003) Synaptic slaughter in Alzheimer's disease. Neurobiology of Aging 24:1023-1027
Consortium EP (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489:57-74
Consortium S (1998) {Genome sequence of the nematode $ C. elegans $: A platform for investigating biology}. Science 282:2012-2018
Costa-Mattioli M, Sossin WS, Klann E, Sonenberg N (2009) Translational Control of Long-Lasting Synaptic Plasticity and Memory. Neuron 61:10-26

Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG (2012) The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome research 22:1775-1789
Donzis EJ, Tronson NC (2014) Modulation of learning and memory by cytokines: signaling mechanisms and long term consequences. Neurobiology of learning and memory 115:68-77
Earls LR, Westmoreland JJ, Zakharenko SS (2014) Non-coding RNA regulation of synaptic plasticity and memory: Implications for aging. Ageing research reviews 17:34-42
Edelman GM, Crossin KL (1991) Cell adhesion molecules: implications for a molecular histology. Annual review of biochemistry 60:155-190
English JD, Sweatt JD (1996) Activation of p42 mitogen-activated protein kinase in hippocampal long term potentiation. Journal of Biological Chemistry 271:24329-24332
Fortin NJ, Agster KL, Eichenbaum HB (2002) Critical role of the hippocampus in memory for sequences of events. Nature neuroscience 5:458-462
Fox AH, Bond CS, Lamond AI (2005) P54nrb forms a heterodimer with PSP1 that localizes to paraspeckles in an RNA-dependent manner. Molecular biology of the cell 16:5304-5315
Furuta A, Wada E, Wada K (2007) Function of glial G-protein coupled receptors. Brain and nerve= Shinkei kenkyu no shinpo 59:717-724
Garay PA, McAllister AK (2010) Novel roles for immune molecules in neural development: implications for neurodevelopmental disorders. Frontiers in synaptic neuroscience 2:136
Garcia J, Ye Y, Arranz V, Letourneux C, Pezeron G, Porteu F (2002) IEX-1: a new ERK substrate involved in both ERK survival activity and ERK activation. The EMBO journal 21:5151-5163
Goodrich JA, Kugel JF (2006) Non-coding-RNA regulators of RNA polymerase II transcription. Nature Reviews Molecular Cell Biology 7:612-616
Gruart A, Muñoz MD, Delgado-García JM (2006) Involvement of the CA3–CA1 synapse in the acquisition of associative learning in behaving mice. The Journal of neuroscience 26:1077-1087
Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL (2008) The vienna RNA websuite. Nucleic acids research 36:W70-W74
Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458:223-227
Hamlin JA, Fang H, Schwob JE (2004) Differential expression of the mammalian homologue of fasciclin II during olfactory development in vivo and in vitro. The Journal of Comparative Neurology 474:438-452
Han M-R, Schellenberg GD, Wang L-S (2010) Genome-wide association reveals genetic effects on human Aβ 42 and τ protein levels in cerebrospinal fluids: a case control study. BMC Neurology 10:1-14
Hangauer MJ, Vaughn IW, McManus MT (2013) Pervasive Transcription of the Human Genome Produces Thousands of Previously Unidentified Long Intergenic Noncoding RNAs. PLoS Genet 9:e1003569
Hardy J, Selkoe DJ (2002) The Amyloid Hypothesis of Alzheimer's Disease: Progress and Problems on the Road to Therapeutics. Science 297:353-356
Harrow J (2016) Identifying and Cataloguing LncRNAs in Human and Mouse. Plant and Animal Genome XXIV Conference. Plant and Animal Genome
Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S (2012) GENCODE: the reference human genome annotation for The ENCODE Project. Genome research 22:1760-1774
Hebb DO (2005) The organization of behavior: A neuropsychological theory. Psychology Press
Hellmich MR, Ives KL, Udupi V, Soloff MS, Greeley GH, Christensen BN, Townsend CM (1999) Multiple protein kinase pathways are involved in gastrin-releasing peptide receptor-regulated secretion. Journal of Biological Chemistry 274:23901-23909
Hermey G, Mahlke C, Gutzmann JJ, Schreiber J, Bluthgen N, Kuhl D (2013) Genome-wide profiling of the activity-dependent hippocampal transcriptome. PloS one 8
Herring BE, Nicoll RA (2016) Long-Term Potentiation: From CaMKII to AMPA Receptor Trafficking. Annual Review of Physiology 78

Hong SJ, Li H, Becker KG, Dawson VL, Dawson TM (2004) Identification and analysis of plasticity-induced late-response genes. Proceedings of the National Academy of Sciences of the United States of America 101:2145-2150
Huang DW, Sherman BT, Lempicki RA (2008) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protocols 4:44-57
Huang DW, Sherman BT, Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research 37:1-13
Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB, Chess A (2007) A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. BMC genomics 8:1
Ichinohe N, Yoshihara Y, Hashikawa T, Rockland KS (2003) Developmental study of dendritic bundles in layer 1 of the rat granular retrosplenial cortex with special reference to a cell adhesion molecule, OCAM. European Journal of Neuroscience 18:1764-1774
Ip JY, Nakagawa S (2012) Long non-coding RNAs in nuclear bodies. Development, growth & differentiation 54:44-54
Ishii N, Ozaki K, Sato H, Mizuno H, Saito S, Takahashi A, Miyamoto Y, Ikegawa S, Kamatani N, Hori M (2006) Identification of a novel non-coding RNA, MIAT, that confers risk of myocardial infarction. Journal of human genetics 51:1087-1099
Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, Barrette TR, Prensner JR, Evans JR, Zhao S, Poliakov A, Cao X, Dhanasekaran SM, Wu Y-M, Robinson DR, Beer DG, Feng FY, Iyer HK, Chinnaiyan AM (2015) The landscape of long noncoding RNAs in the human transcriptome. Nat Genet 47:199-208
Johnson JM, Edwards S, Shoemaker D, Schadt EE (2005) Dark matter in the genome: evidence of widespread transcription detected by microarray tiling experiments. TRENDS in Genetics 21:93-102
Kadakkuzha BM, Liu X-A, McCrate J, Shankar G, Rizzo V, Afinogenova A, Young B, Fallahi M, Carvalloza AC, Raveendra B (2015) Transcriptome analyses of adult mouse brain reveal enrichment of lncRNAs in specific brain regions and neuronal populations. Frontiers in cellular neuroscience 9
Kamichi S, Wada E, Aoki S, Sekiguchi M, Kimura I, Wada K (2005) Immunohistochemical localization of gastrin-releasing peptide receptor in the mouse brain. Brain research 1032:162-170
Kandel ER (2012) The molecular biology of memory: cAMP, PKA, CRE, CREB-1, CREB-2, and CPEB. Mol Brain 5:14
Kapranov P, Drenkow J, Cheng J, Long J, Helt G, Dike S, Gingeras TR (2005) Examples of the complex architecture of the human transcriptome revealed by RACE and high-density tiling arrays. Genome research 15:987-997
Karttunen K, Karppi P, Hiltunen A, Vanhanen M, Välimäki T, Martikainen J, Valtonen H, Sivenius J, Soininen H, Hartikainen S (2011) Neuropsychiatric symptoms and quality of life in patients with very mild and mild Alzheimer's disease. International journal of geriatric psychiatry 26:473-482
Kelleher RJ, Govindarajan A, Tonegawa S (2004) Translational regulatory mechanisms in persistent forms of synaptic plasticity. Neuron 44:59-73
Kim T-K, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S (2010) Widespread transcription at neuronal activity-regulated enhancers. Nature 465:182-187
Knauss JL, Sun T (2013) Regulatory mechanisms of long noncoding RNAs in vertebrate central nervous system development and function. Neuroscience 235:200-214
Kovács KJ (1998) Invited review c-Fos as a transcription factor: a stressful (re) view from a functional map. Neurochemistry international 33:287-297
Kozlova N, Braga J, Lundgren J, Rino J, Young P, Carmo-Fonseca M, Visa N (2006) Studies on the role of NonA in mRNA biogenesis. Experimental cell research 312:2619-2630
Kulahin N, Walmod PS (2010) The Neural Cell Adhesion Molecule NCAM2/OCAM/RNCAM, a Close Relative to NCAM. In: Berezin V (ed) Structure and Function of the Neural Cell Adhesion Molecule NCAM. Springer New York, New York, NY, pp 403-420
Labes M, Roder J, Roach A (1998) A novel phosphatase regulating neurite extension on CNS inhibitors. Molecular and Cellular Neuroscience 12:29-47

Lamond AI, Spector DL (2003) Nuclear speckles: a model for nuclear organelles. Nature Reviews Molecular Cell Biology 4:605-612
Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W (2001) Initial sequencing and analysis of the human genome. Nature 409:860-921
Lee JT (2012) Epigenetic Regulation by Long Noncoding RNAs. Science 338:1435-1439
Lee K, Dixon A, Gonzalez I, Stevens E, McNulty S, Oles R, Richardson P, Pinnock R, Singh L (1999) Bombesin-like peptides depolarize rat hippocampal interneurones through interaction with subtype 2 bombesin receptors. The Journal of physiology 518:791-802
Lépée-Lorgeoux I, Betancur C, Rostène W, Pélaprat D (1999) Differential ontogenetic patterns of levocabastine-sensitive neurotensin NT 2 receptors and of NT 1 receptors in the rat brain revealed by in situ hybridization. Developmental brain research 113:115-131
Leshchyns'ka I, Liew HT, Shepherd C, Halliday GM, Stevens CH, Ke YD, Ittner LM, Sytnyk V (2015) A [beta]-dependent reduction of NCAM2-mediated synaptic adhesion contributes to synapse loss in Alzheimer/'s disease. Nature communications 6
Levenson J, Sweatt J (2006) Memory. Cellular and Molecular Life Sciences 63:1009-1016
Li F, Tsien JZ (2009) Memory and the NMDA receptors. The New England journal of medicine 361:302
Lin C-Y, Kleinbrink EL, Dachet F, Cai J, Ju D, Goldstone A, Wood EJ, Liu K, Jia H, Goustin A-S, Kosir MA, Thepsuwan P, Lipovich L (2016) Primate-specific oestrogen-responsive long non-coding RNAs regulate proliferation and viability of human breast cancer cells. Open Biology 6
Lin S, Fu X-D (2006) SR proteins and related factors in alternative splicing. Advances in experimental medicine and biology 623:107-122
Lipovich L, Dachet F, Cai J, Bagla S, Balan K, Jia H, Loeb JA (2012) Activity-dependent human brain coding/noncoding gene regulatory networks. Genetics 192:1133-1148
Long JC, Caceres JF (2009) The SR protein family of splicing factors: master regulators of gene expression. Biochemical Journal 417:15-27
Lu W-Y, Man H-Y, Ju W, Trimble WS, MacDonald JF, Wang YT (2001) Activation of synaptic NMDA receptors induces membrane insertion of new AMPA receptors and LTP in cultured hippocampal neurons. Neuron 29:243-254
Lucin KM, Wyss-Coray T (2009) Immune activation in brain aging and neurodegeneration: too much or too little? Neuron 64:110-122
Maag JL, Panja D, Sporild I, Patil S, Kaczorowski DC, Bramham CR, Dinger ME, Wibrand K (2015) Dynamic expression of long noncoding RNAs and repeat elements in synaptic plasticity. Frontiers in neuroscience 9
Maeno H, Yamada K, Santo-Yamada Y, Aoki K, Sun Y-J, Sato E, Fukushima T, Ogura H, Araki T, Kamichi S (2004) Comparison of mice deficient in the high-or low-affinity neurotensin receptors, Ntsr1 or Ntsr2, reveals a novel function for Ntsr2 in thermal nociception. Brain research 998:122-129
Malenka RC, Nicoll RA (1999) Long-term potentiation--a decade of progress? Science 285:1870-1874
Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schütz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P (1995) The nuclear receptor superfamily: the second decade. Cell 83:835-839
Markram H (2012) The human brain project. Scientific American 306:50-55
Martin S, Grimwood P, Morris R (2000) Synaptic plasticity and memory: an evaluation of the hypothesis. Annual review of neuroscience 23:649-711
Masters CL, Simms G, Weinman NA, Multhaup G, McDonald BL, Beyreuther K (1985) Amyloid plaque core protein in Alzheimer disease and Down syndrome. Proceedings of the National Academy of Sciences of the United States of America 82:4245-4249
Matthies H, Frey U, Reymann K, Krug M, Jork R, Schroeder H (1990) Different Mechanisms and Multiple Stages of LTP. In: Ben-Ari Y (ed) Excitatory Amino Acids and Neuronal Plasticity. Springer US, pp 359-368
Mattick JS (2001) Non-coding RNAs: the architects of eukaryotic complexity. EMBO reports 2:986 - 991
Mattick JS (2007) A new paradigm for developmental biology. Journal of Experimental Biology 210:1526-1547

Mattick JS (2011) The central role of RNA in human development and cognition. FEBS Letters 585:1600-1616
Mazella J, Béraud-Dufour S, Devader C, Massa F, Coppola T (2012) Neurotensin and its receptors in the control of glucose homeostasis
Mazella J, Botto J-M, Guillemare E, Coppola T, Sarret P, Vincent J-P (1996) Structure, functional expression, and cerebral localization of the levocabastine-sensitive neurotensin/neuromedin N receptor from mouse brain. Journal of Neuroscience 16:5613-5620
Mazella J, Vincent J-P (2006) Functional roles of the NTS2 and NTS3 receptors. Peptides 27:2469-2475
McHail DG, Dumas TC (2015) Multiple forms of metaplasticity at a single hippocampal synapse during late postnatal development. Developmental cognitive neuroscience 12:145-154
Mehler MF, Mattick JS (2006) Non-co ding RNAs in the nervous system. The Journal of physiology 575:333-341
Mercer TR, Dinger ME, Mariani J, Kosik KS, Mehler MF, Mattick JS (2008a) Noncoding RNAs in long-term memory formation. The Neuroscientist 14:434-445
Mercer TR, Dinger ME, Sunkin SM, Mehler MF, Mattick JS (2008b) Specific expression of long noncoding RNAs in the mouse brain. Proceedings of the National Academy of Sciences 105:716-721
Merico D, Isserlin R, Bader GD (2011) Visualizing gene-set enrichment results using the Cytoscape plug-in enrichment map. Network Biology: Methods and Applications:257-277
Michel AM, Fox G, Kiran AM, De Bo C, O'Connor PB, Heaphy SM, Mullan JP, Donohue CA, Higgins DG, Baranov PV (2013) GWIPS-viz: development of a ribo-seq genome browser. Nucleic acids research:gkt1035
Mills JD, Kavanagh T, Kim WS, Chen BJ, Waters PD, Halliday GM, 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
Molyneaux BJ, Goff LA, Brettler AC, Chen H-H, Brown JR, Hrvatin S, Rinn JL, Arlotta P (2015) DeCoN: genome-wide analysis of in vivo transcriptional dynamics during pyramidal neuron fate selection in neocortex. Neuron 85:275-288
Mucke L (2009) Neuroscience: Alzheimer's disease. Nature 461:895-897
Murase S, Owens DF, McKay RD (2011) In the newborn hippocampus, neurotrophin-dependent survival requires spontaneous activity and integrin signaling. Journal of Neuroscience 31:7791-7800
Necsulea A, Soumillon M, Warnefors M, Liechti A, Daish T, Zeller U, Baker JC, Grützner F, Kaessmann H (2014) The evolution of lncRNA repertoires and expression patterns in tetrapods. Nature 505:635-640
Neves G, Cooke SF, Bliss TVP (2008) Synaptic plasticity, memory and the hippocampus: a neural network approach to causality. Nat Rev Neurosci 9:65-75
Noctor SC, Martínez-Cerdeño V, Kriegstein AR (2007) Contribution of intermediate progenitor cells to cortical histogenesis. Archives of Neurology 64:639-642
Nomoto M, Takeda Y, Uchida S, Mitsuda K, Enomoto H, Saito K, Choi T, Watabe AM, Kobayashi S, Masushige S (2012) Dysfunction of the RAR/RXR signaling pathway in the forebrain impairs hippocampal memory and synaptic plasticity. Molecular brain 5:8
Pang KC, Frith MC, Mattick JS (2006) Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. Trends in Genetics 22:1-5
Park CS, Gong R, Stuart J, Tang S-J (2006) Molecular network and chromosomal clustering of genes involved in synaptic plasticity in the hippocampus. Journal of Biological Chemistry 281:30195-30211
Petit F, Plessis G, Decamp M, Cuisset J-M, Blyth M, Pendlebury M, Andrieux J (2015) 21q21 deletion involving NCAM2: Report of 3 cases with neurodevelopmental disorders. European Journal of Medical Genetics 58:44-46
Pisalyaput K, Tenner AJ (2008) Complement component C1q inhibits β-amyloid- and serum amyloid P-induced neurotoxicity via caspase- and calpain- independent mechanisms. Journal of neurochemistry 104:696-707
Ponjavic J, Ponting CP, Lunter G (2007) Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. Genome research 17:556-565

Pousset F (1994) Developmental expression of cytokine genes in the cortex and hippocampus of the rat central nervous system. Developmental brain research 81:143-146
Puchkov D, Leshchyns' ka I, Nikonenko AG, Schachner M, Sytnyk V (2011) NCAM/spectrin complex disassembly results in PSD perforation and postsynaptic endocytic zone formation. Cerebral Cortex:bhq283
Qu Z, Adelson DL (2012a) Evolutionary conservation and functional roles of ncRNA. Frontiers in Genetics 3:205
Qu Z, Adelson DL (2012b) Identification and Comparative Analysis of ncRNAs in Human, Mouse and Zebrafish Indicate a Conserved Role in Regulation of Genes Expressed in Brain. PLoS ONE 7:e52275
Rapicavoli NA, Poth EM, Blackshaw S (2010) The long noncoding RNA RNCR2 directs mouse retinal cell specification. BMC developmental biology 10:1
Raymond CR (2007) LTP forms 1, 2 and 3: different mechanisms for the 'long'in long-term potentiation. Trends in neurosciences 30:167-175
Ricklin D, Hajishengallis G, Yang K, Lambris JD (2010) Complement: a key system for immune surveillance and homeostasis. Nature immunology 11:785-797
Ringman JM, Schulman H, Becker C, Jones T, Bai Y, Immermann F, Cole G, Sokolow S, Gylys K, Geschwind DH (2012) Proteomic changes in cerebrospinal fluid of presymptomatic and affected persons carrying familial Alzheimer disease mutations. Archives of neurology 69:96-104
Rinn JL, Chang HY (2012) Genome regulation by long noncoding RNAs. Annual review of biochemistry 81:10.1146/annurev-biochem-051410-092902
Roesler R, Luft T, Oliveira S, Farias C, Almeida V, Quevedo J, Dal Pizzol F, Schröder N, Izquierdo I, Schwartsmann G (2006) Molecular mechanisms mediating gastrin-releasing peptide receptor modulation of memory consolidation in the hippocampus. Neuropharmacology 51:350-357
Roesler R, Schwartsmann G (2012) Gastrin-releasing peptide receptors in the central nervous system: role in brain function and as a drug target
Rønn LCB, Berezin V, Bock E (2000) The neural cell adhesion molecule in synaptic plasticity and ageing. International Journal of Developmental Neuroscience 18:193-199
Rubin TG, Gray JD, McEwen BS (2014) Experience and the ever-changing brain: What the transcriptome can reveal. BioEssays 36:1072-1081
Ryan B, Joilin G, Williams JM (2015) Plasticity-related microRNA and their potential contribution to the maintenance of long-term potentiation. Frontiers in molecular neuroscience 8
Saunders NR, Dreifuss J-J, Dziegielewska KM, Johansson PA, Habgood MD, Møllgård K, Bauer H-C (2014) The rights and wrongs of blood-brain barrier permeability studies: a walk through 100 years of history. Frontiers in neuroscience 8
Scadden D (2009) A NEAT way of regulating nuclear export of mRNAs. Molecular cell 35:395-396
Scheff SW, Price DA (2003) Synaptic pathology in Alzheimer's disease: a review of ultrastructural studies. Neurobiology of Aging 24:1029-1046
Schneider H, Pitossi F, Balschun D, Wagner A, del Rey A, Besedovsky HO (1998) A neuromodulatory role of interleukin-1β in the hippocampus. Proceedings of the National Academy of Sciences 95:7778-7783
Sequencing ProjectInternational Rice G (2005) The map-based sequence of the rice genome. Nature 436:793-800
Sheng L, Leshchyns' ka I, Sytnyk V (2015) Neural cell adhesion molecule 2 promotes the formation of filopodia and neurite branching by inducing submembrane increases in Ca2+ levels. The Journal of Neuroscience 35:1739-1752
Shumyatsky GP, Tsvetkov E, Malleret G, Vronskaya S, Hatton M, Hampton L, Battey JF, Dulac C, Kandel ER, Bolshakov VY (2002) Identification of a signaling network in lateral nucleus of amygdala important for inhibiting memory specifically related to learned fear. Cell 111:905-918
Sim RB, Kishore U, Villiers CL, Marche PN, Mitchell DA (2007) C1q binding and complement activation by prions and amyloids. Immunobiology 212:355-362
Smith MA, Gesell T, Stadler PF, Mattick JS (2013) Widespread purifying selection on RNA structure in mammals. Nucleic Acids Research 41:8220-8236

Sone M, Hayashi T, Tarui H, Agata K, Takeichi M, Nakagawa S (2007) The mRNA-like noncoding RNA Gomafu constitutes a novel nuclear domain in a subset of neurons. J Cell Sci 120:2498-2506
Stephan AH, Barres BA, Stevens B (2012) The complement system: an unexpected role in synaptic pruning during development and disease. Annual review of neuroscience 35:369-389
Stevens B, Allen NJ, Vazquez LE, Howell GR, Christopherson KS, Nouri N, Micheva KD, Mehalow AK, Huberman AD, Stafford B (2007) The classical complement cascade mediates CNS synapse elimination. Cell 131:1164-1178
Stough S, Shobe JL, Carew TJ (2006) Intermediate-term processes in memory formation. Current Opinion in Neurobiology 16:672-678
Sun Y-J, Maeno H, Aoki S, Wada K (2001) Mouse neurotensin receptor 2 gene (Ntsr2): genomic organization, transcriptional regulation and genetic mapping on chromosome 12. Molecular brain research 95:167-171
Sweatt JD (2016) Neural plasticity and behavior–sixty years of conceptual advances. Journal of neurochemistry
Sytnyk V, Leshchyns'ka I, Delling M, Dityateva G, Dityatev A, Schachner M (2002) Neural cell adhesion molecule promotes accumulation of TGN organelles at sites of neuron-to-neuron contacts. The Journal of Cell Biology 159:649-661
Tacnet-Delorme P, Chevallier S, Arlaud GJ (2001) β-amyloid fibrils activate the C1 complex of complement under physiological conditions: evidence for a binding site for Aβ on the C1q globular regions. The Journal of Immunology 167:6374-6381
Taft RJ, Pheasant M, Mattick JS (2007) The relationship between non-protein-coding DNA and eukaryotic complexity. BioEssays 29:288-299
Tano K, Mizuno R, Okada T, Rakwal R, Shibato J, Masuo Y, Ijiri K, Akimitsu N (2010) MALAT-1 enhances cell motility of lung adenocarcinoma cells by influencing the expression of motility-related genes. FEBS letters 584:4575-4580
Toga AW, Clark KA, Thompson PM, Shattuck DW, Van Horn JD (2012) Mapping the human connectome. Neurosurgery 71:1
Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25:1105-1111
Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protocols 7:562-578
Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L (2010) Transcript assembly and abundance estimation from RNA-Seq reveals thousands of new transcripts and switching among isoforms. Nature biotechnology 28:511
Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, Freier SM, Bennett CF, Sharma A, Bubulya PA (2010) The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Molecular cell 39:925-938
Tsuiji H, Yoshimoto R, Hasegawa Y, Furuno M, Yoshida M, Nakagawa S (2011) Competition between a noncoding exon and introns: Gomafu contains tandem UACUAAC repeats and associates with splicing factor-1. Genes to Cells 16:479-490
Veerhuis R, Nielsen HM, Tenner AJ (2011) Complement in the brain. Molecular immunology 48:1592-1603
Walz A, Mombaerts P, Greer CA, Treloar HB (2006) Disrupted compartmental organization of axons and dendrites within olfactory glomeruli of mice deficient in the olfactory cell adhesion molecule, OCAM. Molecular and Cellular Neuroscience 32:1-14
Wan Y, Qu K, Zhang QC, Flynn RA, Manor O, Ouyang Z, Zhang J, Spitale RC, Snyder MP, Segal E (2014) Landscape and variation of RNA secondary structure across the human transcriptome. Nature 505:706-709
Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics 10:57-63
Washietl S, Kellis M, Garber M (2014) Evolutionary dynamics and tissue specificity of human long noncoding RNAs in six mammals. Genome research 24:616-628

Whitlock JR, Heynen AJ, Shuler MG, Bear MF (2006) Learning induces long-term potentiation in the hippocampus. science 313:1093-1097
Winther M, Berezin V, Walmod PS (2012) NCAM2/OCAM/RNCAM: Cell adhesion molecule with a role in neuronal compartmentalization. The International Journal of Biochemistry & Cell Biology 44:441-446
Wolf SS, Moody TW (1985) Receptors for GRP/bombesin-like peptides in the rat forebrain. Peptides 6:111-114
Wu MX, Ao Z, Prasad K, Wu R, Schlossman SF (1998) IEX-1L, an apoptosis inhibitor involved in NF-κB-mediated cell survival. Science 281:998-1001
Xie SQ, Martin S, Guillot PV, Bentley DL, Pombo A (2006) Splicing speckles are not reservoirs of RNA polymerase II, but contain an inactive form, phosphorylated on serine2 residues of the C-terminal domain. Molecular biology of the cell 17:1723-1733
Yamauchi R, Wada E, Kamichi S, Yamada D, Maeno H, Delawary M, Nakazawa T, Yamamoto T, Wada K (2007) Neurotensin type 2 receptor is involved in fear memory in mice. Journal of neurochemistry 102:1669-1676
Yasojima K, Schwab C, McGeer EG, McGeer PL (1999) Up-regulated production and activation of the complement system in Alzheimer's disease brain. The American journal of pathology 154:927-936
Yirmiya R, Goshen I (2011) Immune modulation of learning, memory, neural plasticity and neurogenesis. Brain, behavior, and immunity 25:181-213
Yoshihara Y, Kawasaki M, Tamada A, Fujita H, Hayashi H, Kagamiyama H, Mori K (1997) OCAM: A New Member of the Neural Cell Adhesion Molecule Family Related to Zone-to-Zone Projection of Olfactory and Vomeronasal Axons. The Journal of Neuroscience 17:5830-5842
Zhao Y, Li H, Fang S, Kang Y, Hao Y, Li Z, Bu D, Sun N, Zhang MQ, Chen R (2015) NONCODE 2016: an informative and valuable data source of long non-coding RNAs. Nucleic acids research:gkv1252
Zhao Y, Yuan J, Chen R (2016) NONCODEv4: Annotation of Noncoding RNAs with Emphasis on Long Noncoding RNAs. In: Feng Y, Zhang L (eds) Long Non-Coding RNAs: Methods and Protocols. Springer New York, New York, NY, pp 243-254
Zuker M, Stiegler P (1981) Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic acids research 9:133-148